

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>

Taking Charge of Our Moral Fate

I will not apologize for the speculative character of the attempt. At this stage, either the question is answered in a vague, fragmentary and tentative way, or it must be left alone: there is not enough sound theorizing and well-regimented evidence in the domain to do otherwise.

—Dan Sperber, *Explaining Culture: A Naturalistic Approach*

Our detective story has been unusually complicated, even for the genre. I will conclude our investigation, first by outlining the central argument of the book in a few steps, then by providing a brief summary of the thinking that supports the central argument, and finally by drawing out some unexpected implications of my analysis.

Here's the central argument:

- (1) Human beings evolved to have a highly flexible moral mind, a general set of competencies for having moralities that can be expressed in very different ways, depending on the social-environmental inputs that stimulate the exercise of those competencies.
- (2) The environment in which the moral mind first manifested itself and the environments that have existed throughout most of human history were profoundly different from the niches that humans have only recently constructed for themselves: in the last three hundred years, some human societies have achieved high levels of surplus reproductive success, solving the problem of achieving sustainable reproduction exceptionally well.
- (3) Where those favorable conditions prevailed, surplus reproductive success allowed the Great Uncoupling: morality was no

longer the slave of fitness, and the powers of the moral mind were unleashed to produce moralities that extended beyond the facilitation of cooperation and were more inclusive, less tribalistic than earlier moralities. The remarkable human capacity for critical, open-ended moral reasoning and the potent motivating force of moral identity fueled this transformation, with the result that hitherto pervasive moral assumptions were revised or abandoned as the circle of moral regard expanded.

(4) The key factor that produced surplus reproductive success and allowed the liberation of morality from the demands of fitness was the human capacity for cumulative culture; through niche construction, it forged institutions that made inclusive moral responses more affordable in reproductive and material terms and created new opportunities for the exercise of the capacity for critical, open-ended moral reasoning, which in turn facilitated changes in moral identities and helped redirect moral emotions such as sympathy and disgust to new objects.

(5) Until now, the processes by which social environments interacted with the moral mind to produce moralities and moral agents were undirected, not subject to deliberate human control; our moral fate was the plaything of morally blind forces, beyond our ken and therefore beyond our control.

(6) If humans learn enough about the moral mind and the interactions between it and specific environmental features, we can in principle take charge of our moral fate: we can exert significant influence on what sorts of moralities are predominant in our societies and what sorts of moral agents we are. Doing so would be perhaps the highest form of human autonomy. It would also be the most profound kind of creativity: the creation of the moral self in a species for whom the moral self lies at the core of our being.

Now to summarize the thinking that produced my central argument. The moral mind is highly flexible: it can generate radically different moralities, depending on the character of the stimuli that different social environments provide. Some social environments are conducive to more tribalistic moralities, some to more inclusive ones. Different inputs to the moral mind mean different

outputs. Some social environments afford greater possibilities for moral progress; in others the odds are against it. So the Tribalism Dogma is wrong: humans don't have a tribalistic moral nature. Our moral nature encompasses the capacity both for tribalism and for inclusion.

Different social environments produce not only different moralities but also different kinds of persons, beings with different moral identities. Some social environments are conducive to humans being the best they can be, given human moral nature; others stunt them. The point is not that the social environment by itself determines our moral possibilities; rather, those possibilities are shaped through the *interaction* of the social environment and the highly flexible moral mind. That interaction doesn't just produce new or better solutions to cooperation problems. Even if morality first was nothing more than a bundle of solutions to cooperation problems, it has become more than that now. So the Cooperation Dogma is just as mistaken as the Tribalism Dogma. Once we abandon the intellectual straightjacket of the tribalism and cooperation dogmas, we can think clearly about morality and about the possibilities for moral change. And we can begin to prepare ourselves to embark on the project of taking charge of our moral fate.

Replacing the tribalism and cooperation dogmas with the recognition that human moral nature is highly flexible transforms the idea of moral progress. We can no longer assume that progress is inevitable or that whatever gains we have made will endure. Progress is neither more nor less natural than regression, and we mustn't take it for granted.

