

# Our Moral Fate

Allen Buchanan

© 2020 The Massachusetts Institute of Technology

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

Library of Congress Cataloging-in-Publication Data

Names: Buchanan, Allen E., 1948– author.

Title: Our moral fate : evolution and the escape from tribalism / Allen Buchanan, The MIT Press.

Description: Cambridge, Massachusetts : The MIT Press, 2020. | Includes bibliographical references and index.

Identifiers: LCCN 2019025836 | ISBN 9780262043748 (hardcover) | ISBN 9780262357876 (ebook)

Subjects: LCSH: Ethics—History. | Progress—Moral and ethical aspects.

Classification: LCC BJ71 .B83 2020 | DDC 170—dc23

LC record available at <https://lcn.loc.gov/2019025836>

---

## The First Piece of the Puzzle: Surplus Reproductive Success

In this chapter, I begin to develop a solution to the Big Puzzle; in the next, I fill in details of the solution. The solution won't be simple. It won't rely on one-liner, magic-bullet answers like the thesis that the circle of moral regard expanded because the circle of cooperation expanded or that stable or efficient cooperation requires that people be treated equally or that it was just a matter of humans somehow finally learning how to engage in moral consistency reasoning in the right way and follow through on the conclusions it yields. Those charmingly simple answers don't work, as I argued in chapter 3. Nor will I attempt to show that becoming deeply inclusive made moralities more conducive to reproductive fitness.

My proposed solution—to be honest, my protosolution—will involve evolutionary thinking, but it may not be wholly accurate to call it an evolutionary solution to the Big Puzzle. It will have significant evolutionary components, but it won't be an evolutionary explanation through and through. Given that it's going to be complicated, perhaps the best way to proceed is to simply state, in a succinct way, the main idea and then flesh it out.

### The Key to Solving the Big Puzzle: A Story of Cultural Evolution

Here's the main idea: *if humans have a biologically evolved moral mind that includes an adaptive plasticity that allows them to respond either tribalistically or inclusively to strangers, and if their tendency to respond tribalistically is triggered by threat cues that mimic the harsher conditions of the EEA, then they can avoid tribalistic moralities to the extent that*

*they use their cultural niche construction abilities to create environments in which those threat cues are absent or significantly mitigated.* In even simpler terms: *we aren't condemned to tribalistic moralities—we have the potential to develop inclusive moralities—because the character of our morality is environmentally conditioned and we can shape our environment.* With that simple idea in mind, here are some of the key elements needed to flesh it out. I'll present them in the form of a historical narrative, based on the results of my investigation in this book so far and my reading of relatively uncontroversial historical, sociological, archaeological, and anthropological primary sources.

1. Even if the EEA was often hostile toward peaceful cooperation with members of other groups—even if, consequently, moralities *in that early environment* were largely tribalistic in character—it doesn't follow that human beings are only capable of exclusive moralities. That would follow only if basic human moral psychology, the moral mind, is rigid rather than flexible, so far as the ability to respond to strangers is concerned.
2. We *know* that the moral mind is flexible regarding responses to strangers, because we observe not only tribalistic moralities but also moralities that are much more inclusive. The Two Great Expansions are a fact.
3. The best explanation of why the Two Great Expansions have occurred includes the idea that the moral mind evolved to feature a special adaptive plasticity: the potential both for tribalistic moral responses and for inclusive ones, depending on the environment. In the EEA, flexibility in responding to strangers would have enhanced reproductive fitness in locales that afforded opportunities for peaceful cooperation. For example, success in competing with other groups could be achieved by welcoming strangers rather than fending them off, responding inclusively rather than tribalistically. Or if, due to a dryer or colder local environment, the threat of foreign parasites was low and if resources were relatively abundant, the risks of trying to cooperate with another group could be acceptable.
4. For the special adaptive plasticity to “toggle” toward inclusion, it must be possible to discern that another human being is the sort of

creature with whom it is safe and beneficial to cooperate on moral terms, even though he or she doesn't resemble you in the way fellow members of your ethnic or cultural group do. This amounts to acknowledging a limited sort of moral status: in other words, you will proceed on the assumption of mutual expectations of basic forms of reciprocity—for example, you will believe that the Other is both capable of keeping promises and entitled to your keeping them. But it doesn't follow that you will only cooperate morally with beings you regard as your full moral equals or that you will recognize as equals people with whom you never expect to cooperate. Intergroup cooperation that only involves trade, out-mating, or military alliances requires only *shallowly inclusive* moralities—limited moral regard for the Other—not the deeply inclusive moralities evidenced by the Two Great Expansions. In fact, when humans in the Neolithic Revolution made the transition from small hunter-gatherer groups to large, complex, multiethnic, multicultural societies based on fixed, year-round abodes relying on agriculture and domestication of animals, they developed highly inequalitarian cooperative schemes that assigned a range of different statuses to various groups of individuals who were incorporated within that more complex kind of society. People learned how to cooperate with people quite unlike themselves, but to do so in ways that were characterized by extreme inequalities. In other words, the moral scaffolding of this form of cooperation did not include the recognition of equal basic moral status. Nor did the new, more complex, large-scale forms of morally structured cooperation require or even encourage the Second Great Expansion. On the contrary, they presented new opportunities for the unrestrained exploitation of many more nonhuman animals than ever before.

