Technological moral enhancement or traditional moral progress? Why not both?[1]

Joao Fabiano[2]

Abstract: A new argument has been made against moral enhancement by authors who are otherwise in favour of human enhancement. Additionally, they share the same evolutionary toolkit for analysing human traits as well as the belief that our current morality is unfit to deal with modern problems, such as climate change and nuclear proliferation. The argument is put forward by Buchanan & Powell and states that other paths to moral progress are enough to deal with these problems. Given the likely costs and risks involved with developing moral enhancement, this argument implies moral enhancement is an unpromising enterprise. After mentioning proposed solutions to such modern problems, I will argue that moral enhancement would help to implement any of them. I will then detail Buchanan & Powell new argument disfavouring moral enhancement and argue that it makes too bold assumptions about the efficacy of traditional moral progress. For instance, it overlooks how that progress was to achieve even in relatively successful cases such as the abolition of slavery. Traditional moral progress is likely to require assistance from non-traditional means in order to face new challenges.

1. A NEW ARGUMENT AGAINST MORAL ENHANCEMENT

For over a decade, there has been intense debate regarding the idea of improving ourselves using technology, i.e. human enhancement. One of the strongest arguments in favour of human enhancement is that we have modified our environment and developed powerful technologies to such a strong extent that without also upgrading our morality we risk going extinct. The proponents of moral enhancement – most famously Persson & Savulescu[1] – argue that in order to deal with modern problems such as global warming and nuclear proliferation, humans need to technologically enhance their moral traits. Our present morality equips us to co-operate well in certain contexts via generosity, altruism, a sense of fairness and the desire to punish cheaters. But we fail to co-operate well in extremely large groups that spread across countries and territories or from different ethnicities and backgrounds. Persson

[1] This article has been accepted for publication in the Journal of Medical Ethics in 2020 following peer review, and the Version of Record can be accessed online at https://doi.org/10.1136/medethics-2019-105915
[2] Department of Philosophy, University of Sao Paulo (jlafabiano@gmail.com). Research supported by grant #2019/22383-5 from the Sao Paulo Research Foundation (FAPESP).

& Savulescu argue that we are being conferred an ever-increasing destructive power and technology is rapidly becoming globalised so that a single individual can cause extreme harm if they set their mind to it. Therefore, the probability of any individual having enough power to destroy the whole of humanity has increased. Hence, they conclude that we have *a moral imperative to pursue moral enhancement*, for not doing so will expose humanity to extreme risks of catastrophes or extinction – what Persson & Savulescu call ultimate harm.

Despite the plethora of counter-arguments against moral enhancement, they mostly involve disbelief in one or more of its fundamental premises. In Buchanan & Powell's most recent book *The evolution of moral progress*,[2][i] the two concede that our present morality might be unfit for dealing with modern challenges and share much of the evolutionary background with proponents of human enhancement. Both authors have argued in favour of other forms of human enhancement, such as cognitive enhancement.[3,4] However, they now argue against the moral enhancement project's claim that technological interventions are necessary to overcome our moral unfitness effectively. Although they do not rule out moral enhancement altogether, they do claim traditional that moral progress (which may include other forms of enhancement) is a better strategy to address our modern challenges. In their own words: "we do not think that BME [i.e., moral enhancement] is likely to be a very effective and plausible means". [2] (p. 373) Given the likely costs and risks involved in developing moral enhancement, their conclusion that moral enhancement is not necessary, effective nor plausible would be enough to produce a powerful rebuttal of the project; in particular, a counter-argument to Persson & Savulescu's central claim that we have a moral imperative to pursue moral enhancement.[ii]

Buchanan & Powell believe our moral psychology is sufficiently plastic to allow us to change our presently unfit morality via traditional forms of moral education. They present evolutionary arguments in favour of their claim to challenge the assumption that evolutionary products are inert to cultural modification. They cite past examples of successful cases of moral progress made despite the obstacle of overcoming our evolved moral failings; the most famous of these cases being the abolition of slavery. While many critics of moral enhancement claim that it is unfeasible, will backfire, or that it is based on mistaken moral grounds, Buchanan & Powell argue that we simply do not need it.

In this paper, I will first quickly review what the present challenges that risk our extinction are exactly, and consider how solving them implies a significant change in our morality; namely, increasing large-scale co-operation. Here, Buchanan & Powell agree with Persson & Savulescu, but detailing the problems at hand will help decide if traditional moral

progress is enough to solve them. Next, I will make the case that moral enhancement is a promising solution for decreasing these risks. I will then move on to state Buchanan & Powell's arguments against the need for moral enhancement and argue that, insofar as they disfavour moral enhancement, they rely on somewhat bold assumptions about traditional moral progress. I will conclude that my arguments here suggest a middle way view that claims we should pursue both moral enhancement as well as traditional moral progress without the favouritism expressed by Buchanan & Powell. Not only can we not be certain that traditional moral progress will be enough, but also the severity of the extinction risks we face does not allow us to dismiss proposed solutions easily. However, one should not take failure in establishing that moral enhancement is not needed as an insurmountable argument in favour of moral enhancement. Undeniably, the many other objections to moral enhancement must be addressed before one can confidently argue in its favour. I do not aim to produce such general conclusion here, but merely to investigate one specific promising objection to moral enhancement.