It is also vital to understand that the social environment that the moral mind interacts with includes whatever morality happens to be pervasive at any given time in a society. That's one important environmental factor and it must be taken into account when thinking about how to tweak or transform the social environment to maximize the probability of moral progress and reduce the risk of regression. Making moral progress in a social environment that is already saturated with tribalism is clearly more difficult than starting from one in which the pervasive morality already includes

progressive elements. This gives us another reason to try to avoid regression.

Once human beings used their capacity for cumulative culture to create niches in which they achieved sufficient surplus reproductive success, the Great Uncoupling occurred: the content of moralities was no longer determined by the demands of reproductive fitness, the space of possible moralities expanded greatly, and deeply inclusive moralities became possible. Neither the Great Uncoupling nor the possibilities of moral progress it created resulted from any conscious project of improvement; they were a matter of highly improbable moral luck.

Control over the social environment, including the features that are critical for the character of morality and the moral development of individuals, is far from democratic: a small portion of the population has disproportionate control over the social environment and hence over what our moral life is like. Yet no one—neither those exercising control nor those who are affected—has paid attention to this reason for being troubled by the growth of inequality in wealth and the inequality of power it inevitably entails. We all know that there is an ongoing contest for power in society; what we have failed to see is that it's ultimately a contest for who will determine our moral fate.

We commonly think of designing institutions for economic efficiency or for well-functioning political processes, but we pay too little attention to the moral effects of different institutional designs. Once we understand that a society's morality and the moral development of individuals within it are the product of the interaction between the moral mind and the social environment, it becomes imperative to extend the idea of scientifically informed institutional design to encompass *moral* institutional design—to think hard and systematically about how the character of institutions either promotes or hinders moral progress. For the first time in our species' history, it becomes possible to liberate our moral fate from the dominion of blind chance and shape it by scientifically informed choice.

To realize this possibility, we need the best scientific minds to focus on constructing a theory of moral change grounded in an

understanding of the interaction between the moral mind and the specific characteristics of human-made environments that shape its expression in particular moralities. That project will require increasing both our knowledge of what the moral mind is like and our knowledge of how its potentials are realized under specific environmental conditions.

I've marshaled support for the dependence of the character of moralities and moral agents on social environmental factors by examining one kind of large-scale morally progressive change that is extremely interesting in its own right, especially for those of us concerned the threat of tribalism: the Two Great Expansions. My investigation showed *how* the creation of new niches in which humans achieved surplus reproductive success enabled the Great Uncoupling—the liberation of moralities from the demands of reproductive fitness—which in turn created the possibility of progressive moral change in the direction of inclusion. Solving the Big Puzzle of how a type of great ape with initially largely tribalistic moralities could have come to have deeply inclusive moralities demonstrated the flexibility of the moral mind in response to different social environments.

The explanation I offered of how the Two Great Expansions came about also supplied a theory of moral regression to tribalism. That theory highlights the fact that tribalism, like inclusive morality, is still evolving. I showed how a new kind of tribalism, intrasocietal tribalism fueled by deeply divisive ideologies, threatens to hollow out the First Great Expansion. My hypothesis was that deeply divisive ideologies are an adaptation (a product of cultural selection) for competition among groups within society under modern conditions and where there is a widespread belief that democracy has failed. That hypothesis is compatible with the thesis that intrasocietal tribalism is contributing to the decline of democracy. We may be witnessing a vicious circle: disappointment in democracy may make tribalism seem like the only alternative, and tribalism may in turn make pessimism about democracy a self-fulfilling prophecy.