5. As human societies became larger and more complex, incorporating) more ethnically and culturally heterogeneous groups, there were selective pressures both for reducing the at-a-glance differences that earlier peoples had used to distinguish Us from Them, and for sophisticated discernment capacities that allowed people to see through the remaining superficial differences and judge that very different-looking folks were still the kinds of beings we could cooperate with. Thus, for example, the expansion of trade in the

Chinese, Egyptian, Minoan, Phoenician, and other early empires probably resulted both in the spread of some commonalities in dress, hairstyles, and so on, and also, at least in some people (merchants in particular), the ability to not be put off by the remaining differences when there were opportunities for beneficial exchanges. People got better at inferring the intentions and above all the trustworthiness (or lack of it) in a wider range of human types. Yet all of this was still compatible with inclusion not reaching all the way to the First Great Expansion. The post-Neolithic world of greatly increased intergroup cooperation was a world of deep inequalities, including the ubiquity of slavery. And the transition to it did not bring about the Second Great Expansion. This more complex world hummed along rather smoothly for millenia, even though it probably never occurred to most people either that all human beings possessed equal basic moral status or that nonhuman animals had any moral standing at all. At most, this new environment moved increasing numbers of people from exclusive moralities to shallowly inclusive moralities.

6. From the time when our ancestors first became cultural creatures, there were selective pressures to develop even greater capacities for culture, because culture was crucial for human reproductive fitness. Cumulative cultural evolution wasn't just a matter of developing new ways of operating more effectively in given environments; it included *niche construction*, the creation of new environments more favorable to the achievement of further cultural innovations, and that in turn produced selective pressures for the enhancement of abilities needed for cultural innovations. Beginning around the second half of the eighteenth century, at first mainly in western Europe but spreading rapidly elsewhere, much larger numbers of people came to live in environments that differed profoundly from the harsher conditions of the EEA. A crucial development that made these new environments possible occurred earlier: starting around 1450 CE, the modern state's monopoly on violence, as the sociologist Max Weber put it (or, as its earlier manifestation was called, the King's Peace), began dramatically reducing the threat of physical violence, a fact that Steven Pinker has convincingly documented, drawing on studies by scholars who specialize in tracking homicide

rates over time. This reduction in the risk of violence made possible the growth of markets, both because it made people more secure in their project of accumulating resources they could use to exchange with others, and because it gave them more confidence that an attempt at exchange wouldn't turn into a violent, unilateral expropriation. Markets have a remarkable feature: they allow mutually beneficial cooperation among strangers, in fact, among people who have no ties of kinship or affection or religion or culture or ethnicity. In this new environment, it paid for people to develop a suite of psychological skills sometimes given the moralizing title "the bourgeois virtues" (Elias 2000, 54–55; McCloskey 2006, 4). To take full advantage of the new opportunities provided by relatively peaceful, market-based societies, people needed to have better impulse control (especially with regard to violence), better ability to read the intentions of strangers from their behavior, better ability to think through the future consequences of their actions and how they would mesh with the consequences of the actions of others, and better ability to signal effectively to others that they were themselves reliable partners in mutually beneficial cooperative arrangements based on voluntary exchange. All of this meant that they needed to become better at reasoning, including moral reasoning, because markets only operate effectively if people generally observe some basic moral rules, and reasoning is required to know how to apply rules to new situations. It's likely that at least at the beginning of market-based societies, individuals who developed these skills had higher reproductive rates, because they did better economically, thus allowing them to live in houses with better sanitation, have healthier diets, and avoid the overcrowded living conditions that bred deadly diseases. Yet even if achieving higher standards of living *initially* conferred benefits in terms of biological reproductive fitness, we know that higher standards of living eventually *reduce* fertility—that wealth and reproductive fitness become negatively correlated at a certain point. In these new conditions, social success or, if you will, "cultural fitness" is not only compatible with low fertility but can even be enhanced by it. Having fewer children and thus being able to invest more in them can produce greater material and social rewards than having more children. Once sufficient

surplus reproductive success is achieved, mating practices need no longer be driven by fitness.

7. At first, the expansion of the circle of regard that market-based, relatively physically secure societies allowed only amounted to accoring strangers the status of potential partners in exchange—extending shallowly inclusive moral responses to more and more people—but it paved the way for further changes in moral beliefs and attitudes that eventually resulted in the First Great Expansion. In the next chapter, I explain that new application of old capacities, assigning a pivotal role to one rather amazing human ability, what I call the capacity for critical, open-ended moral thinking. I argue that this capacity is likely to be widely exercised, in ways that can lead to the Two Great Expansions, only under certain highly unusual conditions.

8. The cultural innovations of the modern era, especially the rise of the modern state and the growth of markets, along with momentous increases in agricultural productivity beginning around 1760 CE, resulted not only in greater physical security, reductions in the frequency and severity of epidemic disease outbreaks, and increased cooperation with strangers; it also produced much greater wealth for more people than ever before. For those who reaped the material benefits of modern social arrangements, extending sympathy beyond the circle of one's close associates and acting altruistically toward strangers—not just being willing to engage in self-interested exchanges with them—became for the first time something one could do without paying a penalty in reproductive fitness or reducing one's own material prosperity to socially unacceptable levels. People who are doing much better materially than they need to in order to survive and reproduce and even live reasonably well according to the standards of their society can afford to care more about strangers, even strangers with whom they will never have any cooperative relationships whatsoever. So cultural innovations that resulted in the creation of the new niche we call modern, “developed” society gave more scope for altruism and for sympathy. For increasing numbers of people, showing moral regard for strangers—even people with whom they weren't engaged in cooperation—no longer imposed the penalties it previously did. As more and more people achieved better material circumstances, their moral horizons widened.