Buchanan & Powell use the term biomedical moral enhancement to refer to most of what I mean by moral enhancement. Occasionally, they talk about moral education being a form of moral enhancement. I will reserve the term moral enhancement to significant changes, brought about via technological interventions, directly targeted at human traits (e.g. co-operativeness, empathy, altruism, etc.) primarily expected to lead to morally better behaviour or motives. I use traditional moral progress in the same way they do to express moral progress that excludes (biomedical) moral enhancement.

2. ULTIMATE HARMS AND PROPOSED SOLUTIONS

In the book *Global Catastrophic Risks*, a variety of experts assess the main risks which could cause the loss of human lives on a global scale.[5][iii] I will briefly mention some of these risks following the book and clarify how the main solutions to each of them involve increased global co-operation.

Sudden anthropogenic climate change is one of the most well-known global risks. The main mitigation strategy is an international effort to decrease the emission of greenhouse gases, which relies on global co-ordination. This is the global risk that has received the greatest amount of attention, yet we have failed to produce an enforceable global strategy to reduce greenhouse emissions, indicating significant changes to our morality are needed.

Super-vulcanism in Toba, Indonesia already drastically reduced human population 75,000 years ago. Dust ejection from volcanic super-eruption produces drastic climate change and a sharp loss of crop productivity resulting in mass starvation. A similar meteorological

effect results from full scale nuclear war. Surviving this effect would require increasing the global grain stockpile. Co-ordinating such a costly increase in order to prevent an improbable but extreme catastrophe is a classic co-operation problem, which requires increasing our ability for large-scale co-operation. While super-volcanic eruptions cannot be avoided, preventing nuclear war is possible but requires widespread substantial nuclear disarmament and radical anti-proliferation treaties. The systems responsible for nuclear threat detection and response have remained the same since the Cold War, and various false alarms have been reported since the fall of the Soviet Union. Widespread substantial nuclear disarmament and radical anti-proliferation treaties would be the only way to reduce these risks significantly, but once again this depends on achieving a level of international co-operation that we manifestly lack.

Infectious diseases have killed more people in the last century than both world wars put together. Smallpox alone has killed well over 500 million people.[6] Now globalisation has made it possible for a disease to spread across the world within days. Our flawed responses to relatively localised epidemics of a deadly disease such as Ebola have evidenced the need for proper global co-operation on the matter.[7]

Lastly, long-standing trends in the advancement of artificial intelligence seem to indicate that we will be able to achieve human-level artificial intelligence in the future. Once achieved, it is likely that this artificial intelligence will be able to improve its own intelligence and reach exponentially high levels of above-human intelligence. Given that higher intelligence is what enables us to control the environment and non-human animals, the introduction of agents with significantly higher intelligence, who may not share our values, would be likely to bring about an extinction risk. Countries are already rushing in an arms race to be the first to develop such a powerful technology, and safety concerns could be overlooked in the absence of an international agreement to prioritise safety in the face of rapid development.

3. MORAL ENHANCEMENT AS A WIDESPREAD SOLUTION

I explicitly discussed how addressing all the global risks explored in the last section relies on large-scale co-operation between groups. Co-operation between groups on a global scale is one of the main proposed targets of moral enhancement. Large-scale co-operation between groups was never a recurrent evolutionary pressure due to groups being small and scattered. Some level of group competition was likely to have been selected for in our evolutionary past. In fact, studies have shown that groups behave in a more individualist and competitive fashion than individuals.[8] Furthermore, most of these risks require both a global

response to decreasing their odds in the form of globally enforceable regulations and strategies as well as a global response for if they were to materialise. One rogue nation or group of individuals would be enough to jeopardise the co-operative effort. Some present risks ought to be addressed as soon as possible and require a radical shift in our co-operative dispositions. Moreover, the reduction of extreme risks taken as a whole can be seen as a public good of a global scale, whose benefits will never be properly observed. No one can directly observe the event of a risk that fails to materialise due to sufficient prevention measures. Therefore, co-operation problems exist not only for solving each particular risk but also for focusing attention on extreme risk itself as a worthwhile cause. Agents might feel safe to take no action given that the number of agents that could act is so high that one single agent defecting would not be noticeable. Moreover, agents have no incentive to act given that if humanity is successful in preventing those risks, their actions might go unrecognised.

