

Moral bioenhancement and agential risks

Good and bad outcomes

Phil Torres

2017

Contents

1. What is Moral Bioenhancement?	4
2. Different Types of Morally Bioenhanced Agents	6
2.1 Apocalyptic terrorists	7
2.2 Idiosyncratic actors	8
2.3 Strong negative utilitarians	10
2.4 Future “ecoterrorists”	12
3. Should Moral Bioenhancement Be Voluntary, Compulsory, or Out-lawed?	14

Abstract

In *Unfit for the Future*, Ingmar Persson and Julian Savulescu argue that our collective existential predicament is unprecedentedly dangerous due to climate change and terrorism. Given these global risks to human prosperity and survival, Persson and Savulescu argue that we should explore the radical possibility of moral bioenhancement in addition to cognitive enhancement. In this article, I argue that moral bioenhancements could nontrivially exacerbate the threat posed by certain kinds of malicious agents, while reducing the threat of other kinds. This introduces a previously undiscussed complication to Persson and Savulescu's proposal. In the final section, I present a novel argument for why moral bioenhancement should either be compulsory or not be made available to the public at all.

Keywords

brain health, cognitive ageing, lifestyle, neuroethics, policy, public health

1. What is Moral Bioenhancement?

According to Persson and Savulescu, moral bioenhancement would aim to enhance our *motivation* to engage in morally good behaviors. It would do this by targeting our “core moral dispositions,” namely altruism and the sense of justice or fairness. The former can be decomposed into two subcomponents: empathy, which means putting yourself in someone else’s shoes, and sympathetic concern, which means caring for the wellbeing of other sentient beings. Persson and Savulescu take the latter to be the motivational part of moral bioenhancement.¹ As for the sense of justice, this refers to our willingness to engage in reciprocal tit-for-tat cooperation. These two moral dispositions – consisting of three parts in total – are potentially manipulable through biomedical interventions because they appear to be *biological* features of our phenotypes. This claim is based on studies involving animals and identical twins, as well as crosscultural gender differences. Such interventions could consist of genetic modifications, neural implants, or pharmaceuticals – the latter of which I will here refer to as *mostropics*, on the model of *nootropics*.²

Of primary concern for Persson and Savulescu is the global risk posed by climate change. This problem arises from large numbers of people, mostly concentrated in the developed world, who engage in unsustainable consumerist practices that endanger future generations with environmental ruination. Such individuals are “too little concerned about others who are beyond their immediate circle of acquaintances, especially large numbers of such strangers, too much preoccupied with the present and imminent future, and feeling too little responsible for their omissions and collective contributions.”³ While our moral doctrines have undergone some degree of progress over time (e.g. cat burning is no longer seen as acceptable), they haven’t been sufficiently “internalized to the degree that they regulate conduct.”⁴ Thus, Persson and Savulescu argue that we should seriously consider biological interventions to augment our moral capacities, as well as morally relevant dispositions “like the bias towards the near future and the conception of responsibility as being causally based,” both of which can “limit the moral dispositions of altruism and justice.”⁵

¹ Persson, I., & Savulescu J. (2012). *Unfit for the Future: The Need for Moral Enhancement* (p. 111). Oxford: Oxford University Press.

² The etymology of “moral” can be trace back to the Latin *mos*, which refers to “one’s disposition” or, in plural, to “mores, customs, manners, morals.” Hence the neologism *mostropics*.

³ Persson & Savulescu, op. cit. note 1, p. 104.

⁴ Ibid: 106.

⁵ Ibid: 110.

Another problem that Persson and Savulescu discuss is terrorism, but they appear less enthusiastic about the efficacy of moral bioenhancement in this domain. (In previous papers this is not the case.⁶) Rather, they argue that liberal democracies will need to implement largescale mass surveillance systems to prevent terrorists from using (what we can term) “weapons of total destruction” (WTDs) to dismantle civilization or cause *Homo sapiens* to go extinct. Nonetheless, if a given society were to administer moral bioenhancements to its citizens in toto, as appears necessary to overcome the challenge of climate change, then any malicious agent within that society would be affected.⁷

In the following section, I will argue that moral bioenhancement would mitigate the dangers of certain types of agents while quite possibly increasing the threats posed by others. In the final section, I will use the insights of Section 2 to attempt to answer the question of whether moral bioenhancement ought to be voluntary or compulsory.