The cultural-innovation-driven developments summarized in items 7 and 8 in the list above amount to this: more people came to live in human-constructed environments that mitigated the harsher conditions of the EEA and therefore made them less subject to the stimuli that toggle our dualistic moral mind toward tribalistic responses. People could afford to become less tribalistic in their moralities; being more inclusive no longer carried the high reproductive fitness cost that it did in the EEA. Cumulative cultural niche construction made it less costly to be more considerate toward and even care deeply about other beings, regardless of whether or not they were seen as potential cooperators.

In other words, *the unprecedented surplus reproductive success of modern societies for the first time in human history made the escape from tribalism possible for significant numbers of people*. But something even more momentous and far-reaching occurred: *surplus reproductive success produced the Great Uncoupling*: the content of moralities was liberated from the demands of reproductive fitness. The surplus reproductive success that humans only recently achieved allowed us to develop moralities that would have been unviable in earlier, less bountiful circumstances, when humanity hovered closer to the edge of survival, and morality was the slave of fitness.

This doesn't mean that people who were lucky enough to inhabit this new environment automatically embraced the Two Great Expansions. But it does mean that a cultural and psychological space was created in which they could afford to do so without paying a prohibitive reproductive or economic penalty.

### **How Cultural Evolution Moved Us Away from the EEA**

Before I speculate about what further changes had to occur for significant numbers of people to walk through the door toward greater inclusion that these new, human-constructed niches opened for them, let me make something clear: the new niches created the possibility of the Two Great Expansions *because they distanced ever increasing numbers of humans from each of the specific features of the locales of the EEA that encouraged tribalistic moral responses*. (a) Public health and sanitation measures (first pioneered in Renaissance

trading centers to prevent the spread of plague from seaports and augmented in the late eighteenth century by the widespread use of vaccinations against smallpox and by scientifically based public health measures a century later) greatly decreased the threat of biological parasites transmitted by strangers. (b) The imposition first of the King's Peace and then of the more systematic and rule-of-law-governed monopoly on coercion of the modern state reduced the threat that strangers posed to physical security. It also reduced the risk that "foreigners" in our midst would free ride on our cooperative enterprises, because it created public enforcement mechanisms that worked in large-scale, anonymous urban areas where older informal social norm enforcement based on kinship and shared culture no longer sufficed. (c) There was a tremendous proliferation of new social practices and institutions for facilitating mutually beneficial cooperation among members of different groups, the growth of various kinds of markets being the most prominent, with the result that the ratio of benefits to risks in encounters with strangers changed radically from what it had been in the EEA. (d) New modes of production increased social surpluses so dramatically that violent competition for resources among groups diminished.

To summarize: if it is true that the prevalence of certain threat cues in many locales of the EEA resulted in moralities that developed *in that environment* being largely tribalistic, and if modern cultural developments have greatly reduced the strength of every one of those threat cues for increasing numbers of people, then you would expect that more people would no longer be so prone to purely tribalistic moralities. When the threat cues aren't present or at least aren't so prominent, we don't need the cultural and psychological equivalents of the water flea's spines and helmets as much as we used to.

### **Why the Cultural Evolution Story So Far Isn't Enough to Solve the Big Puzzle**

Saying that the conditions that toggle the moral mind toward tribalistic moralities have been reduced for many people as a result of the construction of new niches characterized by surplus reproductive

success is one thing. Explaining how and why some of those people actually went on to develop deeply inclusive moralities is quite another. The story of cultural niche construction I've outlined only explains why people who live in that sort of niche would at least develop shallowly inclusive moralities, a limited sort of recognition of the moral status of others, and would have had increasing *opportunities* to go farther than that, to cultivate and act on feelings for strangers and nonhuman animals independently of any prospect of cooperating with them. It doesn't yet tell us why people who could afford (from a biological or cultural fitness standpoint) to become deeply morally inclusive *would in fact become that way*.

### Stepping through the Door That Cultural Evolution Opened

Prepare yourself now for a big shift in my analysis. So far, all the steps I've outlined fit under the heading of an *evolutionary* explanation of movement *toward*, but not yet reaching, the Two Great Expansions. Those speculations include hypotheses about both biological evolution and cultural evolution, because they rely on both biological and cultural selective pressures. The chief role of biological selective pressures was that they produced the moral mind; after that, cultural selection becomes the dominant mechanism for large-scale moral change—so far. In that sense, my account up to this point has been constructed within a broadly evolutionary framework.

Yet as I've just emphasized, my speculations up till now don't get us the whole distance we need to travel to solve the Big Puzzle. At most they make intelligible the possibility that humans have started on a journey that *could* bring them to the Two Great Expansions. At most my eight-stage narrative shows how evolutionary thinking can explain the transition from exclusive morality to shallowly inclusive morality and the creation of the conditions for the *possibility* of deeply inclusive morality.