There have been other proposed widespread solutions to extreme risks, but several of them also rely on a level of global co-operation that we currently lack. For instance, one prominent solution is differential technological development, as proposed by Bostrom.[9] According to this strategy, humanity's focus should be not on developing technological capabilities faster, but instead on developing them in the right order. We should slow the development of potentially harmful technologies while accelerating the development of safe technologies. However, co-ordinating technological development on a global scale seems impossible, given our current levels of parochialism. Taking the lead on developing a powerful harmful technology would confer any group an immense strategic advantage over others. One past instance of this is the development of nuclear weapons. In fact, the differential technological development strategy would prescribe the development of technologies focusing on solving global co-operation problems before developing any other powerful technology that could be unilaterally used to annihilate a large share of the human population, i.e. the development of deep moral enhancement should take precedence. Co-ordinating the order of technological progress in synchronicity is unfeasible in the absence of efficient means of guaranteeing large-scale co-operation. The presence of a single rogue advanced agent would be enough to jeopardise the strategy as it requires full compliance.

The case for moral enhancement's help in addressing the potential causes of ultimate harm is clear. There are, however, important objections to this project. Unfeasibility, conceptual problems and threats to freedom are just a few of the concerns mentioned in the literature. Moral enhancement has faced extensive scrutiny that should be addressed producing an overall assessment of it.[v] However, this task lies outside the scope of this paper. Here, I

wish to address one specific argument disfavouring moral enhancement. The arguments reviewed here in favour of moral enhancement have been presented only so this specific argument can be evaluated.

4. TRADITIONAL MORAL PROGRESS VERSUS MORAL ENHANCEMENT

4.1 Arguments for traditional moral progress

Buchanan & Powell agree we are unfit to deal with the present challenges of unprecedented technological advancement. Notwithstanding, in its strongest form, their argument conclude we do not need moral enhancement to solve this problem because traditional mora progress is a more effective and safer solution. Their main argument is that our evolved morality is not inflexible enough to frustrate traditional attempts towards moral progress.

Buchanan & Powell claim that moral enhancement advocates (whom they call evoliberals) and their opposition (evoconservatives) share the evolutionary assumption that our evolved moral traits are sufficiently innate and inflexible to make substantial moral progress via traditional means impossible, although the two groups draw opposing conclusions. Once we reject the assumption that human nature is sufficiently inflexible to justify a pessimistic view of the prospects of moral progress, the positions of both moral enhancement advocates and evoconservatives are refuted. The two authors proceed to then argue against this assumption both in theory and in practice.

Theoretically, the fact that our moral dispositions are evolved adaptations does not justify the claim that they are insufficiently flexible. An explanation can concern either the past history or the current state of its explanandum; explanations can be either diachronic or synchronic. Suppose we wish to explain the way the human heart is divided into its four chambers. If we say the human heart has four chambers because it evolved out of a common ancestor with a three chambered heart as an adaptation for better cardiovascular efficiency because an additional chamber better separates oxygenated from deoxygenated blood, that is a diachronic explanation. If we instead say the human heart has two small chambers and two large ones, that oxygenated blood flows from the lungs to the upper left chamber to the bottom left and to the body and so on, that is a synchronic explanation. A synchronic explanation provides factors immediately responsible for its explanandum, a diachronic explanation provides the distant historical factors; the former offers proximate causes, whereas the latter distal causes. Accordingly, Buchanan & Powell argue that evolutionary theories provide only

a diachronic explanation that relies on a distant history in order to draw conclusions about the evolved function of certain traits. These traits require a synchronic explanation to be fully understood. That is, evolutionary accounts explain why a certain, otherwise unlikely, biological configuration becomes recurrent by showing how it is an evolutionary adaptation to recurrent past evolutionary pressures. It does not produce an explanation of all the current aspects of the trait. The fact that a trait has a function in order for it to be an adaptation to certain environmental conditions does not entail that the trait cannot be currently modified by culture. In their own words,

> "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."[2] (pp. 351-352)

Empirically, history is full of examples of substantial moral progress overcoming our moral failings. The two authors cite the abolitionist movements and a decrease in inter-group wars as major examples of what they call the inclusivist anomaly, whereby we expand our sphere of concerns beyond those of our own group. According to them, the standard evolutionary account of human morality says our altruistic propensities are necessarily parochial, as humans were subjected to the pressures of inter-group competition, which selected only groups who were able to co-operate internally while competing with out-groups whenever necessary. An unrestricted tendency towards co-operation would have been disadvantageous and selected against.[10] The fact that human societies have been able to develop higher levels of moral sensibility for out-groups, undergoing structural changes for their sake, shows that culture can produce moral progress despite going against our evolved inclinations. The authors also offer an evolutionary explanation for this inclusivist anomaly, claiming that under conditions of low environmental stress, such as the absence of inter-group conflict, violence, scarcity, or parasites, it is actually advantageous to be able to co-operate with out-groups. They argue that our evolved morality is plastic and can be moulded by culture much more than both evoliberals and evoconservatives suggest. Therefore, it seems their main

point of disagreement is over the flexibility of our evolved morality and not necessarily over whether or not it was shaped by natural selection.

