THE ROLE OF FREEDOM FOR MORAL AUTONOMY IN THE PROJECT OF MORAL BIOENHANCEMENT

SILVIYA SERAFIMOVA

Institute of Philosophy and Sociology, Bulgarian Academy of Sciences

silvija_serafimova@yahoo.com

Abstract

The main objective of this article is to analyze why both Harris's criticism to Persson and Savulescu's project of moral bioenhancement (MBE) and his own theory of freedom require the reconsideration of the latter's role for moral autonomy. In this context, I explore how Persson and Savulescu's thought experiment of the God Machine underrates the complexity of human moral self-development.

Keywords: moral bioenhancement (MBE), freedom to fall, the God Machine, moral improvement, moral self-development

Background

In their 2008 seminal paper (Persson and Savulescu, 2008), Persson and Savulescu introduce the concept of moral bioenhancement (MBE) [1]. They argue that if humanity wants to eliminate the danger of ultimate harm, which will make "worthwhile life *forever* impossible on this planet" due to the greater ease or power of causing harm rather than providing benefits triggered by contemporary scientific progress (Persson and Savulescu, 2017: 12), humans should use the benefits of biomedical means for moral enhancement. Certainly, one should take a step from social liberalism in search for what Persson and Savulescu call globally responsible liberalism which "extends welfare concerns globally and into the remote future" (Persson and Savulescu, 2012: 102). Nor should one question the statement that the solution to environmental problems cannot be defined as a purely technological one (Ibid.: 104). The main assumption underlining the aforementioned specifications is that cognitive enhancement is insufficient, unless it is accompanied by a corresponding moral enhancement [2].

However, what raises some concerns in moral terms is Persson and Savulescu's claim that the insufficiency of our traditional means of moral enhancement should be supplied with

biomedical means. Regardless of gradually moderating their views in time, viz. saying that MBE research is still in its infancy and only some safe and effective means (if any) should be compensated in compliance with the traditional means, Persson and Savulescu emphasize the biological basis of altruism and sense of justice as central moral dispositions.

Without going into detail, since the development of the MBE debate is quite vast and is not an objective of this paper [3], I will give only two examples demonstrating why such an interpretation cannot be accepted uncritically. Persson and Savulescu argue that it does not seem that altruistic concern about others' welfare for their own sake or about justice could easily be strengthened by traditional moral education and reflection. In this context, adopting biomedical means may "offer the promise of help on this score, but research into such means has only just began" (Persson and Savulescu, 2017: 9). Among other issues, the question is whether psychopaths can be set in one and the same group with the people who exhibit little altruistic concern, taking into account that Persson and Savulescu introduce the group of the few people who may pose an ultimate harm as a homogenous group.

Discussing the insufficient results of giving oxytocin and SSRI for encouraging sympathy, trust and collaboration, they argue that the intake of oxytocin increases sympathy only towards members of one's in-group (Ibid.). In turn, the intake of SSRI does not provide sufficiently convincing results either. "However…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." (Ibid.). Specifically, it remains unclear how the normative reasons behind the moral obligation to develop biomedical means can be defined, unless there are empirical proofs that we are on the right track.

Some concerns about opening the discussion about the role of altruism and sense of justice as biological dispositions are extended to Persson and Savulescu's claim that moral enhancement has been significantly delayed compared to other aspects of technological progress. According to them, traditional measures of moral enhancement "could not be accomplished to a sufficient degree in time to avert disastrous misuses of modern technology" (Persson and Savulescu, 2012: 106). This is due to the fact that "the degree of moral improvement in the 2,500 years…is nowhere near matching the degree of technological progress during the same period" (Ibid.).

Judging by this definition, I would argue that it brings two major concerns to light: 1) that moral improvement (enhancement) is considered as so slight that it cannot keep up with technological improvement and 2) that traditional moral improvement (enhancement) cannot succeed in making technological improvement safe and morally acceptable so that humanity can avoid coming to ultimate harms, because it has not succeeded in preventing them so far.

Regarding the first concern, it remains unclear what one should understand by degree of moral improvement in this case – whether this is the moral permissibility of building and using technology, the moral permissibility of its speed of development or the general permissibility of technological development as such. Another issue is whether or not we can extrapolate the concept of technological improvement to morality, and then argue for moral improvement rather than moral development.

Regarding the second concern, I suggest that Persson and Savulescu's criticism of the role of availability bias is relevant to their own specification as well. They spontaneously imagine that what has not been observed [regarding traditional moral enhancement] resembles that which has been observed in the past (Ibid.: 18).

