




ARTUR SZUTTA

 University of Gdańsk (Poland)

 0000-0002-1062-4808

 artur.szutta@ug.edu.pl

On an intuition regarding the acquisition of moral virtue

Received: 28.09.2023 / Revised: 20.10.2023 / Accepted: 5.11.2023 / Available: 20.12.2023

Abstract:

This article concerns the issue of the moral enhancement of humans through technology. I propose a thought experiment that allows us to identify a new reason against implementing such enhancement. Achieving virtue through a path that involves one's own effort in making and implementing morally sound decisions deserves greater respect. It also allows us to acknowledge that we are (co)authors of who we become morally. This kind of self-creation seems to be an important part of a meaningful life, and artificial moral enhancement deprives us of it.

Keywords:

moral enhancement, virtue, intuition, meaning of life

How to cite:

Szutta, A. (2023). On an intuition regarding the acquisition of moral virtue [English translation]. *Laboratorium Mentis*, 1(1), 75–85. <https://doi.org/10.52097/lm.8157>

The debate on moral enhancement

According to some authors, the challenges of the modern world, technological development in particular, require us to improve our moral condition. For instance, Julian Savulescu and Ingmar Persson (2012, 2013) argue that contemporary technology may enable immoral people to commit large-scale evil acts, such as destroying a multi-million city or even an entire nation using modern weapons (e.g., dirty atomic bombs, viruses, etc.).¹ The same technological advancement, as they observe, will one day allow us to enhance our moral nature, either at the motivational level (improving our capacity for empathy) or at the cognitive level. Improving our moral abilities through technology, such as genetic manipulation or pharmacology, will not only be an intriguing, perhaps tempting option but also be our moral duty, justified by the need to prevent a very likely great evil.

The idea of moral enhancement is the subject of numerous objections. For example, Allen Buchanan (2009) is concerned that morally enhanced individuals will have a higher moral status and, as a result, greater rights than ordinary human beings. Other authors, such as John Harris (2011) and Michael Hauskeller (2013), argue that as a result of moral enhancement, we will be so determined to do good by our improved emotions, desires, or other mechanisms occurring in our bodies that we will become automatons devoid of free will. Critics of moral improvement through technological interventions also point out the problem of establishing the shared standards by which we should improve people, given the strong disagreement in ethics about basic ethical principles (Schaefer, 2014). Among the objections, there is also the charge of unforeseen consequences of attempts at moral improvement through pharmacology or other techniques (Fabiano, 2018). Critical

¹ The idea of moral enhancement gathers more and more advocates (Specker et al., 2014)

voices have not gone unanswered by supporters of moral enhancement. It can be said that numerous advanced discussions are currently taking place (see Lavazza et al., 2019), in which both sides formulate further arguments, counterarguments, and counter-counterarguments.

In this article, I want to propose a thought experiment that offers a new perspective on the problem of moral enhancement. It allows us to identify a certain intuition that speaks against the implementation of this idea, an intuition that has not been pointed out (at least not directly) in the publications I am aware of.

Assumptions

Before I present the experiment itself, let us make certain assumptions about the technology of moral enhancement. Let us set aside any technical imperfections of current and future methods of human moral enhancement. Let us assume that the outcome of these procedures will be individuals who are indistinguishable in behaviour, consciousness, and motivation from people who would be the perfect embodiment of an Aristotelian virtuous person. Let us also assume that the discussed enhancement of people will be universally accessible; it will have no undesirable effects and will not deprive morally improved individuals of the ability to reflect and act on moral reasons or, more broadly, to make authentic moral decisions. Would we then consider that there are no reasons against realising the idea of moral enhancement?

The experiment

To answer this question, I propose conducting a thought experiment. Let us imagine four parallel worlds indistinguishable except for one detail. Each is inhabited by the same person, John, but in each of these worlds, John has a slightly different history. Accordingly, in each of the possible worlds *a*, *b*, *c*, *d*, there are John_{*a*}, John_{*b*}, John_{*c*}, and John_{*d*}. At a specific time measured simultaneously for all four worlds, let's say at

time t_n , all Johns are morally virtuous in the Aristotelian sense. They have a stable disposition to reliably recognise what is morally right in given circumstances and are motivated in the light of this recognition to effectively do what is good (Aristotle, ca. 300 B.C.E./2014, NE 1116a).

The only thing distinguishing all four men is how they acquired their virtues. John_a became virtuous in the traditional way, initially (when he was still a small child) through imitating other people, then through persistent training in making the right decisions, overcoming his weaknesses and gradually building up the right dispositions within himself, perhaps with more than a few failures (especially at the beginning of his moral development) along the way.

John_b, at a certain time, t_{n-m} , was chosen without his knowledge or consent as the subject of a secret government program, the goal of which was to create and test a moral enhancement project using a certain technology, the details of which I will not specify here to avoid unnecessarily criticisms that would focus on its weaknesses or shortcomings (according to the experiment's assumption, it would be a safe technology).²

The case of John_c differs from that of John_b in that it was not the government that decided to use moral enhancement technology but John_c's parents. They desired a morally virtuous child and were afraid they could not raise him properly using natural methods. John_d, on the other hand, perhaps frustrated by his own poor moral condition, voluntarily enrolled in the moral enhancement program and underwent the procedure with full awareness and consent.

