

# 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>

---

## Large-Scale Moral Change: The Shift toward Inclusive Moralities

### The Late Arrival of the Concept of a Human Being

The Native American group whom most people know as Comanches called themselves “Numunuu,” which is usually translated as “the human beings.” (The name “Comanche” is derived from a Ute word meaning “those who fight us all the time.”) In the not-so-distant past, many peoples called themselves—and only themselves—the human beings. The names that people who called themselves the human beings had for people of other groups were often rather unflattering. They frequently translate as “the enemy” or “the ones we always fight,” or denote dangerous or unclean animals or inferiors of one kind or another or focus on some alleged feature of the Other, often a rather unattractive one.

If the name they gave themselves is an indication, the Numunuu confused being Numunuu with being human. They apparently didn’t understand that non-Numunuu were human beings. And if they didn’t understand that, then they didn’t really understand who they themselves were, namely, only one group of humans among many, just one part of humanity.

Whether we should take the Numunuu’s name for themselves as a clear indication of their concepts is open to dispute, but there’s a big point to be made here. To put it bluntly: people who think that they are the only human beings don’t really know what a human being is, and they don’t appreciate the moral significance of just being human.

If that's the way the Numunuu thought, they weren't unusual. For most of human history—until quite recently, in fact—most human beings lacked the moral concept of a human being, whether their names for themselves and for others reflected that fact or not. They didn't have the idea that everyone, whether one of their own people or a member of other racial or ethnic groups, has an *equal basic moral status*, that all are equally entitled to moral consideration and respect, simply because they are human. Instead their concept of basic moral equality was tied to group membership. In contemporary terms, they didn't have the idea that we all have human rights, that we are all morally entitled to be treated in certain ways, just by virtue of our humanity.

Even today, when people who believe we are all equal in some fundamental moral sense find themselves in dire circumstances—or when propagandists convince them that they are—they may either abandon that belief or fail to behave as if they took it seriously. They “dehumanize” the Other and act accordingly. During the genocide in Rwanda in 1994, Hutus routinely referred to the Tutsis they butchered as “cockroaches.” Hitler called the Jews a bacillus infecting the body of Europe. Racists refer to Africans as monkeys or baboons. President Trump speaks of the “infestation” of illegal immigrants crossing the US border with Mexico, conjuring images of nasty and potentially dangerous insects or rodents. In October 2018 a participant in the Fox network television program *Fox and Friends* said that Central American people headed for the United States carried smallpox, overlooking that this disease was officially declared eradicated worldwide in 1980 (with the last recorded lone case occurring in Somalia in 1977). Representing somebody only as the carrier of a deadly disease characterizes them solely as a threat, undermining the belief that what really matters is that they are a human being, just like us.

Tribalism dehumanizes in another way. The tribalism at work in American society today often doesn't directly or at least explicitly attribute less-than-human status to compatriots on the other side of the ideological divide. It pays lip service to their being equal citizens and may begrudgingly acknowledge that they have human rights. Nevertheless, the ideologies that drive this new form of

tribalism do in a sense dehumanize: they present the Other as *beings with whom one can't reason, either because they are irremediably confused about the nature of society and what is valuable or because they are insincere (or both)*.

If you think genuine human beings are amenable to reason, are reason-*able* creatures—and believe that's part of what distinguishes them from other animals—then you won't regard people you believe to be incorrigibly unreasonable as fully human. For example, when college students shout down speakers on their campus and refuse to let them speak because they've already branded them as racists and hence disgusting and not entitled to be heard, the students are in effect categorizing the speakers as creatures one needn't feel obligated to listen to and engage in reasoned disagreement with.

It's one thing to argue that hate speech is not protected by a proper understanding of freedom of expression; it's quite another to sidestep the question of whether what someone is saying really *is* hate speech by attaching the label "racist" to the speaker (especially if this is done on the basis of a dubiously expansive or inchoate notion of what racism is). When you're in the grip of tribalism, merely labeling someone as a racist (or a socialist) is enough to justify dismissing what that person says or not listening to it, much less trying to evaluate its truth or falsity. And that's especially dubious if you've never thought much about how to define racism (or socialism) in the first place.

Please note that I'm *not* assuming or in any way implying that the Numunuu or other people who lacked the moral concept of a human being were immoral. It's more accurate to say that they had a truncated, incomplete morality. They acknowledged moral duties toward members of their own group, often very demanding ones. Yet they regarded other human beings as beyond the moral pale or at least as having a much lower moral status. They were moral beings, but their morality was incomplete because they lacked the moral concept of a human being.

