

# The Duty to be Morally Enhanced

Ingmar Persson<sup>1</sup>  · Julian Savulescu<sup>2</sup>

Published online: 12 April 2017

© The Author(s) 2017. This article is an open access publication

**Abstract** We have a duty to try to develop and apply safe and cost-effective means to increase the probability that we shall do what we morally ought to do. It is here argued that this includes biomedical means of moral enhancement, that is, pharmaceutical, neurological or genetic means of strengthening the central moral drives of altruism and a sense of justice. Such a strengthening of moral motivation is likely to be necessary today because common-sense morality having its evolutionary origin in small-scale societies with primitive technology will become much more demanding if it is revised to serve the needs of contemporary globalized societies with an advanced technology capable of affecting conditions of life world-wide for centuries to come.

**Keywords** Moral enhancement · Cognitive enhancement · Altruism · Justice · Threats to life on earth

## 1 What is moral enhancement

Suppose that you ought morally to do an action A. Then, according to the time-honoured dictum that ‘ought’ implies ‘can’, it follows that you can do A, in a suitable sense. But it may be quite hard for you to do it, so hard that it is more likely than not that you will fail. If there is some means, M,

that would make it easier for you to succeed in performing A, it may be that you ought morally to apply M. Of course, you are not obliged to apply any effective means to a proper end—not means that are unsafe for yourself or others, too costly, or some such. But, assuming that there are no such sufficiently weighty negative side-effects of applying M, then you ought to apply it to make it easier for yourself to do what you morally ought to do, A. This is an instance of the principle that if you ought to attain an end, you ought to apply suitable means to attaining it. If there is no such a means as M readily available, it may also be the case that you ought to try to make it more readily available, so that you could make it easier for you to do what you ought morally to do.

Applying M to yourself might count as morally enhancing yourself, but this is not necessarily so, unless the expression ‘morally enhancing’ is employed in an exceedingly or improperly wide sense. If you improve your ability to swim, or acquire certain rescue equipment, spectacles which enable you to read the instructions to use them, etc. you may make it easier for yourself to save drowning people, which may be what you ought morally to do, but this would not be tantamount to moral enhancement in any useful sense. In order for something to count as moral enhancement, it must enhance your moral *motivation*, your disposition to (*decide* and) *try* to do what you think you ought morally to do, rather than your capacity to *implement or put into effect* such tryings, to succeed if you try. Moreover, it must enhance your disposition to try to do *for its own sake* what you think you ought morally to do, and not because you are in some way *rewarded* if you try to do this, or *punished* if you do not. The presence of effective sanctions are liable to cause people more often to try to do what they think they ought to do, and to refrain from trying to do what they think they ought not to do, but such actions

---

✉ Ingmar Persson  
ingmar.persson@filosofi.gu.se

<sup>1</sup> Department of Philosophy, Linguistics and the Theory of Science, University of Gothenburg, P. O. Box 200, 405-30 Gothenburg, Sweden

<sup>2</sup> Oxford Uehiro Centre for Practical Ethics, Littlegate House, St Ebbs St, Oxford OX1 1PT, UK

in conformity with moral norms do not make people morally better.

What is it that you try to do for its own sake when you try to do what you think you ought morally to do? It is reasonable to hypothesize that this involves, first, trying to do what makes things go as well as possible for beings for their own sake. There are different theories of what things going well for beings consists in. The oldest and most familiar theory is *hedonism*, according to which things going well for beings consists in their having pleasurable experiences of various kinds, and things going badly for them consists in their having painful or unpleasant experiences. Things will go as well as possible for them if and only if the sum of their pleasurable experiences exceeds as much as possible the sum of their unpleasant experiences. It is rather uncontroversial that hedonism captures a part of the notion of things going well for someone, but dubious whether it captures the whole of it. For the purposes of this discussion—in which hedonism is invoked merely to illustrate what could go into the notion of things going well for beings—we need to make only the uncontroversial assumption that it captures part of this notion. To spell out more fully what could go into the notion of things going well for beings would take us too far afield.