The next part of my story, then, needs to bridge the gap between "some humans *became able* to become deeply morally inclusive" and "some humans *have actually become* deeply morally inclusive." The bridge that I'm about to construct won't sound sufficiently evolutionary to some people. Why? Because it will depend on a new

assumption: that moral thinking and motivation can take on a life of their own, becoming relatively untethered from the sorts of evolutionary forces we are most familiar with, once the process outlined in steps 1 through 8 has occurred.

Nevertheless, I will try to convince you that this part of the story is scientific in the sense that empirical research could turn out to confirm or disconfirm my hypotheses about the power of moral thinking and moral motivation when they operate together in the right sort of environment.

I'm not going to try to provide all the necessary confirming empirical evidence. Instead I'll offer what philosophers call an "inference to the best explanation." I'll argue that the best explanation of the Two Great Expansions, in fact the only explanation available at present that has even a modicum of plausibility, requires that moral thinking and motivation have the power I attribute to them—again keeping in mind the crucial caveat: when they operate in the right environment and on the basis of the right sort of evolved moral psychology, the highly flexible moral mind that all normal humans actually have.

The story I'm going to tell doesn't make morality something supernatural or otherworldly, something that can bring about progress on its own, as it were, regardless of social context and independent of our evolved biological nature and our historically altered moral beliefs and attitudes. It isn't going to be a story that ignores the evolutionary origins of moralities or pretends that those origins have no implications for the possibilities for what moralities can be. It's also *not* going to be the sort of story that moral philosophers have often told in the past: I'm not going to assume that human morality is a purely rational affair and that rationality can expand the circle of moral regard without a lot of help from the emotions. Nor am I inviting you to accompany me on a journey through the history of ideas, presented as if it were an autonomous process. I'm taking seriously the commitment to providing a thoroughly naturalistic explanation.

Finally, it's worth recalling that my speculative explanation doesn't present moral change in the direction of greater inclusion as endogenous to moralities. That is, it doesn't characterize

improvements in moralities as originating exclusively within them. Instead, as I've already said, it relies on the assumption that the use of coercion in a competition for power largely unconstrained by moral scruples played a central role in creating the conditions under which human beings could first develop shallowly inclusive moralities—a necessary condition, as it turns out, for some of them going on to develop deeply inclusive moralities. In other words, the ruthless quest for dominance that resulted in the modern state's near monopoly on violence and its support of markets, including the enforcement of property rights that solidified and even increased gross inequalities, laid the foundation for a fundamental and progressive change in human moralities.

### **Another Necessary Condition for the Two Great Expansions: The Capacity for Critical, Open-Ended Moral Reasoning**

I believe that the Two Great Expansions would not have occurred unless human beings possessed a remarkable ability: what I referred to earlier as *the capacity for critical, open-ended moral reasoning*. In its simplest form, this is the ability to make explicit the moral rules you are following, subject them to critical scrutiny, and modify or abandon them in the light of reasoning about them. The critical light that this capacity shines on our moral life can even illuminate something as fundamental as the rules that govern how and to whom we assign moral standing and equal high basic moral status. Even more remarkably, exercising this capacity can result in a new understanding of what morality itself is about, as occurred when some people began to think that it wasn't just a matter of obeying the commands of some supposed higher authority. It was this remarkable capacity that allowed some people to embrace the Two Great Expansions.

Furthermore, I'll argue that for some people at least, this shift is not just a matter of positive feelings that we evolved to have toward those we cooperate with somehow irrationally spilling over to other beings with whom we don't and never will cooperate. So to that extent, my solution to the Big Puzzle will put me in the "rationalist" camp of moral philosophers. Yet I don't think my account is unscientific or not naturalistic. Nor is my account incompatible with a

plausible sentimental moral psychology. My rationalism is of the moderate variety; I don't deny that emotions play a central role in moralities and in moral change. (On the contrary, as I've already said, the extension of sympathy was instrumental in the shift from shallow to inclusive moralities.) In fact, it doesn't matter much whether you characterize my position as moderately rationalistic or moderately sentimental or hybrid.

Let's start with moral standing, because it's the simpler of the two concepts. To say that a being has moral standing is simply to assert that it counts morally, in its own right. It is not a mere thing; to behave morally, you have to take into account that its interests matter—and not just so far as doing so benefits you or serves some other purpose. For example, to recognize that a pig has moral standing means acknowledging that morality imposes significant limits on how you are allowed to act toward it simply because of the kind of being it is, independently of any strategic purpose that might be served by observing those limits.

The extension of the concept of moral standing to nonhuman animals marks a momentous change in moral thinking, a revolution. For many people nowadays, that shift is grounded in the assumption that sentience that is the basis of moral standing: they believe that any being that can feel pain and pleasure has moral standing.

However, we shouldn't assume that our understanding of the basis of moral standing is static or that the understanding we have of it is fully consistent, much less optimal. In fact, it may be changing right now and in a way that eventually will result in recognizing that some organisms we previously thought didn't have moral standing do have it—and that some animals have a moral status approaching our own.