The problem of freedom in the MBE debates

One of the most frequently cited discussions about the pro and con Persson and Savulescu's project of MBE is that between them and J. Harris. In addition to the different stances on some reliable methods of traditional moral enhancement and the different interpretations of the impact of cognitive enhancement [4], one should examine the crucial role of freedom for moral autonomy. According to Harris, "the space between knowing the good and doing the good is a region entirely inhabited by freedom" including the "freedom to fall" (Harris, 2010: 104). Without the latter, "good cannot be a choice" and "freedom disappears and along with it virtue" (Ibid.). In this context, he specifies that "the sorts of traits or dispositions that seem to lead to wickedness or immorality" are also the same ones "required not only for virtue but for any sort of moral life at all" (Ibid.).

In turn, Persson and Savulescu's critical response, as displayed in a series of publications (Persson and Savulescu, 2012; Persson and Savulescu, 2012a; Persson and Savulescu, 2017) is focused upon what they recognize as contradictions in Harris's interpretation (Harris, 2010). According to Persson and Savulescu, when arguing that the

application of MBE would make one's freedom to do immoral things impossible, Harris wrongly intermingles the claim that MBE undercuts moral reasoning with the one that it undercuts freedom (Persson and Savulescu, 2017: 9). Persson and Savulescu argue that MBE does not rule out moral reasoning, but it should be supplemented with it (Ibid.). The assumption is that MBE can increase the probability that "we shall do what, on the basis of moral reasons, we think that we ought to do" (Ibid.).

Specifically, Persson and Savulescu set the debate about freedom against the background of two major hypotheses: 1) if freedom is compatible with determinism, then MBE techniques will not reduce our freedom because we are always determined by our moral reasons rather than by other factors and 2) if we are free only because our decisions and actions are not fully causally determined by anything, then MBE cannot be fully effective because its effectiveness will be limited by the causal indeterminacy of our freedom (Ibid.). Therefore, irrespective of the choice of determinism or indeterminism, MBE cannot curtail our freedom (Ibid.).

Such an explanation works well merely in functional terms. However, what is of major concern in this context is the particular contribution of MBE. The latter does not curtail freedom, but nothing guarantees that it can contribute to the debate either. Tackling the first option, there is no reason for one to believe that if our freedom of moral decisions is predetermined by some other factors, the application of MBE (even if safe and effective at some point in the future) will have an impact at all. The same conclusion, although with the opposite 'content', can be drawn from the alternative: if we are free because our decisions are not fully causally determined, who needs MBE and not something else?

Against the background of the aforementioned investigations, the main objective of this article is to analyze why the crucial challenges derived from the potential use of MBE regarding the role of freedom, require analyzing the relationship between moral development and moral self-development as irreducible to moral improvement. In this context, I examine how Persson and Savulescu's thought experiment of the God Machine would question the complexity of moral self-development. Specifically, I also explore how Persson and Savulescu's analysis of the similarities between moral bioenhancement and traditional moral enhancement regarding freedom and its diminishment can result merely in justifying the role of collective moral improvement.

The role of freedom for the God Machine experiment

Persson and Savulescu define the God Machine thought experiment as underlined by the necessity of obliterating immoral behavior (Persson and Savulescu, 2012a: 408). According to the experiment, it is 2050 when the science of morality is far advanced (Ibid.). The thought experiment describes a futuristic reality where genetically modified neurons (GMNs) contain 'nanosignalers'. GMNs emit "signatures" of light being controlled via light "in precisely the same range, not visible to the human eye" (Ibid.). Information is "transmitted to bioquantumcomputers that are trillions of times as intelligent and fast as the most powerful supercomputer earlier in the millennium" (Ibid.).

In this reality, the so-called Great Moral Project plays a crucial role. It was completed in 2045, involving "the construction of the most powerful, self-learning, self-developing bioquantum computer ever constructed called the God Machine" (Ibid.). The main function of the God Machine was "to monitor the thoughts, beliefs, desires and intentions of every human being" by modifying these within nanoseconds, "without the conscious recognition by any human subject" (Ibid.). Consequently, the God Machine was developed so that it could give human beings "near complete freedom" (Ibid.). The latter was grounded in the assumption that it "only ever intervened in human action to prevent great harm, injustice or other deeply immoral behaviour from occurring" (Ibid.).

Persson and Savulescu give the example with the avoidance of murdering innocent people. Specifically, the lack of murderers was triggered by the intervention of the God Machine which was activated when a person formed the intention to murder, and eliminated it. Thus, the "would-be murderer would 'change his mind.'" (Ibid.). In contrast, the "God Machine would not intervene in trivial immoral acts, like minor instances of lying or cheating." (Ibid.) "It was only when a threshold insult to some sentient being's interests was crossed would the God Machine exercise its almighty power." (Ibid.).