⁶ See Persson, I., & Savulescu, J. (2008). The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity. *Journal of Applied Philosophy*, 25(3), 162–177; Persson, I., & Savulescu, J. (2010). Moral transhumanism. *The Journal of Medicine and Philosophy*, 35(6), 656–659.

⁷ Sparrow, R. (2014). Egalitarianism and moral bioenhancement. *The American Journal of Bioethics*, 14(4), 20–28.

2. Different Types of Morally Bioenhanced Agents

Talk of “malicious individuals” and “terrorists” is much too vague for present purposes. Therefore I will adopt a more sophisticated framework that we can call the “agential risk framework”. The idea behind this scheme is this: advanced biotechnology, synthetic biology, molecular nanotechnology, and even artificial intelligence are dual-use in nature, and therefore carry with them a certain risk potential to cause, at the extreme, what Persson and Savulescu call “Ultimate Harm”, which is roughly equivalent to an existential risk (in a different phraseology). But this risk potential can only be realized when the relevant technologies are coupled to a suitable agent who, through error or terror, uses this technology to injure civilization. In other words, an existential catastrophe involving advanced technologies requires a complete agent-tool coupling.¹ It follows immediately that there are two ways of mitigating agent-tool risks: intervene on the tool or intervene on the agent. Moral bioenhancement is an agent-oriented strategy for reducing existential risk because it attempts to modify agents such that even in the presence of WTDs, the overall existential risk would remain low.²

The next question becomes, “What kind of agents would bring about an existential catastrophe if coupled to a WTD?”³ This is a far less obvious question than initially appears. In previous publications, I have identified several agents who would intentionally bring about an existential catastrophe scenario if only the means were available.⁴ These agents are: apocalyptic terrorists, idiosyncratic actors, strong negative utilitarians, future ecoterrorists, machine superintelligence, and extraterrestrials.⁵ The latter two are irrelevant to the current discussion because moral bioenhancement is a human

¹ More formally, we can define an agential risk as arising from any agent who could pose a threat to humanity or human civilization if she or he were to gain access to a WTD, where a WTD refers to any technology that could actualize an existential risk.

² See Torres, P. (2017). *Morality, Foresight, and Human Flourishing: An Introduction to Existential Risks*. Durham, NC: Pitchstone Publishing. Sections of this article draw from this book ad verbum.

³ In *ibid*, I elaborate this issue by introducing the “doomsday button test”, which asks: If a doomsday button were suddenly placed before every individual around the world, how many would willingly push it?

⁴ In other words, I will here ignore the issue of agential error, focusing exclusively on agential terror.

⁵ See Torres, P. (2016). Agential risks: A comprehensive introduction. *Journal of Evolution and Technology*, 26(2), 31–47; Torres, *op. cit.* note 9.

intervention. So, let us examine how moral bioenhancements would affect each of these groups in turn. Imagine for the sake of the argument that moral bioenhancements have reached a state of sophistication such that they are universally available, effective in the way that Persson and Savulescu envision, and safe to consume.