What I have in mind are recent revelations about the abilities of certain invertebrates (more specifically, cephalopods) and in particular octopi (Godfrey-Smith 2016, 98–106). Invertebrates don't have a spinal column with a cord running through it, and for this reason, they may not have the kind of neurology that makes it accurate to say that they feel pain and pleasure—at least not the pain and pleasure that vertebrates like sheep and dogs and humans experience.

Until now, a lot of people who think that animals have moral standing have assumed that they have it because—and only because—they are sentient, that is, because they can feel what we recognize as pain and pleasure. By implication, they denied moral standing to animals like invertebrates that they thought weren't sentient (or sentient in the "right" way, namely, the way we are). But now, in the light of recent research on invertebrates, more specifically cephalopods, which include octopi, we are learning that these creatures are much more like us than we thought and that they possess capacities that are very much like the capacities that we believe give us not only moral standing but an especially high moral status, relative to many other kinds of organisms.

Octopi turn out to be not only playful but wonderfully creative in their play, and they can quickly learn to interact playfully and intelligently with humans. One might even say they show a kind of genius in developing novel ways of interacting with items and agents in their environment. They also show great skill in niche construction: for example, they collect pieces of bivalve shells, transport them over considerable distances, and assemble them into shelters for themselves. If niche construction and the ability to act creatively are indications of intelligence, then octopi are highly intelligent.

In the light of this new appreciation of the capacities of cephalopods, some people at least are now beginning to question whether it is only sentient beings, beings that can experience pain and pleasure—at least as we understand those sensations—that have moral standing. In fact, we might even be moving toward the more radical conclusion that some creatures that aren't sentient in the way we are have a rather higher moral status than some creatures who are quite like us so far as sentience is concerned. Eventually we might even go so far as to conclude that if a being is sufficiently intelligent (and conscious and self-aware), then it has the same high moral status we humans do even if it is not sentient.

The point is that part of what we think makes it appropriate to confer a high moral status on ourselves is our intelligence, combined with consciousness and self-awareness. In fact, if we encountered an extraterrestrial being of high intelligence who was conscious and



self-aware but for some reason wasn't sentient (or sentient in the way we humans are), we'd be wrong to conclude that he or she or it didn't have a moral status equal to ours. Instead we ought to conclude that sentience (or sentience as we have understood it) isn't necessary for equal high moral status, though it might be sufficient for having moral standing.

The process of rethinking our notions of moral standing or moral status involves an important kind of critical, open-ended moral thinking: moral consistency reasoning. The idea is that if we think that certain traits give *us* moral standing (or are even so valuable as to confer on us the highest moral status), then we should draw the same conclusions about *other* creatures, if they also have those same traits. We should treat like cases alike, applying our moral principles consistently, reconciling our intuitive moral responses with the values we hold most dear.

Moral consistency reasoning has a remarkable feature: it can produce radical results by proceeding in a rather conservative way. Moral consistency reasoning isn't bootstrapping. You start with values or moral principles you already have, try to identify which of them is most important, and then strive to make your beliefs, intuitive moral responses, and behavior consistent in the light of those moral priorities. Moral consistency reasoning starts at our present moral location, but it can take us to unanticipated destinations.

For that sort of reasoning to reach all the way to a reconsideration of something as morally fundamental as our understandings of equal moral status and moral standing—and for its conclusions to be sufficiently motivating to make significant changes in our behavior—a lot of factors have to be in place. For one thing, we have to know what the other creature is really like—and until recently, humans had little accurate knowledge about many creatures that, at what ultimately turns out to be a superficial level, are quite different from us.

From an evolutionary standpoint, it would hardly be surprising if our "theory of mind"—our ability to infer mental states from behavior—was rather anthropocentric; after all, that's all we needed to cooperate with other human beings. Cephalopods look so very different from us that we may have difficulty seeing just how intelligent they are—and how much their intelligence is like ours—unless



we are scientifically trained to observe them. If our only interaction with them consists of killing and eating them, we aren't likely to appreciate the full range of their abilities.

To conclude that consistency in our moral responses requires reevaluating the moral standing or status of any creature, whether human or nonhuman animal, we also have to be able to see through the irrelevant differences and abandon our narcissistic prejudices, our smug confidence that we alone are so very special. In some cases, religious beliefs, especially those grounded in unscientific creation myths, may also raise a barrier to sound moral consistency reasoning about moral standing and moral status.

Believe it or not, I have encountered graduate students in an elite American university who think that no nonhuman animals feel pain, because God didn't insert souls into them. Quite apart from that bizarre theological view, the Bible contains passages asserting that God gave humans dominion over all living things and suggests that their only value is that we can use them to serve our purposes. In chapter 7, on regression to tribalism, and in chapter 8, on the evolution of moral tribalism to include intrasocietal tribalism, I will have more to say about the cultural conditions under which moral consistency reasoning is likely to function well and about the factors that can derail it. For now, I want to make clear that, given the anthropocentric or "speciesist" biases that humans are prone to, it would be surprising if current moral consistency reasoning concerning the status of nonhuman animals gets it right.

### **The Continuing Evolution of Understandings of the Basis for Moral Regard**

The Second Great Expansion was a true moral-conceptual revolution, because it was change at the deepest levels of morality: an expanded understanding of the kinds of beings that have moral standing. Nevertheless, it may well be that the Second Great Expansion is incomplete, defective in one important way: even if sentience is a sufficient condition for having moral standing, maybe it isn't necessary. Perhaps having the sort of complex mental life that octopi have is also sufficient, even if those fascinating creatures

lack sentience (or sentience as we've understood it so far, in our mammal-centric fashion).