While the aforementioned description of the project reveals the role of MBE for one's negative right of not to be killed and the general prohibitive function of the God Machine as a moral project, it also addresses one's positive right to moral improvement by biomedical or other means. Altruism and sense of justice are so strong that people almost never decide or choose to act immorally (Ibid.). Persson and Savulescu emphasize that "Human beings can still

autonomously choose to be moral, since if they choose the moral action, the God Machine will not intervene. Indeed, they are free to be moral. They are only unfree to do grossly immoral acts, like killing or raping" (Ibid.: 409).

Putting aside the wide discussion about the biological origin of altruism and sense of justice in Persson and Savulescu's writings, I would argue that the 'positive' perspective of the God Machine faces the pitfalls raised against Persson and Savulescu's explanation of the role of freedom within the frameworks of determinism and indeterminism. To be more precise, the concern is, once again, about what Harris defines as freedom to fall versus Persson and Savulescu's statement that people "almost never decide or choose to fall" (Ibid.).

While in Harris's explanation the freedom to fall is underlined by the general human potential for moral development as a crucial part of one's moral autonomy, in Persson and Savulescu's case the choice of arguing for improvement contradicts the role of moral development and moral self-development. Thus, the freedom of moral choice is reduced to what is already predetermined to be chosen as moral. In other words, the 'positive' vision of the God Machine is justifiable only if freedom is recognized within the framework of determinism. In contrast to Harris's case, here the moral reasons are determined by the God Machine. This makes one argue that being free to morally choose what is predetermined to be chosen as moral leads to saying that the choosing agent is free to choose.

Another concern stems from Persson and Savulescu's assumption that if biomedically increased altruism and sense of justice are so strong, they should have already annihilated the opportunity of what Persson and Savulescu call minor instances. For example, if some human agents are biomedically enhanced and showing a stronger sense of justice, they should be able to avoid the minor cases of cheating or lying and only then to gradually eliminate the more complex and versatile out-group moral challenges.

In other words, there are two alternatives. Either the application of MBE means is ineffective and then one faces concerns about indeterminism – namely, that there are other factors anyway (which questions the whole idea of justifying MBE as such) – or it addresses particular forms of collective moral improvement which cannot be called moral (self-)development at all. If so, however, one cannot argue for "changing one's mind" in moral terms, as suggested by Persson and Savulescu, but rather for "switching one's mind". Consequently, an improvement based upon the latter transformation can successfully address merely what

Persson and Savulescu call grossly immoral acts. Then again, the problem is that MBE is morally permissible only if one evaluates its application from the perspective of act utilitarianism.

Concerning what I would call the 'negative' perspective of the God Machine, one can point out another issue affecting moral autonomy – the way in which trivial immoral acts can (gradually or not so gradually) turn into grossly immoral acts in Persson and Savulescu's sense. Disregarding the complex role of moral self-development points towards answering the question of why the distinction between trivial and grossly immoral acts can be successfully grounded merely within the framework of act utilitarianism.

According to Persson and Savulescu, the God Machine decided that only the acts which would have resulted in imprisonment of a person should be prevented (Ibid.). Certainly, cheating in exams does not, in itself, make anyone a murderer. Nor would one imprison anyone for a murder if they have been cheating in exams. Certainly, again, this is true not only for act utilitarianism as such.

However, giving preference to moral improvement over moral (self-)development may lead to missing the stage when the ones unpunished for trivial immoral acts may deteriorate to dangerous criminals. Taking into account that the analysis of moral (self-)development requires one to conduct an analysis of some complex moral and social factors, one may face the paradox that, whilst striving to build the ideal human nature, the God Machine runs against the human condition as such.

The double-bind insufficiency of moral improvement can be illustrated by the following example. Regarding cheating as a trivial immoral act, its improvement should result in telling the truth. However, in terms of moral self-development, no one can guarantee that the one telling the truth is immune from turning into a murderer, since there are multiple (not only moral) factors which play a role in one's moral self-development.

Persson and Savulescu themselves recognize some of the paradoxes derived from the God Machine when arguing that it is not itself a moral enhancement (Ibid.). As already demonstrated, The God Machine only prevents people from acting immorally, although they can still form immoral intentions (Ibid.). However, preserving the intentions in question does not meet the requirements of the freedom to fall, as shown by Harris. Persson and Savulescu's specification that "the God Machine would not compromise autonomy, that is, even if it did

prevent people from acting immorally", viz. "if people voluntarily chose to be connected" (Ibid.), also triggers some concerns. From the perspective of moral improvement, there are some groups of people, such as that of criminals, who should be involuntarily connected (Ibid.: 410), as in real life.