The complex investigation I have just summarized yields several big conclusions. First, evolutionary science doesn't tell us that large-scale moral change is impossible, or that large-scale change

in the direction of greater inclusion is against our nature, or that we've reached the limit of inclusiveness. Instead, its message, properly understood, is that we have a highly flexible moral mind and a capacity to respond either tribalistically or inclusively, depending on the environment. A sound understanding of the origins of human moralities—one that rejects both the Cooperation Dogma and the Tribalism Dogma—implies that to secure further moral progress and reduce the risk of regression, we should develop a scientific, empirically grounded account of how to shape our social environment (and above all how to design institutions) to foster inclusion and, more generally, to create the conditions for achieving and sustaining moral progress. We can no longer rationalize our tribalistic failings with the comforting thought that they're "just human nature." We cannot escape the fact that we are much freer than the Tribalism Dogma portrays us as being. It's up to us whether human societies are predominantly tribalistic or something better.

Second, at this early stage in the attempt to use evolutionary science to explain large-scale moral change, it's not clear whether explanations that are thoroughly evolutionary will suffice. If we assume that the moral mind itself won't change significantly for the foreseeable future, any explanations will necessarily be largely cultural, not biological. Whether the cultural explanations will be genuinely evolutionary through and through may not be ascertainable at present, because the most robust and mathematically rigorous cultural evolutionary explanations are not applicable to the complex, messy phenomena of large-scale moral change. A practically useful understanding of the conditions that determine our moral fate must rely on evolutionary science; but other disciplines, including anthropology, history, sociology, economics, and social psychology, will most likely also play important roles.

Third, large-scale moral change can come about in two quite different ways: either because the change promotes reproductive fitness (whether biological or cultural), through selection operating independently of human intention or design; or as a result of "fitness-independent" factors—under conditions of surplus reproductive success that allow for an expansive set of options for intentional moral change. As an example of the first way, consider the

development of dominance-suppression techniques and norms of fairness and reciprocity that apparently brought about moralities that differed from those of our nearest primate relatives. We have no reason to believe that this type of moral change is over; it could occur again. It all depends on what environmental pressures determine which genes and social practices are transmitted to the next generation and which aren't. The second way that moral change can come about is illustrated by the historical narrative I've provided to explain the Two Great Expansions. Moral change in this case is not driven by reproductive fitness, though of course it can only come about and persist if it doesn't disastrously undermine fitness. In this kind of moral change, "blind" processes of natural or cultural selection play a crucial role but aren't sufficient: they create opportunities for human beings to exercise their moral powers in new ways and to bring about change in part through intentional actions aimed at realizing their evolving understandings of what morality and fidelity to their own moral identities require.

Fourth, large-scale moral change of the second sort—change that doesn't come about because it promotes reproductive fitness—is much more likely to occur under conditions in which highly effective mechanisms for compliance with existing moral rules are *not* in place. This is because moral change requires, in its initial stages, noncompliance with the existing rules. In cases of profound moral change like the Two Great Expansions, it may also require the transformation of the individual's moral identity. Social conditions must permit that kind of change to occur, which in turn requires that religious or secular authorities or less formal kinds of social pressures are not able to stifle the processes that lead people to reconceive their moral identities. Social conditions must permit people to revise their conception of what it is to be a moral person.

Evolutionary scientists emphasize that in their original forms, human moralities only survived if they succeeded in curbing free riding (of both the loafer and bully varieties), reducing conflicts among group members, and coordinating their beliefs and behaviors as to how things ought to be done. In brief, the earliest moralities brought to bear potent pressures for a high degree of moral conformity, since individuals' viability as partners in cooperation

and hence their reproductive success depended on it. Similarly, groups could only survive if their cultural practices successfully produced a high degree of conformity with their moral rules. Moreover, the same scientists stress that these early groups were egalitarian, at least in the sense that they had highly effective mechanisms for preventing any one individual or small subgroup of individuals from determining how things were done. That meant that no one could unilaterally change the rules.

Under these conditions, major moral innovation that was anything other than a “blind” response to selection for reproductive success (whether biological or cultural) would have been extremely difficult, if not impossible. Any would-be moral innovator would most likely be suppressed, pressured into returning to conformity with the moral status quo, and the nonhierarchical character of these earliest societies would have prevented any individual or small group of individuals from simply imposing new moral rules on their fellows.