Nowadays the moral-conceptual framework within which many humans operate includes more than just the idea of moral standing—the idea that the world divides only into two classes of things, those that matter morally on their own account and those that don't. Many people today also believe that the class of beings who have moral standing divides into two major subclasses: those that only have moral standing, and those that have an especially elevated kind of moral standing, a high moral status. People who have embraced the First Great Expansion believe not only that all human beings have moral standing, but also that all humans, and humans alone, have an especially high moral status. This amounts to believing that even if many beings count morally, human beings—all human beings—count a lot more. The idea not only that all humans count a lot more, morally speaking, but also that all of them count equally is the idea of the equal basic moral status of all humans.

It's the coming to be of that idea of basic human equality, along with serious (though admittedly imperfect and fragile) efforts to live that idea, to realize it in our personal behavior, social practices, and institutions, that I've called the First Great Expansion.

In both the First and Second Great Expansions, a change occurred in the way people think, not just in how they feel. Why? Because at least some of the people who have made both shifts, if queried, are able to give reasons for them, and in some cases it was their becoming convinced of the reasons that began the process of their feelings changing. Until recent research on octopi and other invertebrates shook us up, most of us believed that what gives nonhuman animals moral standing is the fact that they are sentient.

Before I read about that research, I relished eating octopi; now I find the very thought of doing so disgusting. For me, the moral emotion of disgust was directed toward a new target. Because of what I now believe about what octopi are like, the thought of eating them causes me to have just the same sort of *reason-based* disgust that many of us feel when we learn the horrific facts about factory farming. I've made an inference: I've moved from the proposition that high intelligence confers moral standing and the proposition

that octopi are highly intelligent to the conclusion that I ought to treat octopi as beings with moral standing. Being consistent requires that. Moral integrity requires that.

I want to emphasize that I didn't change my attitude toward eating octopi automatically and nonrationally by simply imitating someone I looked up to who had stopped eating them on ethical grounds. Sadly, everyone I hang out with who isn't a vegetarian or vegan shows no evidence of worrying about the propriety of eating octopi. I also didn't change as a result of some kind of emotional contagion from others who exhibited disgust at eating octopi. (So far, I've not seen anyone else exhibiting such disgust.) My conversion was not a cognitively empty emotional reaction. Reasoning played a central role, just as it has in many people becoming vegetarians or vegans.

When I say that the growing recognition that at least some non-human animals have moral standing is a case of moral progress, I don't mean that this change in moral orientation has gone as far as it will go or should go. On the contrary, I think it's extremely likely that the Second Great Expansion is far from complete, even at the conceptual level, quite apart from shameful failures of implementation. It may well be that a result of further progress in recognizing the moral standing of some nonhuman animals will lead to a rejection, not of the idea that all humans have the same basic moral status, but of the idea that *only* human beings have the highest moral status. It might turn out that, according to the best understanding of the capacity that gives all human beings *equal* basic moral status, some nonhuman animals also have it.

Suppose you think that what gives all humans high moral status is that they can distinguish between what they desire and what they ought to do—that they have the concept of an “ought.” Some evolutionary thinkers believe that although chimps and bonobos have a limited kind of morality in the sense that, being motivated purely by sympathy, they sometimes act altruistically toward other chimps with whom they interact closely, they don't operate with a sense of “ought.” These evolutionary thinkers think that only humans do.

Suppose they're wrong. Suppose that further study of chimps (or some other animal species) reveals that they do have the capacity to

distinguish between what they desire to do and what they ought to do. If we find out that some nonhuman animals have that capacity, and we think that our having it is what gives us the highest moral status, then we ought to recognize that those animals have the same high moral status that we do. That recognition wouldn't affect the other part of the conceptual shift that the First Great Expansion entails—the recognition that all humans have an *equal* high moral status—but it would require giving up the idea that humans alone have that high status.

### **Toward a Moderately Rationalist Understanding of Basic Moral Change**

The moral psychologist Jonathan Haidt, like Steven Pinker, is a professor whose writings influence opinion far beyond the academy. Haidt has a very nonrationalistic, extremely sentimentalist understanding of human moral psychology. He emphasizes that most of the time, “the emotional dog wags the rational tail” (Haidt 2001, 830). He doesn't say whether “most of the time” means 50.01 percent of the time or 70 percent of the time or 90 percent of the time. But his rhetoric (including the dog-tail analogy) strongly suggests he doesn't mean anything as slim as a bare majority of the time. In brief, he doesn't give much credit to the role of reasoning in our moral lives.

I think he grossly underestimates the role of reasoning in the moral life generally but especially in moral change. Sometimes the rational dog wags the emotional tail and those cases are extremely important for the possibility of large-scale moral change. Reasons regarding who has moral standing are sometimes more than post hoc rationalizations of emotional changes in which reason played no role. Instead, sometimes reasoning stimulates emotional changes. That's a fair description, I think, of both the First and Second Great Expansions, at least so far as some people who have experienced them are concerned.