To want or be concerned that things go well for beings for their own sake is to have an attitude of *altruism*, *sympathy* or *benevolence* towards them. The fact that this attitude is universally considered to be an element of morality is indicated by the fact that something akin to the Golden rule of Christianity—do unto others as you want them to do you—is recognized by several world religions. The thought underlying this rule would seem to be that we want, for its own sake, that things go well for ourselves; so this is what we should do to others. This attitude is also made the centrepiece of morality by such philosophers as David Hume and Arthur Schopenhauer.

It is well-known that the attitude of sympathy, as it occurs spontaneously, tends to be *partial*: we tend to sympathize in particular with members of our family, friends, and people before our eyes. From an evolutionary point of view, it is to be expected that we exhibit so-called *kin altruism*, i.e. altruism as regards our children, parents, and siblings. Kin altruism is straightforwardly explicable in evolutionary terms, since each child shares 50% of each of its parents' genes and on average 50% of each sibling's genes. Consequently, caring about kin is caring about somebody who carries genes similar to one's own. But we seem also to develop concern for other individuals whom we meet on a daily basis and enter into mutually beneficial cooperation with. Such regular close encounters apparently tend to

breed sympathy and liking, other things being equal, i.e. unless there are special reasons for aversive feelings such as hostility, fear, disgust, and so forth. This is sometimes called *the exposure effect*. We are also capable of momentarily feeling quite strong sympathy or compassion for beings who suffer before our eyes. But we have little sympathy for strangers, distant in space or time.

Our sympathy or altruism is, then, limited or parochial by nature. Utilitarianism, which takes sympathy or altruism to be the one and only fundamental moral attitude, opposes this partiality, by declaring in its most familiar form, roughly speaking, our moral goal to be to see to it that things go as well as possible for as many as possible. This goal takes into account every being for whom things can go well or badly—which, according to hedonism, will be beings who can have pleasant and unpleasant experiences—and sees to it that things go well or badly for them in so far as this contributes to maximizing the overall sum of well-being. However, such a maximizing goal is compatible with things going much better for some than others, and this will strike many of us as *unjust or unfair*. We have therefore postulated another moral attitude, alongside that of sympathy or altruism, namely *a sense of justice or fairness*. Accordingly, the moral goal is not just that things go as well as possible for as many as possible, but that how well things go for different beings be as much as possible in line with justice.

What the object or purport of a sense of justice or fairness consists in is however one of the most controversial issues in moral philosophy. Some claim that it consists in getting what you deserve, others in having what you have a right or are entitled to, and still others in equality. Without resolving this controversy, we can ascribe to people a desire to do *what they think* is just or fair, alongside their capacity for altruism, sympathy, or benevolence. We have proposed that moral enhancement consists in boosting the strength of these two attitudes (2012: 108–9).<sup>1</sup>

Now, by what means can these attitudes be effectively enhanced? We have argued that they are biologically based and, therefore, that they are amenable to modification by biomedical means, pharmaceutical, neurological or genetic. In particular, the fact that on average altruism is stronger among women than men is evidence that it is biologically based (see e.g. Baron-Cohen 2003). It does not seem that altruistic concern about the welfare of others for their own sake or about justice could easily be strengthened by traditional moral education and reflection. Consider people who

<sup>1</sup> To a considerable extent, this paper summarizes the account we give in our book *Unfit for the Future* (2012), but in addition to the book we have published a dozen or so papers on moral enhancement, developing or defending this account, some of which we shall refer to in due course.

exhibit little or no altruistic concern, like psychopaths: it has been found that they cannot be induced to have more concern for others by cognitive therapy or reasoning. Biomedical means offer the promise of help on this score, but research into such means has only just began.