Something had to change for large scale moral change to occur that wasn’t a response to the changing requirements for reproductive success. Societies developed new modes of cooperation that worked well enough, even though the mechanisms for ensuring rule compliance no longer functioned as effectively as those that ensured moral conformity in the earliest societies. Cooperation became so productive that it could thrive even if some people violated the existing moral rules or refused to acknowledge their validity because they thought the rules were flawed or simply pointless.

Early human groups depended on virtually every able-bodied person participating in cooperation and being committed to following the moral rules that facilitated it. In contrast, modern societies can afford more free riders or people who are otherwise disaffected without ceasing to function, because productivity has increased so greatly. This too makes moral change easier: people who have not thoroughly internalized norms in the first place or who have become critical of them may be more amenable to modifying them or abandoning them in favor of new norms. Further, if it becomes apparent that society can function well even if some people violate the rules, the question of whether the rules are really necessary can

arise. Here's one example among many: when homosexual behavior was first decriminalized in the UK, some people predicted a collapse of the moral structure of society. It didn't happen. There's no good evidence that compliance with the most basic moral rules diminished because of that particular change. Similarly, it turned out that moral anarchy didn't result from women taking charge of their own sexuality and abandoning traditional ideals of feminine "chastity." Either compliance with these norms didn't play as important a role in social cooperation as the doomsayers thought, or else these norms *had* previously been important, but society was resilient enough to adjust to the change in ways that avoided a breakdown of cooperation and an unraveling of the moral fabric on which cooperation depends.

It's a truism that modern societies afford more scope for individual freedom and the development of individuality than traditional societies. To a significant extent, that's a result of the *failure* of modern societies to achieve high degrees of moral conformity. That "failure" may be a necessary condition for large-scale moral change that is not simply a way of promoting reproductive fitness: moral change that begins with someone deliberately challenging the moral status quo in a context where reproductive fitness is simply not an issue.

One of the most powerful mechanisms for achieving moral conformity in earlier societies was the threat of exclusion from cooperation. As societies became much larger and more complex, it became easier to be a moral innovator because exclusion from one's present cooperative relations wasn't so costly: complex societies, so long as they are not too centralized and tightly controlled by a religious or secular authority, offer alternative opportunities for cooperation, more options for partner choice. In contrast, in the earliest societies, the costs imposed on individuals who abandoned existing norms in favor of new ones were generally prohibitively high, because exclusion from cooperation in one's group meant exclusion from cooperation—and exclusion from cooperation was disastrous.

Another surprising implication of this contrast between the earliest and much later societies is that while the egalitarianism of the former created an impediment to moral change, the inequalities of social and political power characteristic of the latter can facilitate

larger, more rapid moral changes. Individuals with greater power can more easily resist the social pressure to conform to existing norms, so they can afford to initiate moral change. In addition, their greater power may enable them to influence others to follow suit. The “failure” of modern societies to avoid hierarchy—their relative lack of success compared to the avoidance of hierarchy that the earliest societies achieved—contributes to their greater possibilities for moral change, including change that we may reasonably regard as progressive.

Consider another related but much more general point. As I’ve already emphasized, in the earliest human groups, a high degree of agreement on moral rules was probably necessary for successful cooperation. There was strong selection pressure on all members of the group to converge on the same set of rules because successful cooperation depended on it. But in modern societies, successful cooperation often does not require thoroughgoing moral agreement. Under modern conditions, considerable moral disagreement is compatible with successful cooperation, because the rules that govern cooperation are accessible to people with widely different moral views on other matters. In other words, the moral rules that facilitate cooperation do not saturate the whole of life. For example, in most circumstances, one can participate well in running a corporation regardless of whether one is a Catholic or a Jew or a Muslim or an atheist and regardless of whether one thinks abortion is wrong or morally permissible.

The bottom line is that in modern societies, cultural innovations that produced new niches have dramatically lowered the costs, at least for some individuals, of departing from the moral status quo. Such conditions greatly expand the possibilities for moral change. This was the lesson of my discussion of moral pioneers in chapter 6.