Nevertheless, I don't want to make the opposite error of assuming that the causal arrows only or even most of the time go from reasoning to emotions. I think the story is more complicated than

either of those one-sided views. Sometimes emotions get stimulated without much in the way of inferences but nonetheless lead to changes in our moral responses and judgments. This may happen, for example, when we see a video of a starving child and instinctively say “we’ve got to do something about that” — without going through any process that could be aptly called moral reasoning. Yet sometimes new information (for example, about the capacities of nonhuman animals like octopi) gets us thinking and leads us to conclusions that can change our emotional responses.

There has been less consensus on what gives human beings the highest moral status than there has (until the recent revelations about cephalopods) on the idea that sentience confers moral standing. Religious folks often say that because all humans are God’s children, created in his image, we all have the same high moral status. (I find that to be a disturbingly nepotistic view of basic moral status, but that’s neither here nor there for present purposes.) Some moral philosophers believe that it is our practical rationality, our ability to reason about what we ought to do, that makes us so morally special among all the beings that have moral standing. Others think it is something more basic: the ability to distinguish between what is desired and what is good or between what one wants to do and what one ought to do.

A variant on the idea that practical rationality is the basis of equal basic moral status is what some of the most influential contemporary philosophers in the Kantian tradition, including Rainer Forst and Stephen Darwall, refer to as responsiveness to reasons — the ability to justify our actions to others and to be open to serious consideration of the justifications they offer. No consensus exists at present, however, even among contemporary philosophers, as to what grounds moral standing or equal basic moral status (Darwall 2009, 127, 281; Forst 2014, 23–28, 75–76).

Furthermore, what exactly high moral status amounts to and what counts as a proper acknowledgment of it is also not only subject to dispute but evolving as well. I noted in the introduction that at present the modern (post-World War II) conception of human rights is probably the most developed, substantive interpretation of the idea that all humans have a basic (high) equal moral status.

That understanding of what equal basic moral status entails began to take shape earlier, with the development of the concept of natural rights—moral rights that we are all supposed to have simply by virtue of our nature as humans. But filling out the content of the notion of equal moral status using the notion of human or natural rights is only one possibility among others; and even people who think it is the best alternative disagree about what rights are on the list.

I think it's likely that for a growing number of people an appreciation of new information about some nonhuman animals, combined with a sense that attempts to explain why human beings alone have an especially high moral status are less than fully convincing, may result in a blurring of what had previously been a basic structural feature of their moralities. They may become much less confident in the viability of the distinction between the supposedly minimal moral standing of nonhuman animals and the much higher equal standing of humans. That pretty much describes my present predicament.

Now that I've explained why I think the Two Great Expansions aren't a *fait accompli* but rather a work in progress, I can begin to fill out the explanation of how they began to come about. To do so, I first need to explain how human beings could come to *think* that not just members of their own tribe or cultural or ethnic or cooperative group but all human beings have an especially high moral status and that some nonhuman animals also have moral standing. What would the human mind have to be like for those two thoughts to arise in it, given the evolutionary origins of human morality? That's the first key question. In the next chapter, I try to answer it.

Before I make that attempt, another, prior question must be answered: What about the place of the capacity for critical, open-ended reasoning in the evolutionary story? Where does that capacity come from? Is there an evolutionary explanation of it? Is it an adaptation for cooperation?

Two possibilities are worth considering. The first is that this capacity is part of the moral mind, an ability that humans had from the start of their being moral in a distinctively human way. The second possibility is that even if, strictly speaking, this remarkable capacity wasn't part of the moral mind, it developed out of some

feature of the moral mind, perhaps the capacity for the most rudimentary forms of moral consistency reasoning. Because the latter capacity is required for something as basic to moralities as applying moral rules to new situations, it seems reasonable to conclude that it was there from the start, an element of the moral mind itself.

I don't know which option is the best. Fortunately, that doesn't matter for most of what I want to say about large-scale moral change in general and about the Two Great Expansions in particular. I think it is clear that humans have the capacity for critical moral reasoning, and equally clear that it is open-ended in the sense that it has the ability to challenge some of the most fundamental aspects of a person's moral outlook, including her conceptions of equal moral status and moral standing. I'm much more concerned with understanding the role of this capacity in large-scale moral change than in establishing whether it is part of the moral mind or something that emerged on the scaffolding of the moral mind.

Having said that, I fully acknowledge that there might turn out to be a genuinely evolutionary explanation of why we have this capacity that makes it plausible to say that it is part of the moral mind. So let's explore, if only briefly, the possibility that one important form that this capacity takes, moral consistency reasoning, is susceptible to an evolutionary explanation.

There are at least two ways that being able to engage in moral consistency reasoning might have conferred a fitness advantage on individuals in the early ancestral environments in which the moral mind took shape: by making those who exhibit such reasoning more attractive choices for partners in cooperation and by aiding the formation of coalitions. Let's consider each of these in turn, beginning with the partner choice idea.

