THE EVOLUTION OF MORAL PROGRESS

A BIOCULTURAL THEORY

ALLEN BUCHANAN
AND
RUSSELL POWELL

OXFORD
UNIVERSITY PRESS

world of social media. For the first time, scientific information about biases and de-biasing techniques is being developed, and the first rigorous scientific work on social information technologies and their psychological and political impacts is being conducted. Such developments make the redirection of effort toward the improvement of social moral-epistemic resources all the more cogent.

Even if the moral arc has bent (rather recently) toward justice in some important respects, in particular in the dimension of inclusivity, this trajectory is not inevitable or perhaps even probable. To think otherwise would be to dangerously underestimate the amount of cultural and institutional scaffolding that is necessary to bring about, sustain, and advance moral progress. We explained at the beginning of this book that one peculiar feature of moral progress is that over time it tends to become invisible. Yet this invisibility can foster fragility as the inclusivist foundations that we take for granted can suddenly be undermined without anyone noticing until it is too late. If we wish to shore up moral progress, it is crucial that we begin by bringing it out into the light of day.

CHAPTER 11

Biomedical Moral Enhancement and Moral Progress

The Evolutionary Mismatch Problem, Again

Humans in the twenty-first century are confronted with a daunting array of moral problems, from climate change and poverty to the prospects of nuclear war, terrorism, and genocide. These are all ethical challenges that human moral psychology seems ill-equipped to address, given that it evolved to function under very different social and technological circumstances: namely, in small, scattered hunter–gatherer groups packed full of kin, armed with primitive weaponry, and possessing only a very limited capacity for ecological impact. The high levels of cooperation and technological prowess achieved by human hunter–gatherer groups may have enabled them to wipe out continental megafauna and carry on tribal blood feuds, but it did not give them the capacity to destroy ecosystems on a planetary scale and, with them, the human species itself.

The situation is very different for large post-Neolithic societies like the ones we inhabit today, with sophisticated divisions of labor, powerful technologies, gigantic surpluses, and an energy share rapidly rising to the level of a Type-1 Kardashev civilization—one that controls a major share of all the energy found on planet Earth.[1] Humans now engage in niche construction on a truly

[1] In a well-known paper in the *Journal of Soviet Astronomy*, the astrophysicist Nicolai Kardashev classified civilizations into three types: Type

global order, and they have the ability to bring about the Earth's sixth mass extinction event, whether through nuclear annihilation or the unintended side effects of modern economic development and lifeways. Furthermore, modern nations and global markets sustain levels of inequality that would have been inconceivable in pre-Neolithic societies. In the small hunter–gatherer bands that characterized the vast majority of human evolutionary history, internal conflicts were solved through the evolution of a robust egalitarian ethos (see Chapter 4). But our abilities to sustain cooperative egalitarian social structures appear to break down when it comes to massive, complex societies—the circumstances in which humans have lived ever since the advent of the agriculture revolution some 10,000 years ago. Or at least human beings seem not to have discovered, so far, how to combine post-Neolithic revolution social complexity with robust forms of egalitarianism. In addition, over the last few millennia, and especially in the last few hundred years, intergroup conflict has grown orders of magnitude more destructive due both to the sheer size of the groups involved and to the unprecedented power of the weaponry employed. So there is a profound evolutionary mismatch, so the logic goes, between our prehistoric moral psychology, on the one hand, and modern human moral ecology, on the other.

Aligning Human Moral Psychology with Modern Moral Ecology

One way of realigning human moral psychology with modern human moral ecology would be to radically alter our social and technological environment so as to return to pre-industrial—indeed, pre-agricultural—lifeways. Needless to say, this is neither

I civilizations control most forms of planetary energy; a Type II civilization is one that controls most of the energy output of its sun (Type I civilizations control only about one-billionth of stellar output); and a Type III civilization is one that controls energy on a galactic scale.

plausible nor, on most moral accounts, desirable. It is implausible because large, differentiated, hierarchical populations will inevitably supplant small, egalitarian, hunter–gatherer populations through technological and epidemiological interactions, as documented in Jared Diamond's magnum opus *Guns, Germs and Steel.*[2] It is undesirable because on any reasonable account of well-being, humans in modern developed societies (though perhaps not in all post-Neolithic or even industrial societies) enjoy markedly improved well-being as compared to that of prehistoric hunter–gatherer populations, which suffered from exceptionally high levels of homicide, disease, predation, starvation, and child mortality, and hence lower life expectancies. Quite apart from that, returning to hunter–gatherer societies would mean a drastic, indeed catastrophic, reduction in the human population. Thus, returning to pre-Neolithic modes of subsistence is clearly a non-starter.

And so, seeing no alternative solution to the evolutionary psychological mismatch problem and in light of the seriousness of the threats we now face, some liberal political philosophers—whom we call "evoliberals"—have advocated directly altering the biological underpinnings of human moral psychology to meet the pressing ethical demands of the modern world.[3] The basic idea underlying the biomedical moral enhancement (BME) enterprise is that we can use biomedical technologies, such as neurological, pharmacological, and genetic interventions, to enhance human moral capacities, including moral emotions such as

[2] Jared Diamond, *Guns, Germs and Steel: The Fates of Human Societies* (W.W. Norton & Company, 1996).

[3] See Ingmar Persson and Julian Savulescu (2008), "The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity," *Journal of Applied Philosophy* 25(3): 162–167; Ingmar Persson and Julian Savulescu (2012), "Moral Enhancement, Freedom and the God Machine," *Monist* 95(3): 399–421; Ingmar Persson and Julian Savulescu, *Unfit for the Future: The Need for Moral Enhancement* (Oxford University Press, 2014).

empathy, cooperation, and trust,[4] and morally relevant cognitive abilities, such as the ability to understand the temporally distant effects of present actions—capacities that, evoliberals argue, are severely limited by human evolutionary history. If the evoliberals are right, we are on the brink of a revolution in how moral progress is to be achieved. A theory of moral progress ought to take the possibility of this revolution seriously and attempt to assess its prospects. That is the task to which this chapter is devoted.