A trade-off thus exists between moral conformity and moral change. Too little moral conformity, and cooperation breaks down. Too much moral conformity, and moral change becomes difficult, if not impossible. Where you think the balance should be struck will depend on how defective you think the moral status quo is and how confident you are that it can be improved.

The perennial debate between progressives and conservatives (and revolutionaries and reformists) is largely a disagreement about how much moral change is compatible with sufficiently stable cooperation. Since moral change requires moral disagreement, at least in the beginning, this means that conservatives and progressives disagree fundamentally about what the optimal extent of moral disagreement is. Conservatives assume that if cooperation is to be stable, there must be very little moral disagreement; progressives assume that considerable disagreement (at least in the short run, during the process of moral innovation) is compatible with stability or at least with relatively undistruptive social change. In the absence of a much more developed theory of moral change than is available today, both assumptions are little more than articles of faith. Without such a theory, the conflict between progressives and conservatives can't be resolved. That helps explain why it hasn't been resolved after all these centuries.

Without pretending to settle that debate, this book sheds new light on it. Conservatives think that successful cooperation requires that there be very little moral disagreement because they think of a society's morality as a seamless web. In other words, they believe that moral rules and practices are densely interconnected, so that if you change one item, you create an unacceptable risk that the whole thing will unravel. Progressives tend to think that you can change some things without running the risk of changing a lot more than you want to. They think that a morality is less densely interconnected than that. A metaphor more congenial to the progressive way of thinking is a loosely woven fabric with substantial seams or, in evolutionary biology terms, a modularized organism.

In this case, the progressives have evolutionary science on their side: no organism or society could long survive if it were literally like a seamless web. Adaptability through selection, whether biological or cultural, requires that some things can be changed without everything changing (Lewontin 1978, 215–216). A seamless organism or a seamless society would be too fragile, too likely to collapse in the face of random changes in its own components or external shocks. It couldn't adapt.

Nonetheless, even though the seamless-web metaphor, if taken too seriously, gives an unduly pessimistic picture of the possibilities for reconciling stability with progress, the tension remains. It may be hard to know when the thread one is snipping won't damage the rest of the fabric and when it will. The optimal trade-off between stability and change may be hard to identify. If we know little about how a society works, we will have scant hope of striking the right balance between stability and change; if we know a great deal, our prospects will be brighter. So even though the notion that society is like a seamless web is false if taken literally, we still face the question of how dense the connections are, of whether society is more like a seamless web or more like a fabric with prominent seams.

My analysis in this book suggests that there is no general, single answer to that question. In some societies, the seamless-web metaphor is somewhat apt, in others totally misleading. Modern liberal democratic societies are definitely not like seamless webs, though traditional societies or at least the earliest human societies may have been somewhat more like that. Remember, I have emphasized that liberal democratic societies have multiple, independent hierarchies and multiple, independent cooperative schemes. There is no single, overarching cooperative scheme that can only function if there is virtually unanimous agreement on, and high compliance with, one set of moral rules; and there is no sovereign moral authority with a veto on change. In that kind of society, things aren't so densely interconnected: you can change some things without changing everything. So if there is no general answer to the question "Is society like a seamless web?" then there is no general answer to the question "How much moral disagreement is compatible with stability?" It depends on the kind of society you're talking about.

Modern liberal democratic societies are much more loose-jointed, modularized, or seamed than the small, face-to-face societies in which distinctively human moralities first came on the scene. In terms of moral rules, modern liberal democratic societies are considerably less unified than most human societies that have existed before them. That's part of what we mean when we say that modern societies—at least the more liberal ones—are "pluralistic." What I'm suggesting is that to the extent that conservatives fixate on the

seamless-web idea, they may be approaching the problem of the trade-off between moral conformity and moral progress in a nostalgic, backward-looking way, failing to appreciate that the institutional structures of modern liberal societies allow considerable moral disagreement—and that this is the key to their potential for achieving moral progress without unacceptable instability.