The first and maybe the most apparent problem is that Persson and Savulescu face difficulties in delineating the group of those who should be involuntarily connected, and in defining the particular transformations behind their improvement. According to Persson and Savulescu, "this number might not be great if there was also extensive moral enhancement in the manner of improving altruism and sense of justice", coupled with "improved impulse control to prevent backsliding from one's considered judgements" (Ibid.). Theoretically speaking, voluntary MBE is also grounded in act utilitarianism, which overestimates the value of the result as a justification of the means and the utilitarian priority of collective well-being.

The implications of the first part of the problem can be exemplified by Persson and Savulescu's statement that by preventing serious crimes, the God Machine would provide other people with great benefits. In "the absence of perfectly effective moral enhancement, the loss of freedom in one domain of our lives – committing evil deeds – would be worth the benefits" (Ibid). Thus, they reach the conclusion that even in cases where "the God Machine does undermine autonomy, the value of human well-being and respect for the most basic rights outweighs the value of autonomy. This is not controversial." (Ibid.).

However, such a statement is apparently controversial. In addition to the concerns that many ideologies disguise the decrease of human autonomy for the sake of improving their well-being, I would argue that the value of human well-being should also clearly include the well-being of the ones whose autonomy is undermined. If we consider the group of a vague number of people – a number which might not be great – the issue is who, how and under what circumstances can do this in a morally permissible manner. More importantly, the question is when we should stop extending the aforementioned group in an effort to incease our own well-being even when it is not so necessary. Specifically, elaborating on this line of thought can bring us back to Persson and Savulescu's own concern about the moral restrictions of in-group behavior.

Conclusion

Persson and Savulescu's double-bind recognition of traditional moral enhancement in the MBE debate triggers more questions than it answers. To be more specific, if traditional moral enhancement is insufficient in methodological terms and MBE is merely hypothetical at this stage, nothing guarantees that even if the latter is somehow successfully developed in the future, the insufficient traditional moral enhancement can fill in the gaps left by the galloping scientific progress in the long term. In other words, if contemporary moral enhancement is insufficient and lacks the necessary improvement which can meet the requirements of the current technological progress, the incompatibility will increase proportionally in time.

Furthermore, accepting Persson and Savulescu's statement that freedom is not the only value (Persson and Savulescu, 2012a) does not make its partial restriction morally justifiable at all. The same applies to all other values, and even vices as well. Consequently, Persson and Savulescu's critical response to Harris's understanding of freedom, as evaluated from the perspectives of determinism and indeterminism, can be considered as convincing in functional terms alone. However, in both cases (although being underlined by some opposing arguments), the need of MBE is put into question.

Specifically, the God Machine raises Harris's concern about the freedom to fall in a new voice. This is due to the fact that the God Machine's prerogative of replacing some immoral intentions with moral ones turns into a process of switching one's mind. The latter lacks the gradual process of changing one's mind, which occurs as part of one's moral self-development when traditional moral enhancement is conducted. This specification leads me to argue that the God Machine's pitfalls stem from prioritizing moral improvement by some hypothetical biomedical means over moral development by some traditional means (such as education).

While Harris's explanation of the freedom to fall points towards the general human potential for moral development as a crucial part of human moral autonomy, in Persson and Savulescu's case the choice of arguing for improvement contradicts the role of moral development for moral self-development, as well as reducing the freedom of moral choice to what is already predetermined to be chosen as moral. The latter clarification is introduced for the purposes of preserving the freedom to fall merely for what Persson and Savulescu call minor or trivial immoral acts. Judging by the aforementioned clarifications, I draw the conclusion that

the God Machine's 'positive' perspective of tackling freedom is grounded into the vision of collective moral improvement rather than moral development.

In turn, the associated 'negative' perspective addresses some moral concerns about the way in which trivial immoral acts can, gradually or not, turn into grossly immoral acts in Persson and Savulescu's sense. I would argue that neglecting the role of moral self-development, which is irreducible to what Persson and Savulescu understand by moral improvement, makes it difficult to predict the turning point, if there is any, when the people unpunished for trivial immoral acts may deteriorate to dangerous criminals.