A key framing assumption underlying the BME project so conceived is that evolved human moral nature is a source of great, if not insurmountable, resistance to solving the onerous moral tasks that lie before humanity at present. In Chapter 4, we showed that the evoliberal position reposes on the same evolutionary assumption that undergirds the "evoconservative" view: namely, that there are strong evolutionary constraints on human nature, especially in relation to the human capacity for moral inclusion. Recall that evoconservatives conclude from the supposed fact that evolution has produced parochial altruistic dispositions that inclusivist moral norms are futile or unsuitable for beings like us and that we should therefore revise our moral norms to better reflect the limitations of human nature. While they start from the same evolutionary proposition, evoliberals conclude instead that these evolutionary "facts" justify substantial efforts to enhance the biological underpinnings of moral capacities in order to bring prehistoric human moral nature in line with modern moral judgments—particularly given the urgency of the moral problems that we face and the inability of culture to solve them.

This last clause is critical: the evoliberal position rests on the assumption that culture is quite feeble and only minimally

[4] Allen Buchanan, *Beyond Humanity? The Ethics of Biomedical Enhancement* (Oxford University Press, 2011); John Harris, *Enhancing Evolution: The Ethical Case for Making Better People* (Princeton University Press, 2010); Nicolas Agar, *Truly Human Enhancement: A Philosophical Defense of Limits* (MIT Press, 2013).

shapes human moral norms and dispositions. Indeed, much of the philosophical attention to BME has been motivated in part by the belief that cultural forms of moral enhancement (e.g., moral education) have been only modestly effective and are simply not up to the task of mitigating major anthropogenic harms and existential risks. It is the evoliberals' lack of confidence in cultural innovations that leads them to advocate biomedical interventions.

The central question we wish to explore in this chapter is this: how much moral progress is possible, and can major moral regressions be avoided, without the biomedical enhancement of human moral capacities? One way of approaching this question is to look at the extent of moral progress that has already been achieved as this may give us some idea as to the power and limits of more "traditional" forms of cultural moral enhancement. A review of the impressive list of cases and types of moral progress canvassed in Chapter 1 is by itself enough to call this assumption into question. As noted in the Introduction, major moral innovations tend to become invisible once our social moral lives are restructured around them, and thus it is easy to gloss over the truly radical nature of moral progress that has already been achieved. Beyond invisibility, some conservatives might be loath to recognize the radical nature of moral progress because it is in tension with their views regarding moral degeneration or because it is dissonant with their conceptions of traditional society as a ubiquitously positive moral force; evoliberals, on the other hand, might downplay the revolutionary nature of moral progress out of concerns that such self-congratulatory recognition would take the wind, so to speak, out of the reformists' sails. In any case, the point is that the strong evolutionary constraints assumption is belied by the very substantial moral progress that has already occurred. As persuasive as this rebuttal is, it is only by connecting up our history of moral achievements with empirically rigorous investigations of human morality—and, in particular, with the naturalistic theory

of moral progress we have outlined—that we can begin to make meaningful projections about the ultimate scope of moral progress with, and without, BME.

In Part II of this book, we summarized the prevailing evolutionary explanation of morality and explained why it is unable to accommodate cases of sweeping, progressive moral change that we referred to, collectively, as the "inclusivist anomaly." We then sketched an "adaptive plasticity" model of moral psychological development that can accommodate this shift toward inclusivity. This biocultural model, to recap, holds that exclusivist morality is the result of a conditionally expressed moral response that is sensitive to environmental cues that were historically indicative of out-group threat. Such cues, which are detected during the moral development of individuals and feed back into the evolution of cultural moral systems, include (inter alia) signs of infectious disease, indications of resource scarcity, and enculturated beliefs about out-groups.

The present chapter considers the implications of this naturalistic theory of inclusivist moral progress for the plausibility of BME as a solution to some of the most pressing moral problems of our time. We argue that once these problems are recast in terms of moral inclusivity, it becomes clear that BME technology, at least as narrowly conceived by BME proponents, is unlikely to be either necessary or particularly effective in addressing them. On our naturalistic theory, efforts to achieve major inclusivist moral progress and to avert reversions to catastrophic exclusivist moralities do not go against the human evolutionary grain *tout court*. Rather, they only go against the evolutionary grain under certain environmental conditions, and these conditions are both epistemically accessible and within our practical powers to modify. We conclude that cultural moral innovations that make use of our biocultural model of moral progress stand the best chance of solving the evolutionary mismatch problem.

The Evoconservative–Evoliberal Convergence

In Chapter 4 we saw that some thinkers in the secular conservative tradition have appealed to evolutionary theory to lend scientific credibility to long-standing but historically under-evidenced suspicions about the limits of human altruism and the fragility of non-strategic moral relations between peoples. According significant weight to the evolved constraints on inclusivist moral response is not unique to the evoconservative tradition, however. Some liberal moral philosophers have likewise argued that the legacies of human evolutionary history make it difficult to act on the inclusivist moral norms we have come to endorse. Contra evoconservatives, however, these evoliberals contend that rather than giving us reason to trim back our norms, evolved constraints on human morality suggest that a systematic program of BME will be crucial in order to drive major moral progress and to avert future moral catastrophes. Evoliberals conclude not only that BME should be pursued but that in addition it should be given relatively high priority in the allocation of limited social resources.

In a passage worth quoting at length, Ingmar Persson and Julian Savulescu sum up this line of argument, which they have developed in a series of joint publications:

> For most of the time the human species has existed, human beings have lived in comparatively small and close-knit societies, with primitive technology that enabled them to affect only their most immediate environment. Their moral psychology adapted to make them fit to live in these conditions. This moral psychology is "myopic," restricted to concern about people in the neighborhood and the immediate future. But through science and technology, humans have radically changed their living conditions, while their moral psychology has remained fundamentally the same throughout this technological and social evolution, which continues at an accelerating speed. Human beings now live in societies with millions of

> citizens and with an advanced scientific technology which enables them to exercise an influence that extends all over the world and far into the future. This is leading to increasing environmental degradation and to harmful climate change. The advanced scientific technology has also equipped human beings with nuclear and biological weapons of mass destruction which might be used by states in wars over dwindling natural resources or by terrorists. Liberal democracies cannot overcome these problems by developing novel technology. What is needed is an enhancement of the moral dispositions of their citizens, an extension of their moral concern beyond a small circle of personal acquaintances, including [to] those existing further in the future. The expansion of our powers of action as the result of technological progress must be balanced by a moral enhancement on our part. Otherwise, our civilization, we argued, is itself at risk. It is doubtful whether this moral enhancement could be accomplished by means of traditional moral education. There is therefore ample reason to explore the prospects of moral enhancement by biomedical means.[5]

Presupposing the fixed nature of human moral psychology and the feebleness of cultural moral reform, Persson and Savulescu argue that the most effective means of transcending our inability to extend moral concern beyond the group, including to individuals of future generations, is by altering the biological bases of our moral capacities—faculties that evolved in and for a prehistoric world and are desperately in need of an update.