Their argument claims that a belief in an insurmountable obstacle for traditional means of moral progress is flawed, because it is based on an inaccurate view of how evolution shaped human morality. Their conclusion is that

> "We agree with the evoliberals headline that there is an "Urgent need to enhance the moral character of humanity", but we do not think that BME [i.e., moral enhancement] is likely to be a very effective and plausible means by which to do so".[2] (p. 373)

Later on, they clarify that this conclusion does not solely dominate their assessment and that they do not advocate for dismissing moral enhancement entirely. However, they clearly judge moral enhancement to be a much inferior solution to traditional moral progress. In order to evaluate their objection to moral enhancement, their arguments for moral enhancement not being a "very effective and plausible means" should be investigated by themselves . This conclusion requires a significant degree of certainty over which means of moral progress to pursue. I will argue we lack such certainty. The arguments for claiming that traditional moral progress alone can yield the unparalleled and much-needed level of moral progress in order to meet our radically new environment are not that certain. They present convincing arguments in favour of moral progress as an effective way of improving some moral failings, but their arguments do not justify the claim that traditional moral progress should be preferred to meet the unprecedented challenges posited by modern technology and society. I will list several reasons to be uncertain of their reasoning; I do not intend to offer a refutation of their arguments, but to cast doubt over how sure we can be of their conclusion that we should favour traditional moral progress alone or place a substantially greater emphasis on it. Their conclusion requires two claims, that technological moral enhancement is not required to produce the necessary moral progress and that traditional forms will be sufficient. I will now analyse them separately.

4.2 Is moral enhancement needed if morality is flexible?

Regarding their first claim, some degree of flexibility does not mean moral enhancement is unnecessary. One can make the argument for moral enhancement without assuming that human morality is extremely inflexible. An advocate of the moral enhancement project does not need to claim that moral traits are inflexible, only that some of them might be sufficiently hard to change via cultural means so as to justify the pursuit of technological modification.

Buchanan & Powell first major example of the inclusivist anomaly is the abolition of slavery, which was a long process met with extreme resistance, even in the form of wars. Moral enhancement via technological intervention might present a less costly and bloody means to achieve such moral progress. If abolitionists could have developed a technology that would help slavery advocates to perceive slaves as human beings equal to themselves, this could have prevented the American Civil War.[iv] Perhaps we can avoid the bloodshed that could be implied by enforcing an expansion of basic human rights in populations that currently oppose granting those rights to people of all genders, sexual orientations or races. Moreover, it seems that treating members of all groups with some basic concern for their freedom and well-being might indeed have been advantageous in those periods in the prehistoric environment that were relatively peaceful and disease-free. Historical examples might only reveal cases of moral progress that consisted of merely shifting which moral traits we exhibited rather than overcoming them. Under certain stable and peaceful environments, being able to co-ordinate trade and mating between groups would produce higher fitness and there is, in fact, plenty of evidence that early humans engaged in such exchanges.[11] But, under this explanation, respect towards out-groups evolved only to the extent that it enabled such basic trading. The extremely high levels of large-scale co-operation necessary for overcoming the extreme risks listed in Section 2 seem to go much further beyond a basic respect for out-groups and what would have been selected for in our ancestral environment. Therefore, not only would moral enhancement remain a desirable option in place of traditional costly moral progress even when this moral progress is possible via traditional means, but also the extreme and unparalleled levels of moral progress required might be unattainable via traditional means, simply because a reason for them was wholly absent during our evolutionary history.

Buchanan & Powell would have to offer a decisive argument to show that such levels of moral progress are not only feasible via traditional means, but also attainable with lower costs than via technological intervention. The fact that human morality is sensitive to cues present in our evolutionary past is no such argument. As far as one of their prime examples indicates, the abolition of slavery, it is an argument that certain types of moral progress seem to be achievable via traditional means only after much bloodshed. In the next section, another of their exemplary cases of the inclusivist anomaly will be called into question as well.

4.3 Traditional moral progress may not be enough

Their second claim that traditional moral progress is sufficient is also not securely established. Firstly, it is worth noting that a defender of moral enhancement need not claim that

traditional moral progress is unlikely, only that it can be helped by non-traditional technological interventions. Moreover, as I will argue, there is significant room for disagreement regarding how easily we can overcome obstacles to moral improvement without directly interfering with our evolved biology. For instance, even one of the most notable past instances of successful traditional moral progress mentioned by Buchanan & Powell, the abolition of slavery, has not been carried out without its shortcomings.