From the moral standpoint that I (and I expect you) occupy, this incompleteness wasn't a minor defect. A lot can go wrong in how you treat other members of our species if you don't recognize them as fellow human beings in the moral sense. You may think it

is perfectly acceptable to throw them off the land they've always occupied, kill them, enslave them, cause them pain for your own amusement, hunt them for sport, or even butcher and eat them. Humans have done all these things to one another. In fact, until recently in the long history of our species, these kinds of behavior were quite common, including cannibalism. (Given that the Paleo diet probably included substantial portions of human flesh, just how committed to it are you?) So there's another way to put the point that some humans have made moral progress in the direction of greater inclusion: by becoming more inclusive, some moralities have become more moral. If regression occurs, moralities become less moral.

To fend off the inevitable charge of Eurocentrism or colonialism or ethnocentricity (or whatever your favorite bad sort of "-ism" happens to be), let me make one thing clear. To say that many humans (perhaps most humans who ever lived) have had an incomplete morality because they lacked the moral concept of a human being is *not* to say that if your morality includes that concept, it's complete or without defect. Although a morality that isn't deeply inclusive is incomplete and seriously defective, it may have a lot else going for it. I'm only focusing on one important element of morality that I (and I imagine you also) think is especially important.

## **The Two Great Expansions of the Circle of Moral Regard**

Something has profoundly changed; a remarkable transformation has occurred. Nowadays, the idea of a common humanity, of an equal basic moral status shared by all human beings, is pretty widespread: we speak of human dignity, not just the dignity of this or that group. We talk about human rights, not just the rights of Frenchmen or Americans or Hutus or Tutsis and so on. Because it is so momentous, this shift from morality as a purely parochial, exclusive, tribalistic affair to one that is deeply inclusive merits a suitably dramatic name: the First Great Expansion. It's an expansion because it amounts to an enlargement of one's conception of who is a member of the primary moral community, the set of beings who have first-class moral status.

Most people today who acknowledge the fundamental moral equality of human beings came to have that moral orientation in the same way they received other items of cultural inheritance: they learned it through imitation of, and inculcation by, their parents and peers. Most of us didn't think it through on our own, at least not explicitly. Consequently we're not likely to notice this transformation. We take it for granted.

It's worth pausing and reflecting on just how momentous the First Great Expansion is. It makes a great difference whether you believe that all human beings are equal in some fundamental sense or whether you think—as most people did throughout most of history—that there is a natural order of superiority and inferiority. There's a funny thing about moral progress: once it has occurred, it's largely invisible and hence unappreciated. Steven Pinker's work is extremely valuable because he makes what had been invisible visible. That's the first step toward appreciation—and gratitude. I will argue that there is a second step that's just as important: understanding how much luck was involved in human beings' getting to the point where they could make that sort of progress—and how fragile that accomplishment is.

The First Great Expansion involves a radical reconfiguration of our attitudes toward human beings, all human beings. The basic idea is that all are equal at the most fundamental moral level, meaning that all are entitled to be treated in certain ways simply by virtue of their humanity, quite apart from which human group they belong to, whether they can interact with us in mutually beneficial ways, and whether they have done something to earn their being treated in those ways. A few centuries ago, the idea of the equal basic moral status of all human beings took the form of the doctrine of natural rights; today we speak of human rights. The world looks different for those who have embraced the First Great Expansion: when they identify a creature as a human being, they understand that it is their humanity that matters, at least so far as the most basic rights are concerned. Differences among human beings (race, gender, class, nationality, ethnicity, etc.) are ruled out as irrelevant.

A Second Great Expansion has also occurred: the growing recognition that not just all human beings count morally, but that at

least some nonhuman animals do, too (even if they don't have the high moral status that we humans do). Extreme cruelty—gratuitous infliction of pain that was often a source of pleasure for humans, not just an unfortunate side effect of food production—has been common throughout most of human history. As the eminent historian Norman Davies documents, in Europe in the Middle Ages cat burnings were considered a wholesome form of public entertainment (Davies 1996, 543). During the same period, a popular “sport” involved nailing a cat to a post and having young men compete in how many head-buttings it took to them to kill it. Treating nonhuman animals in such ways is now considered by many people—many more than ever before in human history—to be grossly immoral. That's a large-scale change in the moralities of some humans. It amounts to the recognition that at least some nonhuman animals are not mere things—objects that we can use for our own purposes, without moral constraints of any kind. For those who have embraced the Second Great Expansion, the world looks radically different: the line between mere things and beings that count morally in their own right has shifted dramatically.

The First Great Expansion enlarges the pool of humans considered to be first-class members of the moral community. The second enlarges the class of beings that count morally at all.