We might distinguish a weaker evoliberal claim, which holds simply that our contemporary moral problems are so dire that any technical means for reducing our parochial tendencies should be on the table, BME included. It is difficult to fault the more modest claim. Persson and Savulescu are right that the urgency of the problems that humanity faces makes it irrational to rule out

[5] Persson and Savulescu, "Moral Enhancement," supra note 3, pp. 399–400.

investigating potential avenues of BME—there is nothing wrong with having another arrow in our quiver. But there is room for significant disagreement over the relative emphasis that evoliberals place on the biological versus the cultural underpinnings of moral thought and behavior, over how important one believes BME is likely to be relative to cultural reform, and over how BME should be prioritized relative to cultural modes of moral enhancement.

Such relative importance claims are crucial for the evoliberal case, for if BME is to provide effective solutions to pressing global moral problems, it would have to be carried out rapidly and on a global scale. Such a large-scale program of biomedical intervention comes with significant risks of unintended consequences and raises ethical concerns surrounding enforcement and coercion (more on this below). Thus, evoliberals must envision the prospective payoff of BME as sufficiently great, and the expected utility of traditional cultural moral enhancement as sufficiently minimal, to outweigh these risks and concerns. For present purposes, therefore, we will engage with the more substantive evoliberal assertion that BME will be critical for solving our greatest moral problems and for ensuring that further moral progress is achieved and sustained.

From Evolutionary Facts to Psychological Inferences

We have seen that the evoliberal, like the evoconservative, infers from supposed facts about the evolution of morality that human moral psychology is ill-equipped to meet the moral challenges of the modern world. However, one might reasonably question whether evolutionary accounts of morality can tell us that human psychology is hopelessly mismatched to the moral problems we now face, given that evolutionary accounts are etiological and thus do not speak to the current functionality of a trait or to its range of phenotypic expression. The question of moral malleability turns on the nature of morality's proximate (synchronic) causes, not on its distal (diachronic) causes. In other words, what

matters for purposes of gauging the plausibility and durability of moral progress is the nature of the moral psychology we currently possess regardless of how or why morality originated.

Put more technically, synchronic properties, which determine how moralities develop from a complex interaction of genetic, epigenetic, and environmental causes, "screen off" diachronic properties in relation to the alterability of human moral psychology. This is not to say that the etiological properties of traits provide no information whatsoever about the prospects of their alterability. What it says is that, in principle, if we had full information about the synchronic causal structure of human moral psychology, then we would know everything we needed to know about its alterability—and information about evolutionary origins would add nothing to our judgments about malleability.

If this is so, then how might the evoliberal reliance on evolutionary history be justified? Even if synchronic facts about moral psychology screen off diachronic facts with respect to moral plasticity, the evoliberal appeal to evolutionary theory is not necessarily superfluous. This is because although synchronic properties exhaust the facts that ultimately bear on the question of human moral plasticity, the synchronic properties are precisely what are at issue in these discussions. Where the synchronic causal structure of human moral psychology is opaque, evolutionary accounts can permit inferences about the nature of that structure and what it implies for alterability.

What precise epistemic role, then, do evolutionary explanations play in the evoliberal—and, for that matter, evoconservative—logic? As we see it, evolutionary explanation is used to bridge an implicit step in Persson and Savulescu's argument, quoted at length above. This step involves moving from a premise about human moral psychology being adapted for small-group living with rudimentary technology to the claim that social and technological circumstances have changed radically while human moral psychology has remained fundamentally the same. If the latter partial premise concerning the fixed nature of human moral

psychology could be established independently of evolutionary history, then the claims about evolutionary history would do no logical work in the argument, given the screening-off relation described above.

Thus, it must be that *facts about adaptation are taken to warrant the inference of unchangeability*, which is then taken to imply the inability of culture to solve the evolutionary mismatch problem, which then warrants the conclusion that BME will be crucial for major moral progress and to avert global moral catastrophes. Evoconservatives reason in a similar way, although they reject the evoliberal idea that the inherent and unalterable limitations of human nature warrant intervening in the biological underpinnings of human moral capacities—an enterprise they take to be misguided, hubristic, and/or insufficiently respectful of our "given" human nature.[6]

It is worth noting that there are important similarities between the evoliberal emphasis on the necessity of BME and the historical arguments of "reform eugenics" in Scandinavian welfare states. Although the architects of the social welfare state did not subscribe to the biological and social degeneration views that preoccupied conservative eugenicists (see Chapter 7), many reform eugenicists worried about the ability to create and sustain a humane society with a robust social safety net in the absence of substantial efforts to encourage the reproduction of positive social traits or at least a reduction in the incidence of negative ones. Evoliberals, of course, do not make the same scientific mistakes that old eugenicists made, and they do not believe that prosocial traits are inherently possessed by some groups of individuals and not by others; to the contrary, they argue that all humans have the same moral psychological limitations because they all share the same parochial moral psychology that evolved in the Pleistocene. But like reform eugenicists, evoliberals believe that if

[6] Michael Sandel, *The Case Against Perfection: Ethics in the Age of Genetic Engineering* (Harvard University Press, 2007).

we are to create a significantly more just and inclusive world and if we are to avoid moral catastrophes and reversions to exclusivist moralities, then these limitations must be overcome. And because these limitations are evolved limitations, they can only be overcome through biological alteration. Moreover, like the old eugenicists, evoliberals argue that because we are faced with a supreme emergency—in this case, nuclear terrorism and climate change rather than social degeneration—certain coercive restrictions on individual freedom may be morally justified in order to ensure the implementation of BMEs on a massive scale.