It is plausible that much of our moral progress has come about via accidental and non-moral advances. For instance, the lower frequency of wars does not seem to be solely caused by moral progress. Buchanan & Powell might concede as much and claim some of this reduction was not caused by moral progress proper but merely improvements from a moral point of view; i.e. improvements that lead to a morally better situation but that are not improvements in our morality. However, if it is not only the case that there are significant alternative causes for the reduction in armed conflict aside from moral progress, but that moral progress made only a minor contribution, then one of their key examples of moral progress fails. Indeed, more rigorous work on nations' proclivity towards war argues that this has increased in modern times, and that the fact that wars have decreased is a consequence of nations now being more spread out.[12] That is, nations are avoiding war against each other not because they have become more peaceful, but because they have become, on average, more distantly placed.[vi] One may argue that what really matters is the actual frequency of wars. However, intuitively, peace seems to mean a state where people are intentionally not committing violence and not just doing so extrinsically. A prison might have lower violence than a certain neighbourhood, but it might still not be considered a more peaceful place exactly because the individual proclivity towards violence is higher despite the fact that violence itself is not. Proclivity matters for morality. One of the most convincing examples the two authors cite might plausibly be an unintentional consequence of the technological capability to spread across the globe; an improvement from a moral point of view but not moral progress proper.

Although there is an overall trend of decreased violence throughout human history, conflicts have been more sporadic from modernity until now, but also deadlier and more focused on non-combatants who are likely to be under-represented in statistics and whose killing is a greater atrocity. World War II has been the deadliest conflict in human history, with the highest death toll per year, both as a percentage of the global population and in absolute terms.[13] Approximately half of those killed were non-combatants, which decreases the reliability of the data. Recent violence has taken a fat-tailed distribution, one in which there is a low probability of extreme events that are hard to predict. Statistical analysis that takes this

into account concludes we cannot assume our recent peaceful period after the world wars resulted from a real trend of decreased violence instead of a trend of more sporadic, but deadlier, violence.[14] Moreover, even the relatively peaceful period after World War II seems to have come accompanied by the new risk of nuclear annihilation and may have partially resulted from it due to nuclear deterrence.

The two authors suggest this recent decrease in wars was brought about by the rise of liberal democracies, which are less likely to engage in conflicts than more centralised governments. Suppose we grant this claim and reject the plausible possibilities that conflicts have either become unlikelier but deadlier or unlikelier unintentionally. When it comes to technologies with the potential for global or existential catastrophe, it is not enough that on average societies have been shifting away from overly centralised power structures and closer to liberal democracies. However, for a worryingly high risk of global catastrophe to remain, it is enough for one rogue nation to gain access to these technologies. The claim by advocates of strong moral enhancement that moral enhancement alone is necessary and sufficient in order to solve our moral failings requires the assumption that everyone will take their moral pills. On the other hand, Buchanan & Powell's claim requires the assumption that every society with access to powerful technologies will be a liberal democracy.

The costs of spreading moral progress across the globe are not trivial. They cite the spread of liberal democracies as a primary example of this. However, one of the latest intentional attempts to spread democracy, the Iraq War, has cost the world over one trillion dollars[15] and hundreds of thousands of human lives.[16] The total wealth possessed by all Irish nationals and their government is less than that value.[17] Just in financial costs alone, implementing democracy in one rogue nation can cost more than the total economic value of one advanced nation. Although the costs are hard to estimate, it is challenging to conceive how implementing a form of moral enhancement that reduces out-group aggression in one rogue nation would have cost more than one trillion dollars or incur more violations of basic rights than the deaths of hundreds of thousands of human beings (of whom a significant percentage were innocent civilians). One might claim that this is an extreme example of failure for traditional moral progress, but we cannot ignore such cases. As mentioned, it is enough to have one rogue nation with access to powerful technologies in order to jeopardise a global effort to reduce extreme risks to humanity.

Buchanan & Powell's claim that a trait's synchronic aspects (the current state of a trait) screen off diachronic aspects (the history of a trait) is given as support for their rejection of "evoliberals" reliance on evolutionary history. It ultimately supports their rejection of a need

for moral enhancement due to evolved inflexibility. However, although the claim is true, it does not actually support their conclusion. They argue that such a claim enables them to say that the current state contains more possibilities than the causal/evolutionary history would suggest, rendering evolutionary explanations less relevant than a trait's current state. But if the diachronic aspects are constrained (screened off) by the trait's synchronic states, the latter must contain fewer possibilities than the former. This is the case even on a colloquial usage of the term. For instance, when patients are screened off from a certain medical evaluation, fewer patients remain at the end of the screening procedure. A constraining process, screening off, reduces the number of possibilities. In the case of evolution, the diachronic explanation entails that current traits must be an evolved solution to some recurrent evolutionary challenge, but there is a wide range of possible solutions to any given proposed challenge. A full synchronic account would provide a complete description of a current trait, thus screening off which one of those possible solutions came to be realised. But if evolutionary theory is correct, then it cannot give an account that is impossible according to evolution. If current states really contained more possibilities than aetiological history would suggest, then we would have to find violations of evolutionary theory everywhere. We would have to find possibilities not suggested by evolutionary theory, traits that would not have evolved as the result of past evolutionary pressures. Rather, we find traits that are the realisation of one possible solution, out of many, to recurrent evolutionary challenges. Of course, most of these solutions are sub-optimal and complex, this is not only within the possibilities of natural selection but to be expected.