2.1 Apocalyptic terrorists

History is overflowing with examples of violent apocalypticists motivated by eschatological convictions that not only is the end of the world imminent, but also that they have some role in bringing it about. There are two general classes of apocalyptic activists: first, those who turn their focus outwards and believe that “the world must be destroyed to be saved”, and second, those who turn their focus inwards and engage in mass suicides (doomsday cults like Heaven’s Gate). We will focus on the former because the latter are unlikely to pose a grave danger to society. As the terrorism scholar Frances Flannery notes, the apocalyptic worldview is marked by a stark and inflexible Manichaeic dichotomy between good and evil, where those outside of one’s belief community are seen as evil and irredeemable. Flannery refers to this as the condition of “Othering/Concretized Evil”. In her words, “This ‘Othering’ is a conceptual process whereby the ‘ingroup’ (the radical apocalyptic group) ceases to be able to identify in any empathetic fashion with ‘outgroup’ members (everyone else).”⁶ For the apocalyptic terrorist, the scope of empathy is coextensive with her or his community, thereby making the sort of indiscriminate, catastrophic violence unique to religious terrorist groups morally justifiable from their ingroup perspective.

Along these lines, apocalyptic terrorists become infatuated with fulfilling what they see as the wishes of God, which take precedence over the wellbeing of others. If God commands one to kill innocent people, enslave and rape women, and acquire nuclear weapons, then one has a duty to do precisely this. (Note that Osama bin Laden once claimed it was his “religious duty” to acquire weapons of mass destruction, including nukes.)⁷ As a result, the motivating worldview of such individuals limits how much the altruistic subcomponent of sympathetic concern for others extends into the world. The same can be said of the sense of justice, which is foundational to reciprocal cooperation. In fact, the understanding of justice that drives apocalyptic terrorists is that of cosmic justice, whereby God distributes punishments and rewards to humanity based on their cognitive assent to scripture or adherence to behavioral prescriptions. This constitutes the ultimate theodicy that vindicates the existence of evil, and thus catalyzing such a momentous event (i.e. the implementation of cosmic justice by God) through large-scale, merciless violence is warranted by the immense moral significance of the outcome (i.e. eternal life in paradise).

⁶ Flannery, F. (2016). *Understanding Apocalyptic Terrorism: Countering the Radical Mindset*. New York, NY: Routledge, Italics added.

⁷ Yusufzai, R. (1999). *Conversation with Terror*. Time Jan 11.

It follows that moral bioenhancements could indeed potentially mitigate the threat posed by apocalyptic terrorists, since apocalyptic terrorists characteristically suffer from a deficit of altruism and a sense of justice (as defined by Persson and Savulescu). By augmenting these core moral dispositions, terrorists could find themselves less willing to engage in acts that would cause harm to others. But this would have to be done the right way: moral bioenhancements would need to modify not just the intensity, but the scope of empathy, sympathy, and the sense of fairness. Such enhancements would have to demolish the Othering phenomenon that circumscribes moral concern for individuals outside of one's religio-ideological community.

Obviously, there is the crucial logistical issue of getting apocalyptic terrorists to, say, consume a mostropic in the first place. As Robert Sparrow implies, those who are most in need of moral bioenhancement may be the most resistant. This could be overcome to some extent by a society implementing compulsory moral bioenhancement programs, for example, by injecting mostropics into the public water system like fluoride.

While the most notable contemporary apocalyptic terrorist groups are concentrated in the Middle East – the Islamic State being the most notorious instance – there are worrisome apocalyptic groups in the US as well. For example, the Christian Identity movement has influenced multiple far-right organizations motivated by its apocalyptic premillennialism, according to which Jesus will return to Earth (the Parousia) after a catastrophic race war. Hoping to initiate a race war of this sort, “The Covenant, the Sword, and the Arm of the Lord” once planned essentially the same terrorist attack that Timothy McVeigh perpetrated in 1995, namely the Oklahoma City bombing. This plan was (apparently) outlined while the group was “training 1,200 recruits in the Endtime Overcomer Survival Training School.”⁸ Other groups with Christian Identity leanings include the Aryan Nations and the Ku Klux Klan. Thus, if the US were to implement a societywide moral bioenhancement program, one should expect this type of agential risk to become less worrisome.⁹

2.2 Idiosyncratic actors

This category encompasses any malicious individual driven by idiosyncratic motives. Paradigmatic cases include rampage killers and school shooters, some of whom have simply wanted to kill as many people as possible before dying themselves. Consider that the mastermind behind the 1999 Columbine High School massacre, Eric Harris, wrote in his journal, “if you recall your history the Nazis came up with a ‘final solution’ to the Jewish problem. Kill them all. Well, in case you haven’t figured it out yet, I say ‘KILL MANKIND’ no one should survive.”¹⁰ He writes elsewhere that “I think I would

⁸ Ibid: 143.