Let us assume for the sake of argument that evoliberals (and evoconservatives) are right that there are innate adaptive components of human moral psychology that evolved in the EEA and that can result in parochial or truncated moralities. The trouble with the evoliberal line of reasoning is that the inference from innate adaptation to developmental rigidity is not warranted, and without this inference, their argument for the necessity of BME does not go through. The concept of innateness as it applies to cognitive psychological development is famously problematic, in part because of its attendant pre-theoretical associations and conceptual baggage. As Paul Griffiths has shown, although people often associate innateness with developmental rigidity and species natures, these associations are highly problematic.[7]

First, the concept of fixed species natures is indefensible on current post-essentialistic understandings of the evolving biological world, in which blind variation and natural selection, and not essences, are the casual-explanatory foci of biology[8]; to the extent that species natures have been given plausible formulations, these

[7] Paul Griffiths (2002), "What Is Innateness?" *Monist* 85(1): 70–85.
[8] Ibid.

have been consistent with great plasticity in species traits.[9] Thus, if the idea of innate moral adaptation implies rigid species natures, it runs the risk of generating fallacious inferences of inalterability.

Likewise, the fact that a trait is an instance of innate adaptive design does not imply that it is developmentally rigid or that it is insensitive to environmental inputs. In other words, developmental rigidity is not a necessary component of natural selection explanations. It is true that a high degree of environmental invariance—such as the cross-cultural robustness of a trait—is often taken to serve as evidence of innate adaptation. But how invariant an adaptive trait is across developmental environments is a question of contingency that is not answered by the question of whether that trait is, or is not, an adaptation. Indeed, some traits that are not adaptations may nonetheless be highly developmentally insensitive (such as genetic diseases with high penetrance). And likewise, as discussed in Chapter 6, some traits that are culturally acquired (i.e., not innate) are often very difficult to modify, both within an individual's lifetime and over cultural evolutionary time due to scaffolding and constraints that result in substantial cultural inertia.

Equally problematic is the inference from the fact that a trait is shared by all normal members of a reference class of a given species to the conclusion that the trait is "hard-wired" or developmentally rigid. Universally distributed traits could take a very different form, or might not exist at all, if different developmental environments became ubiquitous.[10] The statistically normal environment in which humans currently find themselves (which includes the modern state, powerful technologies, and highly developed global markets) is radically different from the "normal environment" for humans during the vast expanse of their history.

[9] See Grant Ramsey (2013), "Human Nature in a Post-Essentialist World," *Proceedings of the Philosophy of Science Association* 80(5): 983–993.
[10] Tim Lewens (2010), "What Are Natural Inequalities?" *Philosophical Quarterly* 60(239): 264–285.

It is hard to predict how long an environment we now consider "normal" will endure and hence whether the responses we typically have in that environment will persist or be modified in the future.

Likewise, the fact that a similar character state—such as a sense of fairness, parochial altruism, or out-group aggression—is present in both humans and non-human animals (e.g., other primates) does not imply that the trait is ancestral to both lineages, that it is genetically transmitted in humans, that it is culturally unalterable, or that it is even properly described as the same trait.[11] Whether a given trait is universal in existing populations of a species, whether it reflects adaptive design, whether it is genetic in origin, and whether it is developmentally rigid are all contingent severable questions consistent with any configuration of answers.

So, in short, we do not take issue with the evoliberal appeal to evolutionary theory per se, but we do reject their assumption that if the prevailing evolutionary account of morality is correct, this implies that morality is developmentally rigid and has a "deep" biological etiology. In fact, we too appeal to an evolutionary account of exclusivist morality in order to draw inferences about—and to make sense of—the synchronic properties of human moral psychology. But the evolutionary model we propose allows for a wider range of moralities that can develop across cultural developmental environments. Morality may very well be afforded in part an evolutionary explanation, but as we saw in Chapter 6, it is not like a moth's proboscis, a hyena's clitoris, or a peacock's tail—it is instead like a water flea's armor, except infinitely more open-ended and subject to cultural shaping. Flexibility and cultural sensitivity are built, as it were, into the adaptive design of human morality.

The upshot of the naturalistic theory developed in this book is that efforts to advance and sustain moral progress in the form of

[11] Russell Powell and Nicolas Shea (2014), "Homology Across Inheritance Systems," *Biology and Philosophy* 29(6): 781–806.

inclusivity only go against the evolutionary grain under certain conditions—conditions that we can plausibly identify and deliberately modify. The thrust of our push back against the evoliberal argument, therefore, is not that exclusivist morality is a set of predispositions that can be overcome by cultural innovations and moral education but, rather, that the exclusivist predisposition is itself contingent on the presence of certain conditions that are culturally modifiable. If this theory is correct in broad strokes, then it calls into question the need for a systematic program of BME.

Staving Off Moral Catastrophe: A Tale of Two Solutions

Evoliberals propose BME as an antidote to potential moral catastrophes, such as nuclear terrorism, genocide, and climate change—problems which, on their account, stem largely from two factors: the rapid proliferation of powerful new technologies, on the one hand, and evolved constraints on the human capacity for other-regard, on the other. In contrast, the evolutionary model sketched in this book should lead to far greater optimism about the prospects of finding cultural–institutional solutions to these problems. Many of the major moral concerns that rightfully keep evoliberals up at night implicate constraints on moral inclusivity. And both theory and evidence suggest that cultural solutions will be far more effective than BME when it comes to relaxing these constraints.

There is little evidence to think that BME will be capable in the reasonably near future of reducing the incidence and intensity of intergroup conflicts, whereas cultural innovations stand a far better chance of doing so. Wars, ethnic cleansings, and genocides have nearly always been waged between racial, ethnonational, and religious groups.[12] And as we saw in Chapter 7, such conflicts

[12] See Lawrence Keeley, *War Before Civilization* (Oxford University Press, 1996).

are often facilitated by social moral epistemologies that exclude out-group members from the moral community or assign them a relatively low moral status. This is accomplished through calculated dehumanization tactics and the cultural demarcation of moral community boundaries, reinforced by normal cognitive biases (such as essentializing tendencies and cognitive dissonance mechanisms), which "justify" the marginalization, persecution, or annihilation of out-groups.