Diachronic explanations are of little relevance compared to detailed synchronic explanations when deciding if and how it is technically feasible to carry out a focused intervention. A surgeon intending to operate on a human heart would not know where to cut if he was relying on an evolutionary explanation of why the human heart was selected to have its current configuration. A competent surgeon needs a detailed physiological and anatomical model of the human heart, a synchronic account. Evidently, evolutionary accounts can sometimes help elucidate the possible reasons behind the human heart's current configuration, but they are merely auxiliary. However, when it comes to interventions that will have intergenerational or societal consequences, we often look into history. The distant past might be irrelevant when determining if and how it is technically feasible to change a specific feature of our immune system, but the reasons this feature evolved matter for assessing the long-term consequences of this change. For example, suppose it is the case that dust and seafood allergies are caused by the misfiring of a feature of our immune system that evolved to defend us against

being poisoned by certain animals that went extinct.[vii] Then we have reasons to expect that deactivating such a feature would be beneficial. If, however, we were to discover that this feature evolved to protect us against some once lethal but now dormant pathogen, then we have reasons to expect that deactivating this feature would be dangerous. We might have a full synchronic account of this specific feature of our immune system with all its physiological and molecular detail, but without at least hypothesising about its evolutionary history we would fail to see how such a feature protects us against a dormant pathogen. Only evolutionary history would reveal that this feature evolved to protect us against a currently inactive pathogen. Removing this feature would render our immune system defenceless against a lethal pathogen that might become active in the future. Again, if diachronic aspects are "screened off" by synchronic states, the latter must contain fewer possibilities than the former. Thus, a trait's evolutionary history can inform us even when we have a full synchronic account. When intervening in a feature that can have consequences over long time periods, diachronic explanations become more relevant because they can reveal how that feature is likely to change across time. In that respect, moral progress is more like intervening in the immune system and less like heart surgery. Whether traditional or not, moral progress does not consist of focused interventions whose consequences are immediate and easily identifiable; its repercussions are intergenerational and complex. Moreover, moral progress directly engages with safeguarding and enabling the flourishing of humankind. It is no surprise that more profound and ultimate explanations should be more heavily involved.

Given their extensive use of historical examples, Buchanan & Powell are likely to agree that moral progress is not akin to focused interventions. One way to make sense of their claim that synchronic explanations are what matters when intervening with moral psychology is that human history is so recent that it can be considered to contain only proximate events. Perhaps, for them, whatever happened before human civilisations emerged are distal events which are the concern of diachronic explanations such as human evolution, and whatever happened after are proximate events which are the concern of synchronic explanations such as human history. It seems counter-intuitive to place human history with the set of synchronic explanations, thus putting interventions with historical consequences closer to heart surgery than to changes in our immune system. More importantly, their strict separation between diachronic and synchronic explanations gives insufficient weight to the fact that moral traits have been subjected to evolutionary forces from the Pleistocene period to contemporary history, and that inclusivist moralities are also adaptive strategies co-evolving with more ancient moral traits, and as such are subject to environmental cues only to the extent that they had been selected for

in our most recent evolutionary history. They correctly identify inclusivist tendencies that would arise in the absence of human and non-human threats, but do not mention that such a response, also being an adaptation, cannot be completely freely shaped by cultural moral progress. As discussed, we may have evolved an adaptation that enables us to have a basic respect for out-groups in safe environments, but it seems unlikely that we could have evolved the degree of large-scale co-operation necessary to overcome the extreme risks listed here.