⁹ See Torres, op. cit. note 9.

¹⁰ Quoted in Langman, P. (2014). Influences on the ideology of eric harris. Retrieved from https://schoolshooters.info/sites/default/files/harris_influences_ideology_1.2.pdf.

want us to go extinct,” adding, “I just wish I could actually DO this instead of just DREAM about it all,” and “I have a goal to destroy as much as possible ... I want to burn the world.”¹¹

There are also incidents in which individuals targeted not people but civilization. For example, Marvin Heemeyer was a Colorado resident who spent years converting a bulldozer into a “futuristic tank” that he used to demolish large parts of his local town. There were no human casualties, although the resulting damage cost the town 7 million dollars. Despite police efforts, his slow-motion spree continued until the bulldozer became stuck in a basement, at which point Heemeyer pulled out a pistol and shot himself. The psychological template provided by Heemeyer and school shooters suggests the possibility of future individuals who, in the presence of a “doomsday button” connected to a WTD, would push it for misanthropic reasons or simply to “go out with the ultimate bang”.¹²

Many idiosyncratic actors suffer from mental and/or personality disorders. Studies suggest that school shooters often exhibit the latter, most notably sociopathy (also known as psychopathy), which is associated with antisocial behavior, narcissism, impaired empathy and sympathy, and diminished remorse. The psychologist Martha Stout describes the condition simply as a lack of conscience, whereby one has “no feelings of guilt or remorse no matter what you do, no limiting sense of concern for the well-being of strangers, friends, or even family members [and] no struggles with shame.”¹³ She estimates that about 4 percent of the population suffers from the condition, which means that some 300 million sociopaths occupy the planet today and perhaps 372 million will live among us by 2050, if the human population reaches 9.3 billion.¹⁴ This constitutes a huge demographic of potentially dangerous individuals whose behavior is unconstrained by the core moral dispositions that Persson and Savulescu identify. Although not all sociopaths are violent, a disproportionate percentage of prison inmates show sociopathic tendencies.¹⁵

The connection between moral bioenhancement and this category of agential risks is fairly straightforward. As Nicholas Agar – who ultimately argues against moral bioenhancement – writes, “we can imagine a biomedical moral therapy that morally improves a psychopath by restoring a normal aversion to inflicting suffering. Prison psychologists provide moral therapy to psychopaths by talking to them. There’s no reason a drug might not have the same moral therapeutic effect.”¹⁶ If school shooters like Eric Harris had been given mostropics to expand their altruistic sensibilities and

¹¹ Quoted in Langman, P. (2009). *Why Kids Kill: Inside the Minds of School Shooters* (p. 31). New York, NY: Palgrave Macmillan.

¹² See Torres, *op. cit.* note 9.

¹³ Stout, M. (2005). *The Sociopath Next Door* (p. 1). New York, NY: Broadway Books.

¹⁴ *Ibid.* 6.

¹⁵ See Kiehl, K., & Hoffman, M. (2011). The criminal psychopath: history, neuroscience, treatment, and economics. *Jurimetrics*, 51, 355–397.

¹⁶ Agar, N. (2015). Moral bioenhancement is dangerous. *Journal of Medical Ethics*, 41, 343–345.

willingness to engage in reciprocal cooperation with others, the 1999 Columbine High School massacre might not have happened. The same goes for Heemeyer, who was motivated by a distorted sense of justice. If moral bioenhancement was compulsory across society, it's questionable that such tragedies would have occurred, since the enabling condition is a failure of proper moral reasoning about the welfare of others.