At the same time, we also know that institutional context is an important modulator of intergroup conflict. Wars are far more likely to occur when at least one of the states involved is an autocracy, military junta, or monarchy, whereas war between developed democracies is virtually nonexistent; and although democracies may wage war as often as any other type of state, the wars they do wage are significantly less severe than those waged by non-democratic states.[13] Although the causal basis of this robustly evidenced "democratic peace" is unclear and hotly contested, it retains a near law-like status in international relations. Intergroup conflicts are also more likely to occur, and to occur in more severe forms, in the absence of institutions at the international level to ensure that the motives for going to war are legitimate and that the methods used to fight wars are just. All of this gives us good reason to think that there are effective institutional solutions to the problems of intergroup conflict that fuel many of the moral catastrophes that rightfully worry evoliberals. Furthermore, these institutional solutions instantiate inclusivist norms. For instance, democracy in its contemporary forms is premised on the principle that all people are entitled to participate in the political processes of their society; it also protects the freedom of expression, which helps prevent the proliferation of severely defective epistemic practices that underpin exclusivist moralities. Likewise, in the case of institutions for international

[13] R. J. Rummel (1995), "Democracies Are Less Warlike than Other Regimes," *European Journal of International Relations* 1(4): 457–479.

security, both just war norms and the humanitarian law of war presuppose universalizable judgments about war acts and apply the same standards to all parties.

Consider another major evoliberal concern: the impending moral disaster of climate change. Though not immediately apparent, the problem of climate change also has inclusivist moral dimensions, in at least two respects. First, there is empirical evidence to suggest that environmental degradation wrought by the activities of wealthy nations, on any plausible climate change scenario, will fall disproportionately on the world's worse-off populations both within and between nations, mainly because poorer people tend to live at the higher temperatures of lower latitudes.[14] Consequently, the greatest harms of climate change are likely to be morally discounted by the comparably well-off countries and individuals that disproportionately produce them, unless something is done to ensure a more inclusivist response on the part of the better off. Second, we tend to discount the interests of future generations in deciding how we will interact with the environment. If we are to honor our moral commitments to future generations, our moral circle must expand to include not only strangers but also persons who are not yet in existence; in other words, our morality must become even more inclusive than it presently is or than it even would be were morally arbitrary discrimination against existing persons and sentient beings completely eliminated. Bringing future persons into the moral community would require yet further expansion of our capacity for moral inclusiveness, which, according to evoliberals, is at or near its evolutionary limits. This leads evoliberals to advocate BME as a critical solution to climate change.

In contrast, the model proposed in this book indicates that institutional solutions to climate change are far more likely to

[14] R. Mendelsohn, A. Dinar, and L. Williams (2006), "The Distributional Impact of Climate Change on Rich and Poor Countries," *Environment and Development Economics* 11: 159–178.

be effective. There are several reasons for this conclusion. First, the growing recognition that we have moral obligations to future persons is an excellent illustration of our commitment to a subject-centered morality. Thanks to the arrow of time, future generations can neither benefit us nor bite us back, and thus persons of sufficiently distant future generations have no strategic capacities vis-à-vis contemporary people. Nor do contemporary people have sufficiently strong kin relations to distant future generations. Our moral commitments to distant future persons, therefore, must be grounded in a non-strategic, non-group-based conception of moral status.

Indeed, the last few centuries have witnessed a dramatic shift toward subject-centered theories of morality, as documented in Chapters 5 and 9. This remarkable expansion of inclusivist moral norms has, not accidentally, coincided with the amelioration of conditions that foster exclusivist moral response. This began with reduced rates of homicide and theft due to the state's exercise of a monopoly on violence; it continued with meaningful expansions of the rule of law that permitted the peaceful resolution of internal disputes; it increased further with the rise of markets that incentivized mutually beneficial cooperation between strangers and nations; and it culminated in the robust system of international human rights that we see today. It is not much of a stretch, therefore, to think that our moral circle could expand yet further, under the right social and epistemic conditions, to include anonymous individuals who will come to exist long after all existing people are gone. Indeed, this norm has spread quite rapidly over the last decade, as evidenced by Pope Francis's recent encyclical calling for a swift international response to climate change.

Second, difficulties in responding effectively to climate change stem not only from the power of self-interest and the limits of moral inclusivity but also from flaws in social moral-epistemic practices—practices that, qua institutions, are candidates for cultural modification. Few people nowadays believe that present people have no moral obligations to future generations. Much of

the political opposition to meaningful action on climate change in the United States, for example, stems not from a failure of other-regard but from false empirical beliefs—namely, beliefs that the evidence for anthropogenic climate change is nonexistent or ambiguous or that the climate change "problem" is really a scientific hoax or amounts to liberal propaganda. These moral-epistemic deficits can be attributed in part to an inability to identify appropriate expertise, which in the case of certain evangelical communities in the United States, translates into, and is motivated by, an unwarranted skepticism of claims emanating from the scientific community, whose work is often perceived to be in tension with religious doctrine (e.g., special creation). Much of this skepticism is enmeshed in a web of morally exclusivist beliefs, with the work of liberal scientific communities often viewed as a threat to in-group identity and flourishing.

These social-epistemic obstacles to progress on climate change cannot be ameliorated through BME interventions. It is simply not credible to suppose that any genetic or pharmacological intervention could change these complex webs of belief and patterns of epistemic deference. Climate change also poses a series of collective action problems at the international level that only multilateral agreements and institutions can solve in a timely fashion. Collective action problems do emanate from self-interest, but they have time and time again been solved by institutional innovations that create incentives for cooperation, which then foster conditions that are conducive to the development of less selfish moral norms and attitudes.