A significant part of the traditional moral progress achieved by human civilisation has also been carried out by changing our genes and, consequently, biochemistry. There is substantial empirical evidence that human evolution has not only continued since the Pleistocene,[18] but also has accelerated.[19] Scientists expect that this was primarily a result of increased sexual selection,[20] which more heavily selects based on social behaviour than other forms of natural selection. There is preliminary evidence that several of those selective forces have, in fact, influenced genes connected to social behaviour[21]; and hence to moral behaviour as well. Therefore, a strict preference for traditional methods instead of technological moral enhancement does not result in leaving the biological basis of morality untouched, but in letting it be manipulated via cultural means alone. If biological changes form part of the mechanisms through which past moral progress has been realised, then the argument against intervening in our moral traits cannot use past successful moral advances as its basis. Altering the biological basis of our moral traits has been one of the means through which moral progress has happened. While moral progress might have been intentional, its consequences for moral traits were not and took centuries to effect. Why should we strongly prefer slow and unplanned alterations over technological interventions? If sexual selection has been one primary driver of the recent changes in our moral traits exerted by traditional moral progress, then rejecting technological means of directly intervening in our moral traits amounts to a strong partiality for sexual selection over technological manipulation as a means of changing moral traits. The argument for preferring sexual selection over careful and intentional technological manipulation allied with traditional methods is difficult to make. It seems unlikely that our sexual preferences alone would be a factor more conducive to human flourishing or reducing extreme risks than sexual preference paired with intentional technological manipulation of moral traits. In fact, Bostrom has argued that in the absence of a globally co-ordinated policy to control human evolution, uncontrolled evolution will lead to the elimination of the kinds of beings we care about.[22]

5. CONCLUSION

I have argued that some degree of flexibility in our moral traits does not undermine a defence of moral enhancement because significant traditional moral progress has come with costs high enough to possibly exceed the costs of implementing moral enhancement. I argued further that the efficacy and safety of traditional moral progress are not uncontroversial facts, thus traditional moral progress may not be enough to deal with modern challenges. Buchanan & Powell claim that the inclusivist anomaly, the broadening of our moral concern, is the major counterexample to moral traits' inflexibility. I have argued that two of their major examples of such anomaly, the abolition of slavery and the decrease of wars, were costly or not cases of moral progress proper.

Surely, moral enhancement's efficacy and safety are also a concern, arguably more than that of traditional moral progress. Properly addressing such concerns lies outside the scope of this paper, but responses abound to the many proposed risks and ethical concerns of moral enhancement. In *The evolution of moral progress*, their two authors state that "relative importance claims are crucial to the evoliberals case" because moral enhancement comes with significant risks and ethical concerns.[2] (p.351) Therefore, they believe the burden of proof should be on moral enhancement advocates. The arguments here indicate the burden of proof is not so asymmetrical. There are important concerns over the efficacy and safety of traditional moral progress that is substantial enough to match the proposed effects of moral enhancement. Buchanan & Powell's argument disfavouring moral enhancement requires traditional moral progress to be substantially safer and more efficient than moral enhancement. This paper argued this gap is not large enough to lead to their pessimism regarding moral enhancement. Perhaps their other examples of moral progress could better support their conclusion than the abolition of slavery and the decrease in violence. However, most of their other examples are recent; some are still on-going and lack detailed data.

The claim that traditional means of moral progress alone are likely to be sufficient to overcome our current challenges seems not significantly more reasonable than the claim that pursuing moral enhancement is a moral imperative. A solution to our moral failings leading to extreme risks is certainly necessary. Until we know a sure solution, no proposal should be deemed sufficient and none of the possibly feasible solutions should be discarded. Buchanan & Powell seem to agree with this general statement but nevertheless argue moral enhancement is an unlikely solution and that traditional moral progress should be preferred.

A strong preference for traditional methods of moral improvement to solve moral failings that lead to extreme risks should be rejected. Addressing these risks is of major moral relevance, but solutions may not come from traditional moral progress alone. Moral enhancement's probable effectiveness in decreasing known sources of extinction risks provides strong reasons for pursuing its development, but its risks still require further investigation.