The Second Great Expansion, like the first, is incomplete. I'd say it's barely begun. Even today a town in Indonesia holds an annual “barbecue” festival where kittens are bludgeoned to death, roasted with blowtorches, and then eaten. Cockfighting and dog fighting—both of which result in the death of the “contestants”—still occur in the United States, though both “sports” are now illegal there and in many other countries. Furthermore, even though some corporations have yielded to pressure and made their methods of killing food animals and the conditions under which they are raised less brutal, the total *quantity* of suffering that humans inflict on animals is probably much greater today than ever before, because of the sheer scale of factory farming.

This book addresses not only why the Two Great Expansions have occurred for some of humans, but also why they haven't become universal. It also explains why, where they have occurred,



they have been late arrivals—and why they are fragile, alarmingly subject to regression.

Admitting that the Two Great Expansions aren't universal and are only imperfectly acted on even where they have occurred isn't the same as saying they are insignificant, much less that they are merely different ways of thinking, rather than huge behavioral shifts. Many human beings, as individuals and through collective action in their political life, *live* both the idea that every human has an equal basic moral status and the idea that at least some nonhuman animals count morally in their own right.

So far, the most fully articulated and institutionally embodied expression of the First Great Expansion is the culture of human rights. It shapes the legal systems, including the constitutions, of many countries around the world, and scholars have documented that the culture of human rights is now a significant force in international relations, even though sheer power still often determines important outcomes in many instances. Massively expensive efforts by international organizations and global nongovernmental organizations are undertaken in the name of human rights, and sometimes they actually have significant effects, so one can't dismiss human rights as being merely aspirational (Sikkink 2012, 17; Simmons 2009, 3). As impressive as the cultural and institutional impact of the modern idea of human rights is, however, it's important to remember that it is only one way, not the only way, in which the idea of basic moral equality has been worked out or can be worked out. As I'll explain later, people still disagree about what exactly counts as equal basic status and about what properties of a being make it appropriate to ascribe equal basic moral status to it. Nonetheless, the First Great Expansion is a fact, and it's a case of large-scale moral change.

Similarly, the Second Great Expansion, sometimes called animal liberation, is also more than just a change in language or thinking. Many countries, at all levels of government, now have laws that curb some of the practices that cause extreme suffering in animals, and governments often expend a great deal of money and other resources to try to enforce those laws. Just ask any scientist who works with laboratory animals how much more complicated and

costly her work is than it was a generation ago. Scientists in most countries that are at the forefront of medical research must take great care in how they treat lab animals, even those used in experiments designed to discover cures for diseases that inflict great suffering and early death on human beings. Powerful corporations, including McDonald's and Walmart, have found it necessary to change the way they operate in response to pressures from individuals and organizations that are convinced that cows, chickens, and pigs are not mere things.

### The Big Puzzle

How did these two seismic shifts in moral orientation come about? What makes the question so interesting is that behind it lies what I'll call the Big Puzzle: given what evolutionary theorists tell us about how human morality originated in the remote past, how could it have become deeply inclusive? *How could creatures like us, organisms who evolved in the way we did, ever come to embrace the Two Great Expansions?*

That puzzle appears large because, according to a standard evolutionary explanation of how human morality came to be, one would expect it to be purely parochial, exclusive, tribalistic, not at all inclusive. One would expect all humans to lack the moral concept of a human being and all that goes with it, and expect that no one would ever have come to believe that nonhuman animals count morally in their own right. It would seem even more preposterous that humans would ever come to limit the pursuit of their most vital self-interest for the sake of nonhuman animals used in vital medical research. Evolutionary accounts of human "moral origins" (with some difficulty) explain altruism toward nonkin humans; but they seem to be at a complete loss when it comes to explaining altruism toward nonhuman animals.

Later I'll have much more to say about what exactly the standard evolutionary view of the origins of morality is, but for now a short summary will suffice. According to the standard view, morality is an adaptation (in the proper, Darwinian sense): something that first arose and then spread because it contributed to reproductive fitness. How did it contribute to our reproductive fitness? By facilitating

cooperation *within* human groups—including cooperation that allowed a group to compete successfully with other groups.

### **Morality as an Adaptation**

In the standard evolutionary account, human morality came to be because it performed certain functions—managing conflicts, coordinating behavior, reducing free riding, dealing with the problem of how to share scarce resources, regulating sexual behavior—that facilitated successful cooperation *within the group*. And by facilitating successful cooperation, morality enabled humans to pass on their genes to the next generations, as well as to reproduce their cultural practices and traditions. In brief, morality, when it first appeared among our remote ancestors, was all about cooperation, cooperation that contributed to reproductive fitness.

The striking thing about this evolutionary understanding of morality is that it leaves members of other groups, as well as non-human animals, out of the moral picture. Given the standard evolutionary account of the origins of human morality, you wouldn't expect human beings to be capable of inclusive morality, of believing and acting as if they believed that the moral community includes all humans and at least some animals, not just members of their own group, and that all humans have the same high basic moral status.