This is not to say that attempts at moral progress can't backfire—that they can't produce regression or other unintended bad consequences. My point is only that we have no general answer to the question that has divided conservatives and progressives for centuries: how much moral disagreement ought to be tolerated for the sake of moral progress depends on the institutional setting. That should be a familiar idea: I've been arguing all along that the possibilities for moral progress depend on the social environment.

If I'm right that liberal societies can afford to tolerate more disagreement than societies that are more tightly knit, and that moral disagreement is a prerequisite of moral progress, then we shouldn't bemoan the fact that our liberal society has more moral disagreement than earlier ones. Instead we should celebrate it, recognizing that it expands the possibilities for moral progress. Philosophers like Alasdair MacIntyre, who pine for the good old days when most everyone in society was (supposedly) on the same page morally speaking, are unwittingly endorsing moral stagnation (MacIntyre 1981, ch. 2).

In this book, I have tried to solve the Big Puzzle in order to confirm my thesis that the character of our morality and our moral identity as individuals depend on environmental factors that are subject to human control. I haven't offered a prescription for how to sustain the shift toward more inclusive moralities against regression to tribalism. Though I have developed a theory of the current, ideologically driven, deeply divisive intrasocietal tribalism that I and many other people find so dismaying, I haven't fleshed out a program for how to combat it. That would take another book (or two).

Yet I think it is worth emphasizing that the story I've told provides considerable resources for developing strategies to combat both intersocietal and intrasocietal tribalism. And that story opens up the possibility that further scientific investigation will show us

how to curb tribalism effectively. More generally, I hope I've said enough about how moral change comes about to demonstrate the feasibility of doing something we desperately need to do if we're to take charge of our moral fate: develop a science of moral institutional design.

Most obviously, to reduce the threat of tribalistic morality, it's vitally important to maintain the hard-won distance some societies have already achieved from the harsher conditions of the EEA—the conditions that “toggle” the adaptively plastic responses in the direction of tribalism—and to help people lift themselves out of those conditions where they still exist. This means, first and foremost, continuing and amplifying efforts to reduce the risk of diseases, improving physical security, and sustaining and augmenting institutions for mutually beneficial cooperation among groups.

It's also necessary to try to prevent people from coming to *believe* that the harsh conditions of the EEA exist when in fact they don't; and that means finding ways to thwart the efforts of people who try to convince us that we are in EEA-like conditions. It's also vital to sustain and extend the social-epistemic conditions for moral progress, including freedom of expression and association, civil-society organizations that can effectively influence government, a robust culture of reason giving, and communication technologies that provide opportunities for enhancing the human capacity for perspective taking and for extending sympathy to strangers. And it's extremely important to ensure that communication technologies are not under any central or unified control, whether it be public or private.

Doing all of that is a tall order indeed, but it still may not be sufficient to combat ideologically driven intrasocietal tribalism. If I'm right in thinking that this kind of tribalism is an adaptation, in modern societies, for competition among groups, then the only way to eliminate or contain it may be to provide a less destructive form of competition that adequately serves the needs and interests that animate competition among groups.

My hunch is that the ideologically driven intrasocietal tribalism we are seeing today in the United States and a number of other countries is a response to the perceived failure of democracy. In

other words, a metacompetition is occurring, a competition between intrasocietal tribalism and democracy as adaptations for competition among groups within society, and democracy may be losing. (When I use the term “adaptation” here, I’m using the language of cultural selection, not biological [genetic] selection.)

When democracy works the way it’s supposed to, it provides a mechanism for nonviolent competition among groups in society that doesn’t have the destructive consequences of ideologically driven intrasocietal tribalism. It may well be that ideologies of the deeply divisive type flourish when people have lost faith in democracy. If that is so, then an important element in an effective strategy for preventing intrasocietal tribalism from thwarting the full realization of the First Great Expansion will be the revitalization of democracy. The hope is that genuine democracy will outperform intrasocietal tribalism in the competition among adaptations for competition among groups within society.