The outstanding philosopher of biology Kyle Stanford suggests one way in which being able to engage in moral consistency reasoning would make one a more reliable partner in cooperation and therefore be to one's reproductive advantage (Stanford 2018, 11, 19–20). Suppose that you and I are in the EEA. If I do something that you think violates our group's rules, you may refuse to partner with me. If I'm to avoid that potentially disastrous outcome, I need some way of reassuring you that I am in fact reliable. One way I can do



that is by trying to convince you that, in spite of appearances, I am a conscientious person, an individual with a strong moral identity and hence a reliable partner in cooperation. If I can convince you that I wasn't flouting moral rules but instead judiciously acting in a consistent way that recognized the point of the rules and which rules are most important, while discerning that there are justifiable exceptions to rules, that will be highly advantageous for me. According to this way of thinking, the capacity for moral consistency reasoning first came about because it played a strategic role, providing a defense against judgments of others that, if left intact, would result in an individual being excluded from cooperation, with the disastrous consequences for reproductive fitness this would entail.

Richmond Campbell, Jennifer Woodrow, and Victor Kumar offer a second explanation of why engaging in moral consistency reasoning would make one a more attractive partner in cooperation and thereby enhance one's fitness. They suggest that individuals who engaged in moral consistency reasoning would exhibit more predictable behavior than those who didn't, and this, too, would make them more attractive as partners in cooperation, which in turn would enhance their reproductive fitness (Campbell and Woodrow 2003, 361, 367, 371; Campbell and Kumar 2012, 303).

The pathbreaking work of Christopher Boehm provides my first example of the second type of adaptationist explanation of the capacity for moral consistency reasoning: the coalition-building account. Boehm suggests that the capacity to engage in this kind of reasoning was necessary for the formation of punishment coalitions to suppress bullies, which in turn contributed to the reproductive fitness of the suppressors (for example, by keeping control of resources that the bully would have expropriated). Boehm emphasizes that the need to control would-be dominators was a significant force in shaping human morality (he's not clear, as I noted earlier, whether "human morality" means the moral mind or the first moralities).

He emphasizes that in present-day hunter-gatherer groups that he and others have studied—and presumably in our distant ancestors whom they closely resemble—control over would-be dominators is frequently achieved by coalitions. The idea is that people talk among themselves until they reach a consensus that it is time to do



something about the bully in their midst. To determine whether to act and if so when and with what degree of severity of punishment, some basic form of moral consistency reasoning would most likely be necessary. For example, it might be necessary for some members of the nascent punishment coalition to convince others that the current case was relevantly similar to a case that had occurred earlier—that the proposal to punish the current bully in this way was just like a decision that everyone approved of in another case. Moral consistency reasoning could also be valuable for convincing potential coalition members that the proposed punishment wasn't just a pretext for inflicting harm on someone that some members of the group happened to dislike. In brief, the idea is that individuals developed the capacity for moral consistency reasoning because it was useful for persuading other individuals to help curb bullies.

The exceptionally original and systematic evolutionary psychologists John Tooby and Leda Cosmides add another piece to the puzzle (Tooby and Cosmides 2010, 213–230). They, too, think that moral reasoning first developed because it enabled individuals to engage in strategic behavior vis-à-vis other individuals in ways that enhanced reproductive fitness, and they recognize that moral reasoning can be critical and open-ended rather than static. Like Boehm, they emphasize the importance of coalitions, but they focus on coalitions formed for purposes of making war against out-groups, not for suppressing bullies within the group. In their view, individuals who sought to build coalitions for violent competition with other groups needed to develop moral reasoning skills for doing so. Other things being equal, the more people a coalition builder could enlist, the more successful his coalition would be. To attract the greatest number of coalition members, leaders had to engage in reasoning that appealed to interests that were widely shared. They had to learn to identify common interests and convince people that the coalition would promote those interests. They would also have to convince potential coalition members that the goals of the coalition gave due weight to everyone's interests, that some members wouldn't be arbitrarily asked to take risks and bear costs that mainly benefited others, not themselves. So, to be effective, coalition builders would have had to develop moral reasoning

skills that employed the idea of impartiality. To be persuasive, this reasoning would have to be consistent in the sense that it took the interests of each party into account, with no favoritism shown to any particular individual. Successful war-making coalition building would require moral consistency reasoning and also the development of concepts of impartiality and nonarbitrariness.

All three of these accounts might be elaborated to encompass not just moral consistency reasoning but other forms of moral reasoning as well, including reasoning that is critical of the moral status quo in rather fundamental ways. Under the right conditions, such reasoning could lead beyond the immediate strategic context to unanticipated destinations. Remember, the fact that something comes to be because it performs some function that promotes fitness doesn't mean that performing that function is all it can ever do. Even if moral consistency reasoning was originally strategic, that doesn't mean it is always like that.

Suppose that, taken together, the explanations offered by Stanford, Campbell, Woodrow, Kumar, Boehm, Tooby, and Cosmides make a convincing case that the capacity for moral consistency reasoning is an adaptation, a trait that arose because, by serving strategic purposes, it contributed to individual reproductive fitness, and that this capacity worked well because it was flexible, not limited to any particular task. If that were true, then we'd have an evolutionary explanation of the existence of the capacity for moral consistency. Yet we still wouldn't have an evolutionary explanation of how that capacity eventually came to be exercised in ways that led some people to adopt less tribalistic moralities.

Whether or not it is itself an adaptation or a by-product of an adaptation, the capacity for critical open-ended moral reasoning can greatly expand the space of possible human moralities—if it is exercised under the right conditions. In the next chapter, I offer an account of just what those conditions are.