So far as we can tell, BME offers no promising ways of mitigating the in-group/out-group psychological dynamics, let alone the collective action problems, that engender most major moral catastrophes, from war, terrorism, and genocide to climate change and environmental degradation. In fact, BMEs may very well exacerbate these effects. As Persson and Savulescu acknowledge, the prosocial effects of potential BMEs, such as increases in hormones like oxytocin or other factors that enhance empathy,

tend to vanish when kin relations or intergroup psychology are implicated.[15] "Empathy," as the term is usually defined, refers to the combination of perspective-taking and experiencing vicarious emotions for others that are broadly in line with the emotions other individuals are experiencing. Although empathy has been shown to mediate altruism, the problem is that, as Jesse Prinz puts it, "empathy is ineluctably local."[16] Empathy can lead to prosocial behavior when experienced specifically *for stigmatized out-groups*.[17] But enhanced empathy *as a generalized capacity* can exacerbate negative intergroup attitudes when it is not specifically directed toward out-groups, such as in competitive intergroup environments. In such cases, biomedical moral "enhancements" that increase empathy can make moral decision-making worse because they can accentuate exclusivist moral response, strengthening positive attitudes and behaviors toward the in-group, while intensifying negatively valenced attitudes and behaviors toward out-groups.

In addition, because empathy is tightly bound to partiality, it can lead to a wide range of poor moral decision-making—such as unjustly favoring some individuals with whom we contingently empathize over other individuals with whom, contingently, we don't or favoring the lives of concrete individuals over "statistical" lives. Enhancing some of the biological underpinnings of prosociality can therefore backfire in moral decision-making when group identity, locality, and concreteness are at stake—the very features of moral decision-making

[15] See Steven Pinker, *The Better Angels of Our Nature* (Viking, 2011, chapter 9); C. K. W. De Dreu, et al. (2010), "The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans," *Science* 328: 1408–1422.

[16] Jesse Prinz (2011), "Against Empathy," *Southern Journal of Philosophy* 49: 214–233, p. 228.

[17] C. D. Batson and N. Y. Ahmad (2009), "Using Empathy to Improve Intergroup Attitudes and Relations," *Social Issue and Policy Review* 3(1): 141–177.

that Persson and Savulescu propose BME to counteract in the first place.

Thus, it appears that many of the great moral problems we face stem not from a dearth of empathy per se but rather from the fact that the adequate stores of existing empathy are easily manipulated and misdirected in the service of intergroup conflict and local spheres of concern. Prinz concludes that the most effective way of promoting the general moral point of view on which, incidentally, much of the evoliberal normative argument for BME rests, may be to eradicate or reduce empathy in favor of a less parochial and less vicariously emotional "concern" for others.[18]

Yet even a construct such as "concern" will not be a useful target for BME since concern is only generated after an event has already been appraised to constitute a wrong or at least an undesirable state of affairs, and it is the appraisal in particular that we must target if we are to drive moral progress along the dimension of inclusivity. The key issue, once again, is not a general human deficit of concern but rather that concern is not directed in the right ways—toward, for example, the ill-treatment of culturally demarcated out-groups. It is the parochiality of empathy or concern that should be the *Schwerpunkt* of our moral enhancement efforts in the struggle to stave off intergroup moral catastrophes, and BME as it has thus far been proposed fails to engage at this critical locus of the battle.

In theory, if in-group bias has biological roots, this suggests that, again in theory, there may be biomedical interventions that could ameliorate exclusivist response. It is unclear, however, whether such interventions could be carried out without significant unintended costs. If parochialism was a necessary condition for the evolvability of human altruism, as the prevailing evolutionary explanation of morality would suggest, then we might expect altruism and parochialism to be mediated by common proximate

[18] Prinz, "Against Empathy," supra note 16, p. 228.

causes in human psychological development.[19] Indeed, this is precisely what is suggested by studies showing that oxytocin and empathy accentuate exclusivist moral response. Enhancing the biological basis of altruism may thus amount to sharpening both sides of a double-edged sword: by strengthening the biological (hormonal, genetic, etc.) basis of altruism, we may unavoidably exacerbate antisocial attitudes and behaviors toward out-groups, due to the causal developmental dependence of these phenomena. If this is the case, then we must look to avenues for enhancing the moral motivations and behaviors of humans or ways of ensuring that people act as if they are so motivated, which are not causally constrained in this way.

In short, the driving ideas behind the evoliberal argument are (1) that we are likely to discover BMEs that strengthen prosocial attitudes and behaviors toward strangers and out-groups; (2) that these interventions could be carried out without unacceptable or self-defeating costs that result from the developmental interconnectedness of altruism and parochialism; (3) that these interventions could be implemented with sufficient rapidity on a sufficiently massive scale, with entire democratic, autocratic, and theocratic nations, as well as subversive terrorist organizations, incentivized (despite their exclusivist moralities!) to ingest empathy-enhancing pills or to subject their embryos to genetic selection; and (4) that these interventions could be implemented with sufficient rapidity to address imminent catastrophic threats. Each of these points seems dubious.

In contrast, the evolutionary model of moral psychological development outlined in this book not only explains why intergroup conflicts arise and why climate change, by reducing arable land and triggering global refugee crises, can make these conflicts worse. It also suggests a number of concrete avenues for addressing the problem of intergroup conflict itself. In particular,

[19] Samuel Bowles and Herbert Gintis, *A Cooperative Species. Human Reciprocity and Its Evolution* (Princeton University Press, 2011).

it suggests that this can be accomplished through cultural innovations that ameliorate cues that trigger exclusivist moral response, including faulty social moral epistemologies. Thus, in addition to basic moral education, such as teaching individuals to resist their natural proclivity toward essentialistic classifications of human groups, concerted efforts must be made to ameliorate environmental conditions that mimic the dangerous intergroup conditions of early human evolution or create perceptions of the same. For it is only under these "luxurious" circumstances that inclusivist morality can take root, endure, and expand.