One of the most interesting scientific findings so far might be the hormone oxytocin. Oxytocin is naturally elevated by sex and touching, but it can also be boosted by nasal spray. It facilitates maternal care, pair bonding, and other pro-social attitudes, like trust, sympathy and generosity. Frustratingly, however, oxytocin's effect on trust and other pro-social behaviour towards other people appears to be sensitive to the group membership of these people. Experiments indicate that people who receive oxytocin are significantly more likely to sacrifice a different-race individual in order to save a group of race-unspecified others than they are to sacrifice a same-race individual. The effect of oxytocin might then be to increase sympathy only towards members of one's in-group, and not to expand the bounds of our spontaneously limited altruism.

Interesting studies of drugs that keep up the level of serotonin have also been conducted, but it should be emphasized, again, that research in the area of moral enhancement by biomedical means, *moral bioenhancement*, as we have called it, is still in its infancy, and it is too early to judge its prospects. However, in accordance with the argument above, we have a moral duty to try to develop biomedical means—as well as other means—to moral enhancement, and to apply them to ourselves if safe and effective means are discovered. Later in this section and in the next, we shall discuss how urgent these endeavours are.

Against moral bioenhancement, it has been objected by some, like Harris (2010), that it would undercut the reasoning capacity and freedom of people subjected to it. Strictly speaking, the claim that it undercuts moral reasoning and that it undercuts freedom are distinct claims, but they are run together by Harris when he writes that moral bioenhancement will 'make the freedom to do immoral things impossible, rather than simply make the doing of them wrong and giving us moral, legal and prudential reasons to refrain' (2010: 105). As we shall soon see in some detail, moral bioenhancement does not rule out moral reasoning; on the contrary, it should be supplemented with it. It is however intended to increase the probability that we shall do what, on the basis of moral reasons, we think that we ought to do. But, according to Harris: 'The space between knowing the good and doing the good is a region entirely inhabited by freedom... We know how lamentably bad we are at doing what we know we should' (2010: 104). In other words, we know that we are regrettably weak willed.

Throughout the history of philosophy it has been hotly debated what having a free will amounts to. Some philosophers take our will to be free in the more contentious sense

of not being causally determined, while other philosophers take our freedom to be compatible with causal determinism reigning in the domain of mind and action. Suppose, first, that our freedom is compatible with determinism. Then a judicious use of moral bioenhancement techniques will not reduce our freedom; it will simply make it the case that we are more often, perhaps always, determined by our moral reasons rather than by other factors to do what, according to these reasons, we ought to do. We would then act as morally perfect persons now act and would not be any less free than they are.

Suppose, on the other hand, that we are free only because, by nature, our decisions and actions are not fully causally determined by anything. Then moral bioenhancement cannot be fully effective because its effectiveness is limited by the causal indeterminacy which constitutes our freedom. So, irrespective of whether determinism or indeterminism holds in the realm of human action, moral bioenhancement cannot curtail our freedom.

Turning now to moral reasoning, moral bioenhancement could not by itself be sufficient without it. This is particularly obvious in the case of the sense of justice. Needless to say, it takes moral reasoning to determine what theory of justice is the correct one: whether justice consists in getting what you deserve, what you have a right or are entitled to, in an equal distribution of what makes things go well, or something else. Perhaps it will take enhancement by biomedical means to enable our reasoning power to resolve this recalcitrant debate, but if so, this would be *cognitive* rather than moral bioenhancement.

Common sense is firmly wedded to deserts and rights, whose application seemingly leaves little space for equality. We have hypothesized that the concept of desert originates in reciprocity and the tit-for-tat strategy which consists in paying back in the same coin: responding to favours by favours, and ill-treatment by ill-treatment (2012: 34). The idea of rights to ourselves and property acquired is traceable to the special ferocity with which animals defend themselves, their turf, food and so on (2012: 19). With such a background in our evolutionary past, it is to be expected that these notions are deeply embedded in common-sense morality. As a result, the view that the justice requires receiving what you deserve and have a right to has a tight hold on our minds. Consequently, a theory of justice which allots to an equal distribution of welfare a more wide-ranging application than the gaps left by deserts and rights will amount to an extensive revision of common-sense morality. Be that as it may, for present purposes the salient point is that biomedical enhancement of our desire to do what we regard as just rather than undermining the role of moral reasoning and reflection needs their assistance.