2.3 Strong negative utilitarians

Negative utilitarianism (NU) is a relatively obscure ethical theory that nonetheless has some adherents within and outside of academia. It comes in several forms, not all of which are risky in the relevant sense. For example, Roger Chao endorses what he calls “negative average preference utilitarianism”, which he argues avoids the Repugnant Conclusion.¹⁷ A strong, classical interpretation of this view, though, sees the ultimate aim of moral conduct as the total elimination of suffering, independent of how much positive utility there is in the world. This leads to an obvious objection, which is largely why the view has been ignored by moral philosophers: the best way to eliminate suffering is to eliminate that which can suffer. As R.N. Smart put it in a 1958 paper, NU would thus appear to endorse a “world-exploder” whose moral mission is the annihilation of all sentient life.¹⁸ A world with zero suffering is a morally superior world to one in which even a single pinprick stimulates the nociceptors of one's finger. Given the growing WTD types and tokens in the world, a negative utilitarian who adheres to the classical formulation could pose a nontrivial existential threat in the coming years, decades, or centuries.

So, perhaps moral bioenhancement could mitigate this agential risk? It seems unlikely. In contrast with the first two cases above, strong negative utilitarians (SNUs) don't suffer from any obvious deficits in their core moral dispositions. Rather, the aim of eliminating all forms of harm is motivated by a deep sense of empathy and sympathetic concern – the twin subcomponents of altruism. In other words, SNUs are no less motivated by “a capacity to imagine what it would be like to be another conscious subject and feel its pleasure or pain” and “concern about the wellbeing of this subject for its own sake” than those who subscribe to other secular ethical systems, such as hedonistic utilitarianism.¹⁹ Indeed, using a WTD to destroy the world might be seen as the ultimate act of altruism by a passionate SNU: what greater sacrifice is there than killing oneself for the sake of morality, and what greater act is there than to destroy every instance of disutility that exists today or could come to exist in the future? It follows that moral bioenhancement as motivational enhancement could not only fail to

¹⁷ Chao, R. (2012). Negative Average Preference Utilitarianism. *Journal of Philosophy of Life*, 2(1), 55–66.

¹⁸ Smart, R. N. (1958). Negative Utilitarianism. *Mind*, 268, 542–543.

¹⁹ Persson & Savulescu, op. cit. note 1, p. 109.

neutralize this agential risk, but potentially make tokens of this type even more risky in a future world cluttered with WTDs.

Even more, if mostropics were to become widely available, one might expect SNUs to voluntarily consume them to enhance their moral capacities. For example, a SNU could find her or himself hesitant to follow through on actions involving WTDs that she or he believes are moral. The biological instinct of self-preservation, or the worry that a WTD attack could fail (thereby resulting in greater suffering), could be sufficient to prevent one from acting. She or he might then acquire mostropics to surmount this reluctance. Alternatively, if moral bioenhancement were to become compulsory in society, such individuals would necessarily have their moral dispositions augmented. In either case, the threat associated with SNUs increases; only in the absence of moral bioenhancements would it remain at its initial level.²⁰

One might object here that what Persson and Savulescu actually propose is a multifaceted regime of moral and cognitive enhancement plus mass surveillance. For the purposes of this article, I will bracket the surveillance component. Thus, the question becomes: could cognitive enhancement in addition to moral bioenhancement reduce the world-exploding risks posed by SNUs? I believe the answer is (very likely) negative. Consider the question: On what grounds could one assert that SNU is false? This position consists of two central components: (a) a consequentialist mandate to evaluate moral actions based on their consequences, and (b) an axiological thesis that specifies the reduction of suffering as the ultimate aim of moral conduct. All forms of utilitarianism accept (a), so we will focus on (b). Now ask: Are there any facts of the matter about whether this thesis is correct or not? Would it be the case that if only SNUs could, say, reason more abstractly or score higher on psychometric tests, then they would recognize (b) as flawed? On both counts I would answer no. In practice, philosophical arguments for claims like (b) often rely upon thought experiments that characteristically end in, as it were, “the dull thud of conflicting intuitions.” When one hears this “dull thud,” there is nothing much left to talk about. It is considerations like these that lead John Leslie to affirm that “much of the danger of this way of thinking [referring to SNU] may come from the impossibility of actually proving its wrongness.”²¹

²⁰ Note that David Pearce, who holds a form of negative utilitarianism, has argued that classical utilitarianism itself could have existentially catastrophic consequences in the form of a “utilitronium shockwave.” See Pearce, D. (2017). Unsorted Postings. Retrieved from <https://www.hedweb.com/social-media/pre2014.html>; Torres, op. cit. note 9.