Democracy both presupposes and reinforces the commitment to settling disagreements through reasonable compromise, and that in turn requires the kind of mutual respect that deeply divisive ideologies destroy. The question is whether deeply divisive tribalism has already so badly undermined mutual respect and the willingness to compromise that an insufficient number of people with the know-how and resources to do so will be willing to take on the hard task of trying to revitalize democracy. I don’t know the answer to that question.

As I said earlier, I suspect that humans living in complex societies can’t do without ideologies. They need a simplifying, evaluative map of the social world that locates them in it and ascribes a group-based identity to them. If I’m right about that, the goal is not to abolish ideologies but to tame them. We need to think long and hard about what sorts of institutions can help shape ideologies in ways that make them less toxic and enable them to be more productive of valuable social change.

The problem, then, may be not ideologies per se but rather deeply divisive ideologies, because they undermine democracy—the best adaptation for intergroup competition within society that humans have devised so far. In addition to being toxic to democracy, deeply

divisive ideologies erode the gains of the First Great Expansion by relegating the Other to an inferior status, destroying the mutual respect that is essential to regarding another person as one's moral equal. More precisely, they undermine democracy by backtracking on the First Great Expansion.

One often hears that to combat tribalism, we have to learn to listen to each other. That's good advice, but incomplete and taken by itself not very helpful. What's needed are institutions that provide incentives for listening and for compromise, institutions that encourage people who disagree with each other not to operate in the zero-sum, winner-take-all mode.

That almost certainly means, among other things, changing existing political institutions so that they provide powerful incentives for coalition building across ideological boundaries. A system of proportional representation and more extensive use of supermajority requirements for votes on important issues might help. Both of these mechanisms provide incentives for coalition building, and coalition building requires compromise and therefore a commitment to appealing to reasons that can sway those you disagree with. In an environment in which coalition building is necessary for political success, people who show mutual respect and a willingness to compromise should have a cultural fitness advantage; their political culture should reproduce itself over time more effectively; they should come out on top when they compete with groups that remain deeply tribalistic.

Another way in which the right sort of institutions could curb the most divisive ideologies would be to provide incentives for people to "unbundle" the various issues that ideologies clump together. That would lower the stakes: our side could lose on one issue without fearing that we will lose across the board. We would no longer have to regard every minor skirmish as Armageddon.

In evolutionary terms, what we need is an adaptation for competition among groups in society that breaks up the monolithic culture wars into a plurality of independent contests. It's hard to see how we might achieve this sort of unbundling, and more generally do away with the zero-sum, winner-take-all mentality that characterizes deeply divisive ideologies, without abandoning a

two-party-only system that forces voters to choose between fundamentally different bundles.

Yet another proposal for moral institutional design warrants discussion—though I’m afraid some people, especially Americans, may be too quick to dismiss it out of hand. Perhaps we should rethink our understanding of freedom of expression and its limits in the light of what we know about evolved human moral psychology, and then consider how the institutionalization of that vital freedom should take this knowledge into account.

I have argued that certain forms of discourse and certain images evoke tribalistic moral responses by mimicking the threat cues of the EEA. This means that not all “hate speech” is equally dangerous: forms that mimic EEA threat cues carry a special risk. In complex modern societies, where intrasocietal tribalism is fueled by deeply divisive ideologies and made all the more potent by information technologies such as the internet, bombarding the evolved moral mind with words and images that evoke the lethal threats our ancestors faced in the EEA can have disastrous consequences.

If we come to know enough about how those sorts of words and images interact with the moral mind to produce behavior that is physically violent or undermines the minimal mutual respect that democracy requires, shouldn’t we take steps to prevent this from happening? If certain exercises of freedom of expression are especially dangerous, as a result of the evolved moral mind’s potential for tribalism, shouldn’t that fact influence how we understand and institutionalize freedom of expression?