The same goes for altruism: here moral reasoning and reflection have the essential function of determining its

proper range, what kinds of beings it should encompass. Nowadays, most people accept that the distinctions between human races and sexes are morally irrelevant, i.e. that racism and sexism represent immoral discrimination. It is much more controversial to maintain that a speciesism which assigns a higher moral status to members of our own species, *Homo sapiens*, is also indefensible, so that our altruism should extend with full force to non-human animals. An altruism of such an extensive scope, which covers all sentient beings, may necessitate biomedical enhancement for, as remarked, by nature it operates within very narrow confines. The evolutionary explanation of why this should be so is ready at hand: if our altruism and disposition to cooperation extended indiscriminately to strangers, we would be exposed to a grave risk of being exploited by free-riders. Suspicion against strangers is called for, since human beings often try to get the better of each other. Consequently, it is not surprising that xenophobia is a widespread characteristic of humanity. On the other hand, as already noted, studies of the moral bioenhancer oxytocin show that it needs to be seconded by moral reflection demonstrating the moral irrelevancy of various in-group/out-group barriers in order to prevent that the more intense sympathy retains its restricted scope.

Our spontaneous altruism is restricted in another way: the sheer *number* of subjects to whom we have to respond can present an obstacle to our adoption of a proper response. While many of us are capable of vividly imagining the suffering of a single subject before our eyes and, consequently, of feeling strong sympathy or compassion for this subject, we are unable vividly to imagine the suffering of several subjects even if they be in sight. Nor could we feel a sympathy which is several times as strong as the sympathy we could feel for a single sufferer. Rather, the degree of our sympathy is likely to remain more or less constant when we switch from reflecting upon the suffering of a single subject to the suffering of, say, 10 or 100 subjects. Yet the efforts of relieving the suffering of 10 or 100 subjects may well be 10 or 100 times as high as the efforts of relieving the suffering of one subject. Therefore, it is not surprising that, as the number of subjects in need of aid increases, the amount of aid we are willing to give to each subject decreases. Nevertheless, if we thoroughly enough contemplate the fact one group of sufferers is larger than another, we can feel a *somewhat* greater sympathy for the first group, so that we shall be moved to assist them rather than the other group. But this presupposes that the quantity of the sympathy of which are capable is not so slight that it is expended on those who are nearest and dearest to us.

A further plausible object for inclusion within the ambit of altruism is future generations. Since the threats that are most urgent to deal with in order to survive and reproduce tend to be located in the immediate future, and

through most of our history we have had tools to affect only the immediate future, we exhibit a *bias towards the near future*, i.e. we are more concerned with positive and negative events that happen in the nearer than in the more distant future. This bias manifests itself when we are relieved if something unpleasant due to happen to us in the immediate future is postponed, and disappointed if something pleasant in store for us is postponed. The bias towards the near future is not a discounting of possible future events in proportion to how probable we estimate them to be. For we could be greatly relieved when an unpleasant event, such as a painful piece of surgery, is postponed for just a day, even though we take this delay to make it only marginally less probable. To the extent that our lesser concern for what is more distant in the future is out of proportion to its being estimated as less probable, it is arguably irrational. The bias towards the near future is frequently the explanation of why we exhibit weakness of will by choosing, against our better judgement, to have a smaller good straightaway rather than to wait some extra hours for a greater good. It will soon emerge that there are other reasons than such reasons of prudence or self-interest why it is useful to overcome this bias. This is a further task for our powers of moral or practical reasoning.

Yet another function reserved for these powers is to settle the soundness of deontological principles like *the act-omission doctrine*—i.e. the doctrine that it is harder morally to justify doing harm than omitting to benefit—and *the doctrine of the double effect*, i.e. the doctrine that it may be permissible to cause harm as a side-effect when it is not permissible to cause this harm as a means to an end, or as an end in itself. It is especially the rejection of the act-omission doctrine that will make a great difference in practice. This is due to the fact that those of us who live in the more affluent parts of the world have resources to relieve a lot of the vast amount of suffering which occurs in poorer parts of the world.