### **A Nonsolution to the Big Puzzle**

Let me forestall a quick solution to the Big Puzzle that you might be tempted to offer—a solution that, if correct, would warrant tossing this book aside right now. You might think that the view that morality is all about cooperation is correct and that it can explain how the circle of moral regard gets expanded, namely, by expanding cooperation to include more participants. In other words, you might think that we now recognize the basic moral equality of all other human beings because, through cultural innovations over the centuries, we have expanded the circle of cooperation. After all, it's frequently said that we are all linked together now in a global economy. If morality is just an evolved device for facilitating cooperation, then

as more people are included in cooperation, then the moral circle will expand to include them.

That explanation doesn't work. Most obviously, it doesn't explain the Second Great Expansion, the fact that many people now recognize that some nonhuman animals have moral standing—that they count morally in their own right—even though they aren't and never will be participants in our cooperative schemes. The idea that morality is all about relations among cooperators might explain why, if you happen to be a shepherd, you might think you ought to treat your sheepdog reasonably well; but it can't explain why you think you shouldn't consume factory-raised poultry. In fact, it can't even explain why you would believe you shouldn't burn cats—who are notoriously noncooperative—if doing so gives you pleasure. It also can't explain why human beings to whom we are not bound by sympathy based on kinship or close association, and who are so severely disabled that they can't participate in our cooperative schemes, are nevertheless now widely regarded as having the same basic human rights as those who can participate in cooperation and those we happen to be attached to.

In addition, the fact that human cooperative networks are now global does nothing to explain why some (though not enough) of the world's most fortunate people try to alleviate the deprivations of distant strangers with whom they will never interact at all, much less cooperate. Nor can it explain why anyone would come to believe that they should try to include those people in the cooperative schemes they participate in when they don't really need to, rather than excluding them from cooperative arrangements altogether or dictating terms of "cooperation" that deny them anything approaching basic equal moral status. In chapter 3, I'll have a lot more to say about why the expansion of cooperation doesn't explain the enlargement of moral regard represented by the Two Great Expansions.

## **The Puzzle Deepens**

It's not just that evolutionary thinkers, by focusing on the origins of morality, haven't tried to solve the Big Puzzle (or even acknowledged

that it needs to be solved). In addition, some of the best work by evolutionary developmental psychologists, including Tomasello and the creative students he has nurtured, actually makes the Big Puzzle bigger. That work emphasizes that very young children tend to learn whom to cooperate with—and to do so in a way that demonstrates moral values like fairness—by noticing certain resemblances between themselves and others. Even children as young as four months interact morally and cooperatively with individuals who are like themselves in terms of ethnicity, language, and dress, but not with people who don't resemble them in these ways (Tomasello 2016, 94).

These criteria for partner choice make sense, because whether someone is like you in those ways may be a fairly accurate predictor of whether they will be enough like you in their beliefs and values to make mutually beneficial cooperation with them work. Also, the more someone is like you, the more experience you will have had with people like them, and the more likely it is that you'll be able to predict whether they are genuine cooperators or free riders, whether they'll abide by simple rules of reciprocity.

Yet it's clear that the resemblances that Tomasello's cadre and other theorists of moral learning have identified do *not* provide a reliable guide when it comes to knowing whether someone who is quite different from you could be a valuable cooperative partner. More importantly, the fact that humans rely on those resemblances to determine who would make a reliable partner in cooperation does nothing to explain why a lot of people nowadays would ever have come to believe that other individuals are worthy of basic moral respect *regardless of whether we do or ever will cooperate with them*.

On the contrary, relying on those sorts of resemblances can lead us to conclude that the Other is not worthy of respect. That's how racist cues (like skin tone and hair texture) work to promote tribalistic morality, rather than overcome it. So this kind of valuable scientific work on how children become moral cooperators doesn't solve the Big Puzzle; *it deepens the puzzlement*. That fact should make one wonder whether there is something very wrong about trying to understand the Two Great Expansions as progress in *cooperation*. More fundamentally, it should make one wonder whether morality is really just about what makes cooperation work, whether it is

really nothing more than a device for successful cooperation. Maybe morality now, for some people, is more than that, even if that's all it was originally.

### **Morality Is More Than What It Originally Was**

Many traits, across the whole range of organisms, came about because they helped the organism do something that promoted its reproductive fitness. However, these traits can also produce other results, including reductions in fitness. The philosopher Stephen Stich gives a striking example: the frog's lightning-fast, elongated tongue may have evolved to enable it to catch insects and to gobble up their small round eggs, because that enhanced its reproductive fitness. But some frogs have been observed to flick out their tongues to grab and swallow BBs (small round lead pellets fired from air guns and shotguns)—and that's not good for them.