A Nontraditional Approach to Traditional Moral Enhancement

The effort to modify conditions that trigger exclusivist moral response involves several interrelated components. The first involves creating an environment of physical and economic security, both internationally and in microenvironments within otherwise secure nations. This can be accomplished by fostering economic productivity and social surpluses by instituting markets, effective property rights, and the rule of law more generally; by encouraging the genuine democratization of political institutions; and by creating institutions that allow for mutually beneficial intergroup cooperation and the peaceful resolution of intergroup conflicts, as now exist at both domestic and (to a lesser but still meaningful extent) international levels. These institutional interventions can reduce and ultimately eliminate many of the ancient trigger conditions that cue the development of exclusivist response. Importantly, none of these cultural innovations require intervening at the level of individual moral capacities, as BME promises to do.

Second, by significantly reducing the incidence of infectious disease and perhaps other diseases and disabilities that mimic infectious disease outcomes, we can lessen the effects of yet another major cue type—namely, signs of parasite stress, which also

signal the presence of intergroup threat and, consequently, modulate intergroup attitudes and behaviors. Although not all disease and disability is infectious, and thus not all disease and disability indicates the existence of parasite threat, the evolution of adaptive plasticity is a heuristic process that is epistemically incapable of precisely discriminating between disease cues based on their etiology and epidemiology. Although there has been much experimental work on the antisocial priming effects of parasite stress, which has been shown to increase xenophobic and ethnocentric response,[20] the response parameters of the parasite stress cue remain unclear. A broad-strokes approach to reducing general rates of disease through sanitation, vaccination, and broader public and private healthcare initiatives is likely to have a significant attenuating effect on the development of exclusivist moralities.

Although such interventions are biomedical in nature, they do not fall under the rubric of BME proper, insofar as the latter refers to the direct modification of specifically moral capacities. We might nevertheless think of attempts to reduce cues of parasite stress in the service of ameliorating exclusivist moral response as "indirect" BME interventions, which are likely to be more efficacious, cost-effective, and logistically feasible than direct BME interventions when it comes to tempering exclusivist moral tendencies.

Third, we must ensure that inclusivist cultural innovations, such as the protections afforded by the recognition and institutionalization of human rights in the domestic and international spheres, are not dismantled by social-epistemic practices that are designed to engender perceptions of out-group threat conditions, including, preeminently, propaganda designed to evoke racial or ethnonational violence. Enhancing social moral-epistemic

[20] This literature is reviewed in Corey Fincher and Randy Thornhill (2012), "Parasite-Stress Promotes In-Group Assortative Sociality: The Cases of Strong Family Ties and Heightened Religiosity," *Behavioral and Brain Sciences* 35: 61–79.

practices so as to increase their reliability in producing correct, morally relevant beliefs cannot be accomplished biomedically, let alone through direct BME interventions that target the moral capacities of individual people. In focusing on enhancing individual cognitive and affective capabilities, BME has tended to overlook the fact that morally relevant knowledge is the product of social practices and that the potential for moral progress often turns on the epistemic virtues of those social practices.

An important difference between our approach and that of the evoliberal is that our approach targets population-level statistical effects on the social development of morality, whereas the evoliberal approach aims for immediate impact on individual moral development. This difference in our respective approaches bespeaks an important philosophical difference in our respective conceptions of morality. On our view, morality is not an epiphenomenon that supervenes on the aggregate of individual moral capacities and judgments constrained by evolutionary history. Rather, it is a dynamic social phenomenon that causally feeds back into the processes of individual moral development that produce it, which in turn serve as causal inputs into the social evolution of moral systems.

This feedback process resembles the biological phenomenon of "downstream niche construction," wherein organismic adaptations shape the ecological environments in which they and their adaptations continue to develop and evolve.[21] In social moral evolution, this feedback between exclusivist morality and environmental conditions can drive both moral progression and moral regression, depending on the direction of the changes to the initial conditions of the system. In focusing on individual moral response generated by a fixed prehistoric human moral nature, BME proponents tend to overlook this dynamic causal structure

[21] F. J. Odling-Smee, K. N. Laland, and M. W. Feldman, *Niche Construction: The Neglected Process in Evolution* (Princeton University Press, 2003).

of human morality—and thus to underestimate the pivotal role of culture in the evolution of morality and moral progress.

Objections

We envision several objections to our critique of the evoliberal position.[22] The first and weakest objection maintains that what we have identified as the inclusivist anomaly is not really an anomaly at all since there is a massive rift between what might be called the "official moral declarations" of societies as expressed in attitudes, international documents, and laws, on the one hand, and the actual practice of human beings in those societies, on the other. In other words, inclusivist morality is essentially aspirational, and the fact that it remains essentially aspirational demonstrates the force of our "hard-wired" psychological constraints on moral inclusivity. Perhaps there are only weak constraints on the shape of theoretical, doxastic, or official morality—but the actual lived morality of human beings, so the objection goes, is strongly shaped by our evolutionary history. We do not find this objection persuasive. Throughout this book, we have made the case that the inclusivist shift is not merely aspirational, giving many examples where inclusivist commitments have become embodied in large-scale institutional changes that are quite costly to the societies that implement them. The British abolition of the slave trade and then of slavery in the British Empire and the regulatory constraints on the use of animals in experimentation are two compelling examples. The facts that some forms of slavery still exist and that factory farming is still common are not objections to our theory. Indeed, such an uneven implementation of inclusivist norms is to be expected, given that inclusivist moral progress is a late arrival in human history and only emerges under limited, luxurious conditions.

[22] We are grateful to Ingmar Persson and Julian Savulescu for raising these points.

A second, stronger objection holds that our critique falls well short of a refutation of the evoliberal position since many of the evolved psychological biases that the evoliberal argument identifies do not implicate moral inclusivity. For instance, evoliberals argue that evolved preferences for short-time horizon preferences—or what they call a "bias toward the near future"—also undercuts a great deal of moral progress, especially in relation to climate change and other global collective action problems. Moreover, the objection continues, the adaptive plasticity account of moral exclusivity developed in this book fails to explain all major aspects of moral inclusivity. In particular, it fails to account for an extra layer of altruism/sympathy stratification within the group, namely at the level of kin. Even if all cues of out-group threat were eliminated, we would still be stuck with the moral parochialism that emanates from nepotism and cronyism—bias for family and friends, for the near over the far, for the concrete over the statistical. To expect anything else from unenhanced human beings would be utopic. The most plausible way of overcoming these dimensions of exclusivity, the evoliberal concludes, is through the biomedical enhancement of the fundamental moral motivations and dispositions of human beings.