We have hypothesized that the act-omission doctrine involves *a conception of responsibility as causally based*, a feeling that we are responsible for an event in proportion to our causal contribution to it. We do not see ourselves as causes of what we let happen, so we feel little responsibility for it. The notion of responsibility as causally based is proportionally diluted when we cause things together with other agents, e.g. when we together destroy a lawn by each of us walking across it from time to time, since our own causal contribution to the deterioration of the lawn then decreases compared to what it would have been had we destroyed the lawn single-handedly. The parallel to our minimal sense of responsibility for anthropogenic climate change should be obvious.

Together with the other upshots of revisionist moral reasoning that we have considered—the extension of moral

concern to include all sentient beings, now and in the future, and a conception of justice as requiring equality for all these beings—the rejection of deontology will yield a morality which makes heavy demands on those of us who are fortunate enough to be well off. Thus, it is very far from the case that moral bioenhancement, as we conceive it, replaces or renders redundant moral reasoning and reflection. Rather, they must work in tandem in order for us to undergo moral enhancement in the sense of becoming better at trying to do not merely what we think that we ought morally to do, but what we really ought morally to do. Reason instructs us about what we really ought morally to do, and as this might well be more taxing than what common-sense morality lays down—which is often hard enough for us to accomplish—moral bioenhancement is called for to increase our chances to make whole-hearted attempts to do what we really ought morally to do. It will soon transpire that modern scientific technology further exacerbates the demandingness of our moral obligations.

However, before we turn to this matter in the next section, it is worth noting that the likelihood of our living up to a demanding morality can be increased in a less direct way than by boosting altruistic concern and a concern for justice. Human beings have a number of strong motivational dispositions that are liable to interfere with their acting out of these central moral dispositions. Consider, for instance, the so-called seven deadly sins: gluttony, greed, lust, envy, wrath, pride and sloth. There are plausible evolutionary explanations of why humans have strong desires for food, material possessions, sex, social reputation and status, relaxation, and negative reactions of envy and wrath towards those who compete with them for these goods. But these desires can obviously obstruct humans acting altruistically and justly. Consequently, if these motivational dispositions are suitably modified by biomedical means, this could count as moral bioenhancement in a wider sense. It would not be moral bioenhancement in a narrower sense because it is not modification of attitudes whose objective is things going well for other beings. Analogously, bioenhancement of some traits like courage, conscientiousness and steadfastness could be said to be tantamount to moral bioenhancement in a wider sense to the extent that it renders us more liable to act in accordance with the central moral dispositions of altruism and a sense of justice and try to do what we ought morally to do.

## 2 Why moral enhancement is especially urgent now

In Sect. 1 we have argued that we ought morally to seek and apply suitable means to make it easier for us to do what we ought morally to do. Additionally, it has been seen that

called for revisions of common-sense morality may make morality more demanding and, so, make the discovery of such means more urgent. We would now like to argue that modern scientific technology has created a situation in which morality has become more demanding especially for people in affluent nations, thus making it even more urgent to seek and apply suitable means to make it easier for us to do what we ought morally to do. We shall also see that this technology gives rise to a need to amplify our inhibition against doing things that we morally ought not to do.

It is easier for us to harm each other than it is for us to benefit each other. To give an everyday illustration: most of you probably have access to a car and live in densely populated areas. Whenever you drive, you could easily kill a number of people, by ploughing into a crowd. But very few, if any, of you have the opportunity every day to save the lives of an equal number single-handedly. Indeed, it might be that none of you have ever had that opportunity, since this kind of situation obtains only when, first, a large number of lives is threatened, and, secondly, you are also in a position to eliminate that threat. The claim is not that we are *never* capable of saving as many individuals as would die if a threat were not successfully foiled. It is that in order to save such a number of lives, we have to find ourselves in situations in which these lives are under a threat that we could avert, and this is a comparatively rare event often beyond our control. By contrast, we frequently have the opportunity to kill many on our own.