Evolutionary science offers lots of other examples of items that evolved to do one thing but are capable of doing a lot more. These include human genitalia, which evolved to enable reproduction but are now employed for a wide range of activities that have nothing to do with generating new human beings (including, in the case of some edgy conceptual artists, using them to paint pictures). Similarly, the cognitive architecture of the human brain evolved to be what it is because it enhanced our reproductive fitness by helping us solve problems in the ancestral environment that we needed to solve to survive; but now it can be used to construct astrophysical theories, write symphonies, and play video games.

If you confuse the statement that the moral mind came about because it facilitated cooperation (and thereby enhanced reproductive fitness) with the statement that everything of importance about human morality can be explained by showing that it facilitates cooperation, you'll be making a huge mistake. More specifically, you won't be able to explain the Two Great Expansions by trying to show that people whose moralities include those enlargements of the circle of moral regard have an advantage in cooperation. If you make the mistake of thinking that, because the basic elements of human moral psychology are an adaptation for cooperation,

everything about human morality can be explained by showing how it contributes to cooperation, then you'll be blind to the possibility that some features of the moralities of some human beings require a different kind of explanation. You won't be in a good position to explain significant moral change that isn't just a change in the solutions that morality offers for cooperation problems. And you won't be alive to the urgent question of whether your society is moving forward or backward, morally speaking, when it comes to inclusion versus exclusion.

Here's another reason to be careful not to slide from "moralities came to be because they facilitated cooperation" to "morality is a type of cooperation" or "morality is all about cooperation." Some moralities, including ancient Stoic morality and most forms of Buddhism, encompass ideals of individual excellence that have nothing to do with cooperation. In fact, some moralities actually advocate withdrawing entirely from cooperation, presenting the life of the meditative hermit as the ideal of moral excellence. If the die-hard cooperation dogmatist replies "Oh well, but those aren't really moralities," he'll have to justify that claim without indulging in definitional tricks and circular reasoning. It won't do to say that genuine moralities are all about cooperation, and that hence Stoicism and the like aren't moralities, because that just assumes precisely what is at issue.

### **Cautious Darwinianism**

My goal is to take human biological evolution seriously, without being a biological determinist. To do so, I'll have to give both culture and biology their due in the story of how some moralities became more inclusive. My explanatory framework for large-scale moral change will assign a significant role to biological evolution—but only up to a certain point, the point at which the moral mind emerges. After that, understanding the interplay between the evolved flexibility of the moral mind and cumulative culture becomes the main resource for explaining large-scale moral change.

The trick is to show how "descent with modification" has occurred in human moralities despite the fact that the basic features

of human moral psychology that together constitute the moral mind have probably not changed much, if at all. Let me rephrase: the trick is to show how the features of the moral mind, *combined with our capacity for niche construction*, explain *why* human moralities change, and more specifically why certain large-scale changes have occurred. The specific explanatory task, which I'll use to make some more general points about how to explain moral change in a wider range of cases, is to provide an account of how the same moral mind that produced tribalistic moralities under the selective pressures of the earliest environments has only very recently been expressed in more inclusive moralities.

### **An Overview of the Investigation**

This chapter has added detail to my initial description of the Two Great Expansions and shown why appreciating these amazing transformations—and, more generally, understanding moral change—requires getting the evolutionary science about “moral origins” right. In chapter 2, I set out, in the needed detail, the standard attempt to understand human morality in evolutionary terms, listing the main features of what evolutionary thinkers call the Environment of Evolutionary Adaptation (EEA), the set of ancient conditions that created the pressures of natural selection that forged both the moral mind and the first distinctively human moralities.

Notice that I just said “that forged distinctively human moralities” (plural), not “that forged a distinctively human morality” (singular). Most evolutionary scientists and nonscientists who take evolutionary theory seriously refer to the origins of human morality, not moralities in which it was expressed. That's not good, for two reasons. First, it's pretty likely that a distinctively human kind of morality didn't develop in just one early human group and then spread to all others. On the contrary, that new kind of morality probably originated in a plurality of human groups. And we have every reason to believe that these original moralities were not identical; they surely varied, depending on features of the local environment in which they developed. That's one reason to talk about the origin of human moralities, rather than human morality.



Even if you assume that there was one original human morality, a second reason still applies: if you refer to human moralities, you'll avoid the ambiguity of the term "morality," which can refer either to the moral mind or to some morality or moralities in which the moral mind is expressed. In other words, using the phrase "the origins of human morality" or "moral origins" encourages the confusion I've shown must be avoided: the confusion between the character of the moral mind and the character of the moralities it initially underwrote.