²¹ Leslie, J. (1996). *The End of the World: The Science and Ethics of Human Extinction* (p. 12). London, UK: Routledge.

2.4 Future “ecoterrorists”

According to Flannery, this type of agential risk will probably grow this century due to the increasingly salient effects of environmental degradation. Already, there are many deep ecology extremists who believe that the destruction of humanity would be a great boon for the biosphere, and that such destruction should therefore be pursued. For example, the Finnish eco-fascist Pentti Linkola argues that Western societies are guilty of a perverse “overemphasis on the value of human life”, to which he adds that “on a global scale, the main problem is not the inflation of human life, but its everincreasing, mindless over valuation.”²² He claims that another world war would be “a happy occasion for the planet” and suggests that, to avoid an ecocatastrophe, “some transnational body [or] small group equipped with sophisticated technology and bearing responsibility for the whole world” should attack “the great inhabited centres of the globe.”²³ Linkola has also avowed that “if there were a button I could press, I would sacrifice myself without hesitating, if it meant millions of people would die.”²⁴

The Gaia Liberation Front (GLF) takes this a step further by endorsing total human extinction using advanced weaponry. As its “Statement of Purpose (A Modest Proposal)” puts it, nuclear war would result in too much collateral damage, sterilization is too slow, suicide is relatively faster but impractical, but bioengineering offers “the specific technology for doing the job right – and it’s something that could be done by just one person with the necessary expertise and access to the necessary equipment.” The statement continues:

Genetically engineered viruses ... have the advantage of attacking only the target species. To complicate the search for a cure or a vaccine, and as insurance against the possibility that some Humans might be immune to a particular virus, several different viruses could be released (with provision being made for the release of a second round after the generals and the politicians had come out of their shelters).²⁵

A similar statement from a 1989 article in *Earth First!* exhorts the following:

Contributions are urgently solicited for scientific research on a species specific virus that will eliminate *Homo shiticus* from the planet. Only an absolutely species specific virus should be set loose. Otherwise it will be just another technological fix. Remember, Equal Rights for All Other Species.²⁶

²² Linkola, P. (2004). *Can Life Prevail? A Revolutionary Approach to the Environmental Crisis* (p. 132). Helsinki, FL: Tammi Publishers.

²³ Milbank, D. (1994). A strange finnish thinker posits war, Famine as Ultimate “Goods.” *Asian Wall St J*, 24; Linkola, op. cit. note 26, p. 131, respectively.

²⁴ Milbank, op. cit. note 27.

²⁵ Gaia Liberation Front. Statement of purpose (A modest proposal). Accessed on 11/10/2016. Retrieved from <http://www.churchofeuthanasia>.

²⁶ Quoted in Dye, L. (1993). The marine mammal protection act: Maintaining the commitment to marine mammal conservation. *Case Western Reserve Law Review* 43(4), 1411–1448.

However ghastly these statements are, it appears unlikely that moral bioenhancements would mitigate this type of agential risk. The reason is that for many in this category, moral components like empathy and sympathetic concern are located at the heart of their motivating ethical systems. For them, the Singerian “circle of moral concern” extends far beyond the human species to include most or all other sentient beings – or even the Gaian system as a whole. Some also see human-caused environmental degradation as a morally catastrophic injustice toward the biosphere. Thus, in tit-for-tat fashion, the ecoterrorist – broadly construed – might reason that since humanity is destroying the environment, so too must humanity be destroyed. Given these considerations, it could be that moral bioenhancements not only fail to reduce the threat of ecoterrorists but actually exacerbate it. One might even expect ecoterrorists to actively pursue the use of moral bioenhancements to enhance their altruistic concern for nature. An effective mostropic could make a hesitant ecoterrorist more likely to follow through on her or his genocidal (in the case of Linkola) or omnicidal (in the case of GLF) goals to save the biosphere from “Homo shiticus.”