We could distinguish between two related aspects of the greater easiness or power to cause harm. First, the *magnitude* of the harm we can cause can be greater than the magnitude of the benefits we can provide: e.g. we can generally kill more individuals than we can save the lives of, wound more than we could heal the wounds of, and cause pain that is more intense than the pleasure that we could cause.

Secondly, there are normally many more *ways or means* of causing harm of a given magnitude than there are ways of benefiting to the same degree. This is because there are more ways of disturbing a well-functioning system, like a biological organism, than of improving it to the same extent. Thus, arbitrary interferences with well-functioning systems are much likelier to damage them than to improve them. Their degree of organization or integration tends to decrease in the course of time because most changes in them will damage them. This is a part of what is known as *entropy*. If we remove *any* of the countless conditions which are necessary to maintain the functioning of an integrated system, we shall interrupt its function, but in order to improve its function, we shall have to discover a condition which fits in so well with all or most of these conditions that the function is enhanced. Such conditions are likely to be far fewer, so this task is much harder.

This is why it is in general easier to kill than to save life. But, imagine, contrary to the present argument, that it would be as easy to save life as to kill; it would still not follow that, if we save a life, we could claim credit for as much life-preservation as we are guilty of life-destruction if we end it. This is again because there are countless conditions which are necessary for an organism to remain alive. If we remove any of these conditions, we are guilty of ending the life forever. But if we prevent the removal of such a condition, we cannot claim the whole credit for the good things that continuation of this life contains, since there are other conditions which are necessary to ensure them.

Suppose that life is good for the organism as long as it lasts. If we remove any of the conditions which are requisite to sustain it, we kill the organism, thereby depriving it of all the future good that its life would have contained had it not been ended. Thus, by removing any of those conditions we are guilty of causing it a harm which equals the loss of the goodness of which it is deprived. But if we had instead saved the organism from death at the same time, we cannot claim credit for all the good that the future has in store for it, since this saving is only one of indefinitely many conditions which are necessary for it to lead this good life in the future. Consequently, the benefit we would bestow upon an individual by saving its life at a time would be less than the harm we would do it were we to kill it at the same time, for our saving is not sufficient for it to receive the future good life, whereas the killing is sufficient to deprive it of it. Therefore, even if it had been as easy to save life as to kill, which we have denied, it would still not be true that our capacity to benefit would be as great as our capacity to harm by these means.

Now, as scientific technology increases our powers of action, the easiness of harming is magnified. Of course, our capacity to benefit also increases, but the power to harm maintains its clear lead. With the invention of nuclear weapons during the last century our power to harm reached the point at which we could cause what can be called *ultimate harm*, which consists in making worthwhile life *forever* impossible on this planet. Since such a harm would prevent an indefinitely large number of worthwhile lives that would have been led in the future had it not occurred, its negative (instrumental) value is indefinitely high.

To fabricate a nuclear bomb out of fissile material, such as highly enriched uranium or plutonium, is comparatively difficult, though it might in the imminent future be within the capacity of a well-organized terrorist group. Biological weapons of mass destruction are far easier to fabricate than nuclear weapons—indeed, a single individual could do so. To illustrate, some scientists in Australia inadvertently produced a strain of mousepox that is lethal in almost 100% of mice. This study of the genetic modification of mouse pox was published on the internet, making it indiscriminately

available. Mousepox is similar to human smallpox. Knowledge of such experiments could enable a group of terrorists to genetically engineer smallpox to produce a new strain with a mortality of near to 100% instead of 30%, and with a resistance against current vaccine. These terrorists could then fly around the world and disseminate their product. Since the incubation period of smallpox is one to two weeks, the disease would have spread widely before it was even detected, and even after detection there would be no effective way of preventing further dissemination.