Chapter 2 explains exactly how the standard evolutionary "moral origins" story set out in chapter 1 generates the Big Puzzle, the apparent inconsistency between our supposedly tribalistic moral nature and the Two Great Expansions. Chapter 3 canvasses some contemporary attempts to explain moral change in the direction of inclusion and shows why they don't succeed. Chapter 4 goes back to square one, providing a revisionist account of the origins of human moralities in the EEA, the environment in which distinctively human moralities supposedly first appeared.

This revisionist account doesn't so much solve the Big Puzzle as begin to dissolve it. I argue that it is likely that there wasn't just one EEA; there were variations in the early ancestral environment, and these variations fostered a special kind of flexibility: the ability to respond to strangers in tribalistic ways when that kind of response was required for reproductive fitness, but in inclusive ways when being inclusive was conducive to fitness.

The key to understanding how the character of morality can vary depending on features of the environment is to appreciate the connections between (1) the distinctively human capacity for cumulative culture and three other important human traits: (2) the impressive flexibility of moral sentiments, beliefs, and behaviors in response to environmental changes that the moral mind allows, in particular the ability to extend sympathy to "strangers" and at least some nonhuman animals; (3) the capacity for niche construction (the ability to create new environments in which we are subject to new pressures for natural selection and in which different potential moralities can be realized); and (4) moral identity (the powerful motivation to convince others and ourselves that you take moral norms seriously,

which you have to do if you're to be regarded as a reliable partner in cooperation). Working together, these four features explain how human beings became supercooperators—a species whose capacity for cooperation is much more robust and flexible than that of any other species.

Chapter 5 explains these four features, describes how they work together, and begins the task of developing an explanatory framework for large-scale moral change, something that is essential if we are to understand how moral progress can come about and how it can be protected from regression. Focusing on the relationship between the flexibility of human moral responses and the cumulative cultural capacity for niche construction, I show how humans first made the transition from exclusive, purely tribalistic moralities to *shallowly inclusive moralities*, moralities that facilitate cooperation with people of different ethnic or cultural groups and extend to them the *limited* sort of moral regard that such cooperation requires.

Shallowly inclusive moralities fall short of the *deep inclusion* manifested in the Two Great Expansions, but they take a necessary step in that direction. The basic idea of chapter 5 is that, through cumulative cultural innovations, human beings constructed new niches in which they could *afford* to develop more inclusive moral beliefs, concepts, rules, and attitudes—environments in which extending limited moral regard to strangers no longer risked reproductive suicide. I also show how having a shallowly inclusive morality actually is highly beneficial, in both social and reproductive terms, again in certain environments. The fundamental message here is that a human being who is capable of both exclusive morality and shallowly inclusive morality is more adaptable than one who can only respond morally in an exclusive, thoroughly tribalistic way.

Chapter 6 fills out the solution to the Big Puzzle by developing in more detail what cognitive capacities, motivations, and social conditions are required for people to incorporate the Two Great Expansions into their moral orientation, to go beyond shallowly inclusive to deeply inclusive moralities. Here I focus a good deal of attention on two remarkable human traits that most likely were necessary for the radical moral reorientation that the Two Great Expansions

represent: the capacity for critical, open-ended moral reasoning and the powerful motivating force of moral identity. Critical, open-ended moral reasoning is crucially important for moral change because it enables us to challenge the moral status quo. Moral identity is crucial because it can give us the courage to exercise our capacity for critical, open-ended moral reasoning and act on the results.

In this sixth chapter, I also explore the question of whether we can provide an evolutionary explanation of the existence of the capacity for critical, open-ended moral reasoning and, more important, whether there is an evolutionary explanation of its coming to be exercised by some human beings in a way that leads to a profound rethinking of which beings have moral standing and which have the highest, equal moral status. In addition, I explain why anybody would be *motivated* to exercise the capacity for critical, open-ended moral reasoning in a way that leads toward the Two Great Expansions. The answer to this question is that the source of the needed moral motivation is moral identity: the robust motivation to be, and to be regarded by others as, moral.

There's a rather broad consensus among evolutionary scientists that the powerful motivation to sustain moral identity evolved because it was vital for the kind of cooperation that only humans engage in—and for the reproductive success of individual humans, whose survival depended on their being regarded by their fellows as desirable and reliable partners in cooperation. Under the right conditions—in the new niches that humans constructed using their capacity for cumulative culture—the motivational power of moral identity could and did lead some human beings toward the Two Great Expansions.

Chapter 7 examines regression to tribalistic, exclusive moralities. Given the increases in racism and xenophobia revealed by recent elections and public political discourse in the United States and much of Europe, this is a topic of more than theoretical interest. I argue that the key idea to understanding regression to tribalism is this: if people actually find themselves in conditions that are like the elements of the EEA that were hostile to inclusive morality—or *if they become convinced that they are in those conditions even when they aren't*—they tend to revert to tribalistic moral thinking and behavior.