Once again, though, one might object that moral bioenhancements in concert with cognitive enhancements could mitigate this type of agential risk. Unfortunately, this does not appear promising, since many ecoterrorists are of above-average intelligence. For example, the Unabomber, Ted Kaczynski (who falls into the neoLuddite more than the deep ecology tradition), was a Harvard-educated mathematician who wrote about the perils of industrial megatechnics eloquently enough to earn praise from many intellectuals in the US (the majority of whom found his actions indefensible). Similarly, Linkola is considered a genius by some in his home country of Finland. And Flannery notes that, within the Earth Liberation Front (ELF) more generally, “in terms of demographics, the members of ELF are most often male, well educated, [and] technologically literate.”²⁷

Complicating matters even more is the fact that empirical science affirms that the globe is warming and the biosphere is wilting due to overpopulation, pollution, CO2 emissions, overexploitation, ecosystem fragmentation, habitat destruction, and so on. Our species really has been a highly destructive force in the natural world, even bringing about the sixth mass extinction event in life’s multibillion year history. Thus, it appears unlikely that a morally and cognitively enhanced ecoterrorist would abandon her commitment to destroying either civilization or humanity itself.²⁸ This “mixed bag” analysis of moral bioenhancements poses the following question:

²⁷ Flannery, *op. cit.* note 13, p. 188.

²⁸ For further discussion, see Torres, *op. cit.* note 9.

3. Should Moral Bioenhancement Be Voluntary, Compulsory, or Outlawed?

Agar asserts that moral bioenhancements could be dangerous because they are likely to result in an “unbalanced excess in influences on moral thinking.”¹ The present article argues that even if bioenhancements were to achieve a proper balance among our cognitive, emotional, and motivational modules, they could still make the world more dangerous. This leads to the question of whether moral bioenhancement ought to be administered by society – and if so, how? There are three options here: (1) make moral bioenhancement compulsory throughout society, (2) make moral bioenhancement optional, perhaps with government incentives such as tax breaks or credits for those who enhance, and (3) do not make moral bioenhancements available.² Persson and Savulescu initially preferred (1), whereas Vojin Rakic argues for (2), claiming that it would solve Sparrow’s objection that moral bioenhancement forced upon a population would implicate the state in a “controversial moral perfectionism.”³

Let us adjudicate this issue by examining the potential consequences of these options, given the agential risk analysis above. We have tentatively concluded so far that moral bioenhancements would mitigate the agential risks of apocalyptic terrorists and idiosyncratic actors while exacerbating those of SNUs and ecoterrorists. It follows that option (1) would probably reduce the threat of apocalyptic terrorists and idiosyncratic actors while inflating the threat of SNUs and ecoterrorists. Whether (1) is advisable thus depends on a careful examination of the relative threat significance of each category of risk. For example, it could be that moral bioenhancements have a greater impact on certain agential risks than others. If the undesirable effect on SNUs or ecoterrorists is disproportionately large compared to the desirable effect on apocalyptic terrorists and idiosyncratic actors, then (1) might not be advisable. Alternatively, if an agent type like ecoterrorists are already extremely motivated to destroy humanity, then forcing an entire population to consume moral bioenhancements might not make much of a difference to the overall threat level. One should also consider the

¹ Agar, *op. cit.* note 21, pp. 343–345.