The expansion of technological prowess is likely to put in the hands of an increasing number of people such weapons of mass destruction. Now, if an increasing number of us acquires the capacity to destroy an increasing number of us, it is enough if very few of us are malevolent or deranged enough to use this power for all of us to run a significantly greater risk of death and grave injury. Killing with weapons of mass destruction is usually psychologically less difficult because it is done at a distance and, so, does not activate the inhibitory mechanisms that the sight of blood and guts constitutes. As a substitute for this instinctive repulsion against killing, we need to strengthen our sympathy and sense of justice.

The advance of scientific technology has also produced another kind of threat to our survival. It has produced an explosion of the human population and its colonization of the whole planet, by giving it knowledge of how to make an extensive use of natural resources. The human population is now over seven billion and is expected to grow to around ten billion by 2050. Population growth is bad enough, but it is coupled with a sharp rise in consumption in some populous countries like China, India and Brazil, bringing them closer to the standard of living in Western democracies. The human impact on the Earth is a function of three variables: the size of the human population, the average level of welfare, or the GDP per capita, and the efficiency of technology, i.e. how much welfare it could generate out of natural resources. ‘Overshoot Day’, that is to say, the day when we have consumed more than the Earth produces in a year and released more waste than it can absorb, has in the last years occurred alarmingly early—in 2015 earlier than ever, on August 13. This means that in 2015 humans used up approximately 1.4 times more than what the Earth can provide in the same period of time. Clearly, this overconsumption is untenable, but it seems unlikely that we could stop it just by making technology more efficient, for to achieve sustainability, technology must be made *radically* more efficient. Besides, there is the problem that if technology is made more effective, the surplus tends to be spent on more consumption.

So, the astonishing progress of scientific technology has not produced the bright future prospects for humanity that one might have hoped. Quite the contrary, the future of

humankind looks darker than ever. The prominent British scientist Martin Rees estimates that ‘the odds are no better than fifty–fifty that our present civilisation on the Earth will survive to the end of the present century’ (2003: 8). Such an estimate would have been wildly implausible with respect to any other 100 year period before 1950s, before humans acquired a capacity to blow up the Earth with nuclear weapons, and when only eruptions of super-volcanos or hits by massive asteroids presented such catastrophic threats. It then seems indisputable that contemporary scientific technology has markedly increased the risk of world-wide catastrophe, even if Rees’s estimate of the risk is exaggerated. Perhaps human civilization will end sometimes this century in a war with weapons of mass destruction over the dwindling resources of this planet. On the other hand, we cannot rid ourselves of our advanced technology because we need it, and even improvements of it, to provide a huge—and increasing—human population with a decent standard of living without depleting the resources of the planet. It is unacceptable to let billions of human continue to live in misery. Thus, we face a dilemma: we need sophisticated technology for the foreseeable future, though it comes with a horrifying risk.

We face this dilemma because we are not capable of handling this powerful technology in a morally responsible way. Technology has progressed so quickly that there is now a huge mismatch between our technological and moral capacity. As already indicated, it is reasonable to hypothesize that our moral psychology has been shaped by evolution to suit entirely different social circumstances than the current ones. For most of their 150,000 year long history, human beings lived in small, close-knit societies with a primitive technology that allowed them to affect only their most immediate environment. Therefore, evolution is likely to have made human beings psychologically myopic, disposed to care more about what happens in the near future to themselves and some individuals who are near and dear to them. Since we now have the power to perform actions which affect sentient beings world-wide and for centuries into the future, our moral responsibility and obligations have become overwhelmingly more extensive. This is especially so if some of the revisions of common-sense morality reviewed in the preceding section are accepted. For instance, if the act-omission doctrine is rejected, the failure of affluent people to aid those who live in poverty and misery will be more gravely wrong. And if the conception of justice is revised in an egalitarian direction, we have moral reasons to strive not only for a more equal distribution of welfare for the present generations, but for a more sustainable way of life which leaves future generations with as many resources as the present generations possess.