REFERENCES

1. Persson I, Savulescu J. The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity. *J Appl Philos*. 2008;25(3):162–77. http://doi.wiley.com/10.1111/j.1468-5930.2008.00410.x
2. Buchanan A, Powell R. The evolution of moral progress: a biocultural theory. Oxford University Press; 2018. http://doi.org/10.1093/oso/9780190868413.001.0001
3. Buchanan A. Beyond humanity? Oxford University Press; 2011.
4. Powell R. Convergent evolution and the limits of natural selection. *Eur J Philos Sci.* 2012; http://doi.org/10.1007/s13194-012-0047-9
5. Bostrom N, Cirkovic M, editors. Global catastrophic risks. Oxford University Press; 2011.
6. Koplow DA. Smallpox the fight to eradicate a global scourge. Berkeley, Calif.; London: University of California Press; 2004.
7. Moon S, Sridhar D, Pate MA, et al. Will Ebola change the game? Ten essential reforms before the next pandemic. *Lancet.* 2015;386(10009):2204–21. http://dx.doi.org/10.1016/S0140-6736(15)00946-0
8. Shi Z. Are Groups More Pro-Self Than Individuals? Individual-Group Comparisons on Social Value Orientation and Ethical Decision Making. Loyola University Chicago; 2014.
9. Bostrom N. Superintelligence: Paths, Dangers, Strategies. Oxford University Press; 2014.
10. Bowles S, Gintis H. The Coevolution of Institutions and Behaviors. In: A cooperative species: Human Reciprocity and Its Evolution. Princeton University Press; 2013.
11. Brooks AS, Yellen JE, Potts R, et al. Long-distance stone transport and pigment use in the earliest Middle Stone Age. *Science*. 2018;360(6384):90–4. http://doi.org/10.1126/science.aao2646
12. Braumoeller B. Is War Disappearing? In: APSA Chicago 2013 Meeting. 2013. p. 1–28. https://ssrn.com/abstract=2317269
13. Human Security Research Group. Human Security Report 2013 The Decline In Global Violence: Evidence, Explanation And Contestation. 2013. https://reliefweb.int/sites/reliefweb.int/files/resources/HSRP_Report_2013_140226_Web.pdf
14. Cirillo P, Taleb NN. On the statistical properties and tail risk of violent conflicts. *Phys A Stat Mech its Appl.* 2016;452:29–45. http://dx.doi.org/10.1016/j.physa.2016.01.050
15. Amadeo K. The Ongoing Costs of the Iraq War. *The Balance.* 2018 [cited 2018 Jul 24]. https://www.thebalance.com/cost-of-iraq-war-timeline-economic-impact-3306301
16. Hagopian A, Flaxman AD, Takaro TK, et al. Mortality in Iraq Associated with the 2003-2011 War and Occupation: Findings from a National Cluster Sample Survey by the University Collaborative Iraq Mortality Study. *PLoS Med.* 2013;10(10). http://doi.org/10.1371/journal.pmed.1001533
17. Credit Suisse. Global Wealth Databook 2017. 2017. http://publications.credit-suisse.com/tasks/render/file/index.cfm?fileid=A8BD95FB-A213-1EE7-59CC7F2F001A11AF
18. Courtiol A, Pettay JE, Jokela M, et al. Natural and sexual selection in a monogamous historical human population. *Proc Natl Acad Sci.* 2012;109(21):8044–9. http://www.pnas.org/cgi/doi/10.1073/pnas.1118174109
19. Hawks J, Wang ET, Cochran GM, et al. Recent acceleration of human adaptive evolution. *Proc Natl Acad Sci.* 2007;104(52):20753–8. http://www.pnas.org/cgi/doi/10.1073/pnas.0707650104

20. Harpending H, Cochran G. Genetics and Social Behavior. *Emerg Trends Soc Behav Sci.* 2015;1–15.
21. Richards RA. Sexual Selection: Its Possible Contribution to Recent Human Evolution. In: eLS. John Wiley & Sons; 2016. http://doi.org/10.1002/9780470015902.a0021788
22. Bostrom N. The Future of Human Evolution. In: Death and anti-death: Two hundred years after Kant, fifty years after Turing. Ria University Press; 2004.
23. Buchanan A, Powell R. The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress. *Ethics.* 2015;126(1):37–67.
24. Author 1. XXXXXX
25. Wong L, Huang CH, Lee BW. Shellfish and house dust mite allergies: Is the link tropomyosin? Allergy, *Asthma Immunol Res.* 2016;8(2):101–6. http://doi.org/10.4168/aair.2016.8.2.101

ENDNOTES

i: For very similar arguments but without explicit mentioning of moral enhancement see also [23].

ii: The central thesis of the book is not about moral enhancement itself but about moral progress in general. I will not be concerned with the central thesis. I will only address the arguments that relate to moral enhancement and only in so far as they do. Some arguments that I argue to be unsuccessful in giving reasons for dismissing moral enhancement can still be successful in other regards. My goal is to test moral enhancement against a potentially powerful objection, not to build a case against the book. Moreover, although the book contains a couple of passages suggesting that the use of non-moral enhancement and biomedical interventions could count as traditional moral progress, almost all of its mentions of enhancement and biomedical interventions are negative as evidenced by looking at the book's Index for those terms.

iii: Unless stated otherwise all data cited in this section are from the cited book.

iv: At the very least, it would have helped to advance its legitimate goals with less bloodshed. This bloodshed was by no means trivial even when comparing with other wars of the time, killing 3% of the American population and using particularly gruesome tactics. For instance, the strategy of total destruction during the Atlanta Campaign or the futile massacre of blacks during the Battle of the Crater.

v: For a review of arguments against moral enhancement and their shortcomings see [24].

vi: Here I do not wish to make the already refuted counter-argument that many instances of moral progress were just moral declarations empty of practical consequences. Countries did in fact engage in less armed conflicts, practices did change. But the reason of such a change was not the fact nations were less willing to engage in war than before, it was because they were less physically capable due to, on average, higher geographical separation.

vii: This is merely hypothetical; for a more refined hypothesis attempting to explain links between substances found in both dust mites and shellfish and associated with poisonous organisms see [25].