It is true that some forms of exclusivity, such as favoritism toward kin, are not a response to out-group threat cues; it is also true that some moral problems stem not from exclusivist dispositions but rather from biases toward the near future, the tendency to favor concrete lives over statistical lives, moral intuition asymmetries between act and omission, and so forth—and thus one might be inclined to conclude that such biases would be unaffected by creating environments in which out-group threat cues are absent. This conclusion would be mistaken. Part II argued that certain "luxurious" conditions must exist in order for inclusivist norms to arise and take root. In particular, it argued that such conditions are necessary for the effective operation of the human capacity for open-ended normativity on a social scale, which plays a crucial role in driving

all types of moral progress. Importantly, open-ended normativity can drive progressive shifts in any of the dimensions of moral progress discussed in Chapter 1, not merely along the dimension of increasing inclusivity.

For example, it is open-ended normativity that allows for the critical evaluation of norms concerning the proper subjects and territory of justice, which enables us to come to view nepotism and cronyism as morally problematic and to attempt to reduce these biases. Indeed, there are now many institutional practices in government and in private organizations that are explicitly designed to curb partiality and nepotism—and some of these inclusivity devices are quite successful. Again, while progress on these fronts has been limited, such efforts have only just begun, and it is far too early to opine on their ultimate efficacy. Likewise, markets give individuals, both as personal decision-makers and as agents of corporations, incentives to take future consequences seriously, as do criminal and tort laws. Also, constitutional design theorists have emphasized that some features of sound constitutions function to mitigate biases toward the near future. So, once again, it is hard to see how evoliberals can conclude that non-biomedical changes are utterly incapable of coping with these evolved features of human psychology.

Finally, the evoliberal might object that the lofty institutional environments that we envision mitigating moral megaproblems are unrealistic, or worse utopian, goals. The aims of the evoliberal project may be unrealistic and utopian as well, the objection goes on, but the magnitude of the threats we face justifies pursuing both of these projects. Our response to this objection is that while the aims of both projects are indeed daunting, there is an important asymmetry that should not be overlooked: namely, we already have actual examples of how cultural innovations, especially the development of institutions, have mitigated some of the damaging effects of some forms of exclusivity. This is true, for example, with the modern state,

the rule of law, democratic governance, markets, the system of human rights, and so on.

In contrast, we have *no* examples of how biomedical interventions of the sort evoliberals propose can help solve serious social problems. In fact, we know that the threat of punishment exerts a much stronger influence on prosocial behavior than does oxytocin or other BME variables, and it does so without reconfiguring the interests of the relevant actors. This is confirmed by laboratory studies involving economic games, as well as in the real world where legal institutions, such as contract enforcement, property rights, tort law, and criminal law exert a more profound and positive influence on promise-keeping, non-exploitation, and non-aggression than BMEs are ever likely to do. And the introduction of these secure interactive environments created conditions under which moral norms and motivations could begin to shift in the direction of inclusivity.

Moreover, implementing BMEs with sufficient speed and on a sufficiently massive scale would require *momentous cultural innovations* in the form of unified international political will; powerful enforcement mechanisms to ensure compliance and to prevent free-riding, which would presumably include state-based coercion and the attendant costs associated with substantial restrictions of personal and religious freedom; social consensus on the permissibility and desirability of mandatory biomedical interventions, and so on. As noted above, evoliberals in effect appeal to the "supreme emergency" exception discussed in Chapter 7 to justify these restrictions on individual freedom. Our point, however, is that deploying BMEs in an efficacious way would require robust cultural innovations—the very power and prospect of which evoliberals want to deny. If biomedical interventions would only work if massive cultural innovations could be achieved, then evoliberals cannot consistently argue that culture is too feeble to cope with the problems to be solved by biomedical interventions.

Conclusion: From an Optimistic Induction to a Pessimistic Conclusion

Persson and Savulescu make much of the fact that there has been little moral progress in over 2500 years since the first great teachers of morality.[23] Not surprisingly, we disagree vehemently with this historical assessment: there has indeed been monumental moral progress over the last few centuries, as demonstrated throughout this book, especially in Chapters 1, 5, 6, 7, and 8. Moreover, we can make sense of the fact that most of this moral progress has taken place roughly over the last two hundred and fifty years, despite the ancient history of moral philosophizing, once our focus shifts away from free-floating moral reasoning and toward the conditions and institutions that are necessary for sound moral reasoning to flourish and become socially efficacious—conditions that have arisen only in the most recent eye blink of human history. This leaves us cautiously optimistic that further moral progress lies ahead and that we have the cultural resources necessary to push it along.

In arguing that a nontraditional approach to traditional moral enhancement is much more promising than BME, we do not mean to understate the daunting nature of the task that lies before us. Establishing lasting economic, political, healthcare, and security infrastructures, as well as international institutions that successfully prevent violent conflict within and between states, solve collective action problems, and help ensure that basic human rights are respected, is clearly a monumental undertaking. Indeed, it might turn out that the needed cultural innovations will not be achieved soon enough to avert catastrophe. Nevertheless, in light of what we are coming to know about morality and its evolution, we believe that cultural innovation is the best hope for preventing moral catastrophes like genocide, nuclear war, terrorism, and climate change.

[23] Persson and Savulescu, *Unfit for the Future*, supra note 3, p. 106.

Research on BME is still in its infancy, however, and we think it is reasonable to view biomedical intervention as one potential instrument in our diverse moral enhancement toolkit. Like many ethicists writing on this topic, we see no in-principle objection to using biomedical technologies in conjunction with cultural modes of moral enhancement to bring moral motivations and behaviors in line with the norms we have come to endorse. Nevertheless, the foregoing analysis leads us to the pessimistic conclusion that BME is unlikely to play a necessary or even major role in the future of moral progress or in solving the greatest moral dilemmas of the coming centuries. We agree with the evoliberal headline that there is an "urgent need to enhance the moral character of humanity," but we do not think that BME is likely to be a very effective and plausible means by which to do so.