Chapter 8 shows how tribalistic moral thinking and behavior have evolved since their earliest manifestations: how creatures who previously often excluded from full moral status only actual strangers—members of other societies—have come to practice elaborate forms of “intrasocietal tribalism,” relegating other groups *within their own society* to an inferior status. This happened in slave societies and caste systems, but it also in the United States and other liberal constitutional democracies like Canada, Sweden, and Denmark during the era of eugenics, when certain members of society were ruthlessly segregated and subjected to forcible sterilization because they were thought to harbor defective genes that were responsible for all major social problems. Another, more recent American example of intrasocietal tribalism is the hysteria during the early 1990s over the threat posed by an allegedly distinct class of urban African American men dubbed “super predators”—a term that evokes primal EEA-like fears if there ever was one. Today’s deep division of American society between “liberals” and “conservatives” is also a form of intrasocietal tribalism.

Chapter 8’s account of the evolution of modes of intrasocietal tribalism relies on the same evolutionary story that chapters 1 and 2 told to explain why our remote ancestors often exhibited exclusive moral attitudes toward members of other societies. But it adds something new and important: it shows how *ideology*, a more recent cultural phenomenon, though one that is grounded in our original evolved moral nature, makes this new kind of tribalism work—and why it is so hard to stop it. In chapter 9, the book’s final chapter, I sum up the main conclusions of my investigation, including several quite general insights about the possibilities for large-scale moral change that the detailed examination of the Two Great Expansions has yielded. Here I also offer guidance for how to resist tribalistic regression, focusing on the key role of institutions in determining whether the social environment is friendly to inclusion or to tribalism. Most importantly, I make the case for taking seriously the fact that how the human environment is shaped—and by whom—determines what sort of morality will be dominant in a society and whether individuals will be progressive moral agents or stunted specimens of what they could have been.

If I'm right that, given the flexibility of the moral mind and the cumulative cultural capacity for niche construction, the space of possible human moralities is quite large, there is good news and bad news. The good news is that we are not morally tribalistic by nature: whether tribalistic moralities dominate depends on what sort of niches we create for ourselves. The possibilities not just for moral change but for moral progress may be so great that we can't even imagine them.

The bad news is that it's misleading to say that *we* construct the niches whose character determines the possibilities for human morality. The process of niche construction is far from democratic. The sobering fact is that some people—a very small proportion of the human population—exert disproportionate influence over the character of the niches that humans now occupy, because power is distributed extremely unevenly among human beings. The powerful few wield disproportionate influence over what sorts of moralities are likely to be dominant and over what sorts of moral agents we are likely to be. The problem of domination that the earliest human societies had to solve was simple: how could the group control bullies who were physically aggressive toward other members of the group and hogged resources? In modern societies, the problem of curbing domination is more complex because new, more complex forms of domination have evolved through cultural processes. (In fact, one could interpret human history as a coevolutionary arms race between modes of domination and modes of counterdomination, an idea that's the object of my current research.) One crucial aspect of the modern problem of how to suppress domination has gone unnoticed: how can we develop institutions that prevent some people from shaping the moral environment in ways that prevent us from being as good as our moral minds make us capable of being?

It's a pretty safe bet that most of the people who wield the most power over niche construction aren't guided by any thoughts about how the niches they construct will influence the character of human moralities. On the contrary, their aims and motivations for wielding that power are typically shortsighted, often self-interested, and utterly uninformed by any scientific understanding of how the moral mind expresses itself differently in different environments.

While some hierarchy, some inequality in power—if it is multiple, and the various hierarchies are independent of one another—may actually be necessary for moral progress, as I'll argue later, hierarchy also carries a great risk. If hierarchies enable some people to shape the moral environment in ways that make it hostile to inclusion or other forms of moral improvement, then inequality will hinder moral progress and create a risk of regression.

If my explanation of large-scale moral change turns out to be valid, it won't be correct to say, as many people do, that culture enables humans to “stretch the evolutionary leash” of a fundamentally tribalistic human moral psychology. Instead, the right conclusion will be that our flexible moral mind, when combined with our capacity for cumulative culture and the ongoing niche construction that it facilitates, enables us to be *both* tribalistic and exclusive, depending on the human-made niches we occupy. In other words, some human beings have developed inclusive moralities, not *in spite of* our evolved moral psychology, but *because of it*—in particular, because of its flexibility in the face of different environments. If one wants to speak of human moral nature, one should say that it is dualistic, encompassing the potential for both inclusive and exclusive moralities.