On a macro-methodological level, the major criticism to the God Machine is that it can work if the Great Moral Project is predetermined as an act utilitarian project where the effectiveness of switching immoral intentions with moral ones is judged by the results, as well as by the number of the affected. On the other hand, it is act utilitarian framework that makes the shift to the voluntarily acceptance of MBE so complicated. In addition to Persson and Savulescu's own hesitation in delineating the small group of people who should be obliged to accept MBE, I also raise the concern that the lack of analysis of moral self-development makes no room for 1) justifying the potential extension of this group (if necessary) in normative terms, 2) exploring the well-being of the ones whose autonomy is undermined, and 3) stopping the extension of the group when increasing our well-being is not so necessary.

In conclusion, we may ask the question whether instead of discussing the God Machine, one may go back to the ancient Greeks' vision of the God from the Machine. While the God Machine displays a model of negative freedom (freedom from grossly immoral acts), the God from the Machine brings back the hope for the model of positive freedom, where humans can be assisted in their most helpless situations. While in the former case the machine represents the climax of moral improvement as a technological development, in the latter case it simply solves the irresolvable in the plot of a tragedy.

However, one may find something common between the two models which was already provisioned by the ancient Greeks. They had realized that such an intervention is undesirable. As Aristotle says in *Poetics*, the proper use of the God from the Machine is only for events external to the drama, which lie beyond the range of human knowledge and require to be reported or foretold (Cf. Aristotle, 1995: 1454b5). I suspect that is also the case with the God Machine as such.

NOTES

[1] Persson and Savulescu develop their conception of MBE in a series of papers in which they begin with advocating compulsory MBE and end with encouraging voluntary MBE (e.g., Persson and Savulescu, 2008; Persson and Savulescu, 2012; Persson and Savulescu, 2017), as well as outlining the necessity of arguing for MBE as filling in the gaps of traditional moral enhancement, and suggesting a symbiosis between the latter and MBE.

[2] For some concerns about Persson and Savulescu's "crude conceptual distinction" between cognitive and moral enhancement in the service of which types of enhancement should be urgently pursued or discouraged, see Carter and Gordon, 2014. However, Persson and Savulescu do not accept the idea of cognitive enhancement as something bad once and for all. Cf. Persson and Savulescu, 2017: 13.

[3] For the debates between the proponents and opponents of MBE, see Persson and Savulescu, 2017; Sparrow, 2014; Diéguez and Véliz, 2019; Agar, 2013; Hauskeller, 2013; Sparrow, 2014.

[4] While Persson and Savulescu emphasize that cognitive enhancement should be unspeeded unless effective means of moral enhancement are found and applied (Persson and Savulescu, 2008), Harris argues that there are good reasons to believe that "moral enhancement must, in large part, consist of cognitive enhancement" (Harris, 2010: 106). Harris also explains that if technological progress increases the potential of individuals to do harm, it is not clear that moral enhancement in Persson and Savulescu's sense will reach the stage of providing safe and effective means for moral transformation (Ibid.: 108).

REFERENCES

Agar, N. (2013). Why Is It Possible to Enhance Moral Status and Why Doing So Is Wrong? – In: *Journal of Medical Ethics*, 39, 67 – 74.

Aristotle (1995). *Poetics.* Translated by Stephan Halliwell. Harvard University Press.

Carter, J. A. and E. C. Gordon (2014). On Cognitive and Moral Enhancement: A Reply to Savulescu and Persson. – In: *Bioethics*, 29 (3), 153 – 161.

Diéguez, A. and C. Véliz (2019). Would Moral Enhancement Limit Freedom? – In: *Topoi*, 38(1), 29 – 36.

Harris, J. (2010). Moral Enhancement and Freedom. – In: *Bioethics*, 25, 102 – 111.

Hauskeller, M. (2013). *Better Humans?* Durham, NC, Aucmen.

Persson, I. and J. Savulescu (2008). The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity. – In: *Journal of Applied Philosophy*, 25(3), 162 – 177.

Persson, I. and J. Savulescu (2012). *Unfit for the Future*. Oxford, Oxford University Press.

Persson, I. and J. Savulescu (2012a). Moral Enhancement, Freedom and the God Machine. – In: *The Monist*, 95(3), 399 – 421.

Persson, I. and J. Savulescu (2017). The Duty to Be Morally Enhanced. – In: *Topoi* (*Dordr*). 38(1), 7 – 14. doi: 10.1007/s11245-017-9475-7. Epub 2017 Apr 12. PMID: 31798198; PMCID: PMC6887531.

Sparrow, R. (2014). Egalitarianism and Moral Bioenhancement. – In: *American Journal of Bioethics*, 14(4), 20 – 28. doi: 10.1080/15265161.2014.889241. PMID: 24730485.