Now, we do not deny that during the course of human history there has been some measure of moral progress;

for instance, the idea that all human beings are of equal worth is now more widely spread than ever before. But to no small extent people pay only lip-service to this doctrine; their behaviour often gives evidence of racial discrimination, and in recent times we have witnessed even outbreaks of genocide, e.g. in ex-Yugoslavia and Rwanda. Xenophobia lurks beneath the civilized surface.

Since human moral development has been relatively modest so far in the course of history, and since such a great moral improvement in short time seems necessary for us to handle responsibly the enormous powers of modern scientific technology, it is imperative to put a lot of effort into research on biomedical means of moral enhancement, as a supplement to intensified moral education of a traditional sort. With the help of the means that scientific technology has put in our hands we have altered the conditions of life on earth as radically as the most potent of natural forces. We now need to turn these means to our own nature to change it to fit in with these altered circumstances.

Some of our critics have fastened on the problem that we call the ‘bootstrapping’ problem (2012: 124): moral bioenhancement must be deployed by human beings who are morally imperfect, so there is an ineliminable risk that this technique will be misused, as other pieces of scientific technology have been misused. There is indeed no guarantee that this technique will not be misused, but if we do not try to enhance ourselves morally, we run a grave risk of undermining our own existence. It should be kept in mind that we would like humankind to survive not merely to the end of the present century—which, according to Rees, is less likely than not. We want humankind to survive for many centuries with advanced technology. Moral bioenhancement might well be necessary to ensure this. Although it may take decades before effective means of moral bioenhancement are discovered, and although its implementation may be quite risky in the beginning, it is likely to get more secure the longer it is been implemented, and the more morally enhanced we become.

As we have argued elsewhere (2013, 2014, 2015), we are defending a *cautious* as opposed to a *confident* proposal about moral bioenhancement. Our proposal is cautious in that it admits (1) that moral bioenhancement is not fool-proof, that it can be misused as other scientific technologies, (2) that it is not sufficient by itself to accomplish a necessary moral enhancement of humanity; traditional methods of moral instruction are also needed, as well as suitable reforms of laws and social institutions, and (3) that the science of moral bioenhancement is still in its infancy, so it is too early to tell whether any effective and safe biomedical means to moral enhancement will be found. We argue merely that the need for moral enhancement is so great that it is important that the possibility of moral enhancement by biomedical means is explored. Such

means are possible in principle because the central moral dispositions of altruism and a sense of justice are biologically based, and the use of such means do not undermine the freedom requisite for morality.

Although being morally enhanced by biomedical means does not necessarily make anyone less free, we do not rule out that moral bioenhancement could be justifiably imposed without the informed consent of the subjects. In particular, it could be so imposed on children where it is likely to be most effective. But fostering children by traditional methods also inevitably involves a massive coercion of children for their own good. In fact, coercion in the shape of effectively implemented laws is an inescapable feature of functioning social life. Even liberal democracies, in spite of their avowed neutrality in matters of ideology and value, penalize killing, raping, stealing, and so on. They would fall apart if they did not. So, coercion *per se* is not morally wrong, and some instances of moral bioenhancement could be among the cases in which it is justified.

**Acknowledgements** Julian Savulescu was supported by the Wellcome Trust: grant number WT104848/Z/14/Z.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/)), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Baron-Cohen S (2003) *The essential difference*. Basic Books, New York
- Harris J (2010) Moral Enhancement and Freedom. *Bioethics* 25:102–111
- Persson I, Savulescu J (2012) *Unfit for the future*. Oxford University Press, Oxford
- Persson I, Savulescu J (2013) Getting moral enhancement right: the desirability of moral bioenhancement. *Bioethics* 27:124–131
- Persson I, Savulescu J (2014) Against fetishism about egalitarianism and in defence of cautious moral bioenhancement. *Am J Bioeth* 14:39–42
- Persson I, Savulescu J (2015) *The art of misunderstanding moral bioenhancement*. Cambridge quarterly of healthcare ethics 24:48–57
- Rees M (2003) *Our final century*. Heinemann, London