This book, then, is a detective story. Like any good detective story, it starts with a mystery to solve, the Big Puzzle: how could creatures whose first moralities were tribalistic come to have deeply inclusive moralities? As in any good detective story, the route to solving the mystery will include false leads, dead ends, and flashes of insight that occur only when we question the assumptions that originally framed our understanding of the mystery itself. Solving (or dissolving) the mystery will take a great deal of effort on my part and considerable patience on yours, but it will be worthwhile, because solving it is vitally important—assuming that we care about what it means to be human, about what it is to be moral, and about what the prospects for human moralities are.

Anyone who is curious about what human beings are and how they got that way should find my detective story interesting—if I can tell it in a sufficiently engaging way and convince you that my beginning of a solution makes sense. This is a detective story

about you and me, not just about some abstraction called *Homo sapiens*. Morality is not something that is out there and with which we engage while remaining the same. It isn't an alien force that constrains us from without; it lies at the core of our being. Our morality in large part determines who we are. And—though unwittingly and utterly without design—we have determined what our moralities, and hence ourselves, are like, by determining what sort of social environment we live in.

As the detective story unfolds, it will become clearer just what I mean by “tribalistic moralities.” But to avoid a misunderstanding that could abort the investigation before it even gets started, let me say right now what I *don't* mean by that phrase. I *don't* mean just any morality that includes a prominent role for group-based identity. For reasons that I'll explain later, I think that humans generally have a strong desire, indeed a need, for identity, for a sense of who they are, and that an important element of identity, for practically everyone, is a perception of oneself as belonging to a group (usually more than one). The question I focus on in this book is not “how could we have evolved to attach so much importance to the group-membership aspect of our identity?” though I think that question is profoundly interesting. Instead, I home in on a particular kind of group-membership identity, one that involves a serious kind of moral exclusion of the Other.

Many, perhaps all, kinds of group-membership-based identities not only involve a distinction between Us and Others but also tend to regard Us as superior in some respect. That's not what I mean by “tribalistic moralities.” The kind of group-membership identity I target goes beyond that. It represents the Other in such negative terms that it becomes an obstacle to achieving the First Great Expansion.

### **What Tribalism Is and Isn't**

A pair of concrete examples will clarify what I mean and don't mean by “tribalistic moralities.” If you think that in general the francophone Swiss are more cultured or nicer than Swiss people whose primary language is German, that doesn't mean you have



a “tribalistic morality” (though it probably does mean you’re a francophone Swiss). Thinking that foreigners suspected of terrorism may justifiably be tortured because they aren’t Americans and hence aren’t morally entitled to the legal protections of the US Constitution does show that your morality is tribalistic. That kind of thinking means that you have not yet made the First Great Expansion, or you’ve regressed from it, because it implicitly denies that foreigners suspected of terrorism have fundamental human rights and assumes that the only question is whether they have the particular legal rights that are specified in the Constitution. People who think like that just don’t get the idea of *human* rights, or their moral reasoning in this case displays a pretty gross inconsistency. Perhaps their fear in the wake of 9/11 or other terrorist attacks has disabled their ability to apply the idea of human rights consistently in the case at hand.

As I noted in the preface, there are somewhat less extreme forms of tribalism than those that deny that members of some other groups have even the most basic human rights (including the right not to be tortured). But the most extreme and milder forms of tribalism share certain common features: the tendency to clump individuals and issues, the tendency to think in grossly simplified, black-and-white terms, and the assumption that we are in a winner-take-all struggle for the highest stakes in which compromise is not an option because of the degenerate character of our opponents and the momentous moral disagreements that divide us from them.

Given our evolved nature, I don’t think it’s feasible or desirable to disregard group-membership identity or to try to prevent it from playing a significant role in morality. Yet I think that reducing the power of the particular kind of group-membership identity that I call “tribalistic morality” is an extremely important undertaking, and achieving this goal is an especially significant kind of moral progress.

Before we plunge into the investigation, a disclaimer and a plea for tolerance are in order. The disclaimer is that I am not providing a thoroughly scientific theory. Much of what I’ll say in developing a solution to the Big Puzzle will be speculative. That’s not surprising, given that scientific thinking about the evolution of morality



is new and quite undeveloped—and that it hasn't yet extended to grappling with the Big Puzzle, because it has been restricted to the origins of moralities, not their ongoing evolution. So my attempt to solve that mystery won't be "hard" science by a long shot.

Nevertheless, it is worth emphasizing that in the early stages of any line of research there is a legitimate role for speculation—a larger role than when considerable progress has been made. Speculation in the early stages is acceptable and even necessary, but only if it is responsible. That means two things: first, it must be consistent with the best existing theories that are relevant to its subject matter; and second, it must be liable, eventually, to being disconfirmed or confirmed by solid empirical research. I believe my thinking satisfies those criteria. If I'm right, then my speculations merit your tolerance.