² See Rakic, V. (2014). Voluntary moral bioenhancement is a solution to Sparrow’s concerns. *American Journal of Bioethics*, 14(4), 37–38. 38

³ *Ibid.* Also Rakic, V. (2014). Voluntary moral enhancement and the survival-at-any-cost bias. *Journal of Medical Ethics*, 40, 246–250.

total number of token agents within each agential risk category: a regime of compulsory bioenhancement might have a huge desirable impact on apocalyptic terrorists and only a small undesirable impact on ecoterrorists, but if there are far more ecoterrorists than apocalyptic terrorists, the net result could be bad.

Now consider option (2). As argued in the previous section, there is at least some reason for suspecting that if moral bioenhancements were freely available, SNUs and ecoterrorists would voluntarily use them while apocalyptic terrorists and idiosyncratic actors would not. This could result in a problematic situation: apocalyptic terrorists and idiosyncratic actors would remain at their initial levels of threat, unchanged by noncompulsory enhancement technologies, while SNUs and ecoterrorists would probably seek out mostropics or other enhancers to intensify their moral motivation to eliminate human suffering and exterminate *Homo sapiens*, respectively. This would be the worst possible outcome: the overall existential danger associated with agential risks would increase.

Option (3) avoids the analytical complexities of option (1) and the repugnant outcome of option (2) by making moral bioenhancements universally unavailable to the public. Narrowly considered, it appears to be the best of the three options. Yet it introduces problems of its own since it would do nothing to mitigate anthropogenic climate change (arising in part because of a tragedy of the commons), which is the primary phenomenon that Persson and Savulescu believe moral bioenhancements could address. Indeed, climate change itself could fuel some types of agential risks; it is a “threat multiplier” that the terrorism scholar Mark Juergensmeyer has explicitly argued will foment apocalyptic terrorism in the future – a point that I elaborate elsewhere.⁴ In fact, a 2015 study proposes a more or less direct causal link between climate change and the emergence of the Islamic State (IS) during the Syrian civil war, and IS is an apocalyptic organization par excellence.⁵ Even more, recall from above that Flannery believes environmental degradation will further foment catastrophic ecoterrorism in the coming decades. As Gary Ackerman warns, modifying a maxim of Juergensmeyer’s, radical conditions will breed radical beliefs – and radical beliefs in a world replete with WTDs could all but guarantee Ultimate Harm.⁶

In conclusion, this article focuses on the conception of moral bioenhancement adumbrated by Persson and Savulescu, according to which moral bioenhancement aims to enhance human moral motivation by targeting our core moral dispositions. Using the agential risk framework, I have outlined several reasons for qualified skepticism about the efficacy of moral bioenhancements as a means of reducing overall existential risk. In particular, it could exacerbate two types of agential risks. I should emphasize

⁴ See Juergensmeyer, M. (2017). Radical religious responses to global catastrophe. In R. Falk, M. Mohaty, & V. Faessel, (Eds). *Exploring Emerging Global Thresholds: Toward 2030*. Hyderabad, India: Orient BlackSawn; Torres, op. cit. note 9.

⁵ See McCants, W. (2015). *The ISIS Apocalypse: The History, Strategy, and Doomsday Vision of the Islamic State*. New York, NY: St. Martin’s Press.

⁶ Personal communication.

here that my conclusions are tentative, given the almost negligible work that has been done on this topic. Thus, perhaps the most robust conclusion of this article is that additional research is desperately needed, focusing not just on moral bioenhancements in particular, but on agential risks in general.

ORCID

Phil Torres <http://orcid.org/0000-0003-4420-9159>

Author Biography

PHIL TORRES is the founding director of the XRisks Institute. His most recent book is *Morality, Foresight, and Human Flourishing: An Introduction to Existential Risks*.

The Ted K Archive

Phil Torres
Moral bioenhancement and agential risks
Good and bad outcomes
2017

Bioethics. 2017;1–6. <wileyonlinelibrary.com/journal/bioe> VC 2017
Correspondence: Phil Torres, XRisks Institute, 189 Viburnum Way, Carrboro, NC.
27510. Email: philosophytorres@gmail.com
John Wiley & Sons Ltd.

www.thetedkarchive.com