There are several different approaches to limiting exercises of freedom of expression that use EEA-like threat cues to stimulate violence or to destroy the mutually respectful disagreement and compromise that democracy requires. Some are more problematic than others. Government censorship might be unacceptable—too subject to abuse and error. Other approaches, such as voluntary agreements among social media platforms to exclude discourse and images that evoke tribalistic responses, might be not only less dangerous but also more effective. The central point is that these strategies would be more targeted and scientifically informed than

attempts to limit freedom of expression in the name of something so vague as the grab-bag category of “hate speech.”

At present, I’m not willing to take a stand on whether any such effort would be permissible, all things considered, much less to advocate one approach as superior to others. I simply want to stimulate a conversation about the implications for freedom of expression of the analysis of the evolutionary roots of tribalism I’ve offered in this book. This is a conversation we desperately need to have.

My aim in this book hasn’t been to provide a prescription for how to design institutions to combat tribalism. I simply want to stress that *for the same reasons that the right sort of institutions were critical for achieving gains in inclusion, an effective response to tribalism must be institutional, too.* In both cases, it’s a matter of using our formidable powers of niche construction wisely; and in a world in which human life is thoroughly shaped by institutions, that means paying close attention to moral institutional design. Simply exhorting people to be more civil or more tolerant of different points of view won’t be enough without structural changes in the heavily institutional niche in which we live. If I’m right that intrasocietal tribalism is flourishing because of the perception that democracy is failing, then the most important focus of institutional design should be to improve existing democratic institutions.

Any effort to design institutions—or for that matter any attempt to change anything—carries the risk of unintended bad consequences. For the foreseeable future, the safest way to design institutions for better moral results might be to proceed defensively: to concentrate on structuring them so as to minimize the risk of the worst sorts of regression or at least to curb the more dangerous regressive tendencies that are already at work.

Yet opportunities may arise for responsible design with a more positive goal. The better we understand the general conditions for progressive moral change, the better our prospects become for getting good results through moral institutional design. Some possible moral improvements may be hard to predict, simply because the human genius for constructing new niches seems to know no bounds. Nevertheless, as our knowledge of the human moral mind and its interaction with different environmental factors increases,

it may eventually become feasible—and morally mandatory—to undertake more ambitious institutional design, to use what we've learned about recent progress in the direction of inclusion to extend that dimension of progress still further.

I've made the case that humanity can for the first time take charge of its moral fate if we learn enough about the moral mind and how it interacts with specific environmental features to be able to engage in scientifically informed moral institutional design. I want to make clear, however, that I'm not advocating that something so profoundly important as what sort of moralities we have and what sort of moral beings we are should be determined by some elite, scientific or otherwise. To be morally and prudentially acceptable, any attempt to maximize human moral potential through scientifically informed moral institutional design would have to take place within the political processes of a liberal democratic order—if those resources are still intact. Or if democracy has already broken down under the onslaught of tribalism, some other kind of bottom-up, not top-down, process would have to be used. Getting the science right is necessary, but very far from sufficient, for a responsible effort take charge of our moral fate. The citizens of a democratic society ought to show proper deference to scientific expertise when it comes to the facts, but scientific expertise is not an entitlement to social control.

My worst fear, however, is not that scientists will dominate us with knowledge about how the moral mind works; it is that contemporary intrasocietal tribalism is undermining the credibility of genuine science. The ability of the scientific community to produce knowledge of great potential value—whether in the case of climate change or moral institutional design—will be of no avail if a substantial portion of the population is so blinded by ideology that they are unable to distinguish between genuine experts and mere pretenders or if they think that genuine expertise has been disabled by ideological bias (as when Rush Limbaugh proclaims that we can't believe climate experts' predictions about global warming, because they are liberals).

As I close this detective story, I hope you will share my feelings: a sense of hopefulness liberated from the pessimistic thought that

we are morally tribalistic by nature, tempered with the sober recognition that the moral progress we have achieved is fragile and may collapse if we don't stop the regression we are now witnessing. On balance, I'm guardedly optimistic. We've gotten beyond tribalism before; perhaps we can learn to escape the new forms it's now taking.