² It would be a natural reaction, while reading about John_b's situation, to raise objections regarding the moral acceptability of the very methods of artificially implementing moral dispositions. I acknowledge that such objections could be justified and provide an important starting point for discussion. However, I want the reader to disregard them here because the goal of this experiment will not be the evaluation of the methods themselves but the very idea of artificial moral enhancement.

For this article, we can assume that at time t_n , all the Johns mentioned above are indistinguishable in terms of their behaviour, motivation, cognition, and emotions. By behavioural indistinguishability, I mean their observable actions from a third-person perspective (including activities of the brain and other relevant body parts) are identical. Motivational indistinguishability means they are guided by the same moral reasons in decision-making. Cognitive and emotional indistinguishability³ implies that the world, all its objects, their characteristics, and relationships between them are cognitively perceived the same way by all the Johns, and they have dispositions to experience the same emotions in the same circumstances.

An intuitive assessment

Despite the indistinguishability described above, the different histories and ways of acquiring virtue seem to lead us to distinct evaluations of the four men. It is intuitive to claim (assuming that most readers will share this intuition) that we should admire John_a more than the other three Johns.⁴ This admiration should encompass not only the current abilities of John_a but also the effort, perseverance, and willful commitment throughout his moral growth process. John_a's current state is, to the greatest extent (compared to the other Johns), a result of his own merits. It seems that accounting for merit in the achievement of virtue should play a role in the moral assessment of the Johns.

³ I neither assume nor reject here that emotions are, by definition, non-cognitive. Such an assumption is not relevant to the purpose of the experiment presented in this paper.

⁴ I believe that, to some extent, John_d deserves a slightly better moral evaluation and, to some extent, our admiration (though to a lesser degree than John_a) because he himself requested moral improvement.

Discussion

Now, does the above intuition have any implications for the moral enhancement debate? It seems that (assuming its validity) this intuition provides *prima facie* reasons against the implementation of artificial moral improvement. Still, the proponents of such enhancement could formulate two arguments against the abovementioned intuition.

First, the mere fact of intuition regarding John_a's greater merit and admiration for him does not necessarily mean that he deserves a higher moral evaluation. This intuition may be a relic of the evolutionary process that is no longer useful as it does not take into account our technological development. At best, such an intuition could relate to the conditions in which people lived before the technology enabling human enhancement emerged. Evolutionary changes (including moral intuitions) are slow, while technological development is rapid. Hence, our mismatch of intuitions with the current civilisational situation, including the possibilities of moral improvement.⁵

Second, even if we ascribe some initial credibility to the intuition mentioned above, it might still have to yield to another intuition concerning the comparison of the likelihood of success in achieving virtue through traditional methods with the probability of attaining it through technology. From observation, we know that very rarely, if ever, people manage to achieve virtue through their own practice based on a series of autonomous decisions. Meanwhile, technology, at least in light of the assumptions adopted in the experiment, provides, if not complete certainty, a significantly higher probability than the traditional method of achieving virtue (or at least something indistinguishable from virtue).

⁵ An advocate of this argument, for example, is Peter Singer (2005). For more about the criticisms of the thesis on the credibility of moral intuitions, see Szutta, 2018.

Thus, when faced with a decision at the societal/political level, whether to introduce moral enhancement through technical means or stick with traditional educational methods, we should weigh both intuitions: on the one hand, the intuition concerning the lower moral status of artificially produced virtue (or something deceptively similar to virtue),⁶ and on the other, the probability of success in implementing virtue by one method over the other. If we could, with a relatively high probability, conclude that, for example, in a society of one million, we would have a million (or even half a million) people *à la* John_{b, c, or d}, wouldn't that be a more desirable state of affairs (which we would be morally obliged to realise) than an equally large society with a few virtuous exceptions? Wouldn't this be a more desirable situation and its realisation our moral duty, even if, as a result of moral enhancement, individuals were deprived of the possibility of (full) merit and (full) admiration for their moral condition?⁷

⁶ I don't want to resolve here whether we can indeed describe the dispositions of Johns_{b-d} as virtues, but it's worth considering the thesis that virtue is four-dimensional, meaning it encompasses not only the condition of the subject at time t_n but the entire process from t_1 (the moment in a person's existence when she may start slowly becoming virtuous) to t_n (the moment of acquiring virtue). Another component/feature of virtue, alongside its four-dimensionality, would be, in light of the experiment presented here, the self-determination of the subject. In other words, virtue would have to originate from within, not from outside the subject. Similar intuitions regarding the acquisition of virtue (or moral improvement in general) are held by Jason Erbel (2018), Ruben Herce (2019), and Tartaglia (2020, ch. 5).

⁷ The above argument can be presented in another form. We can imagine John_e, who consciously gave up on moral improvement in favour of the traditional form of moral development but still failed to achieve virtue or even come close to it. Wouldn't we consider that John_e deserves a negative moral evaluation for wasting time on unsuccessful attempts when he had an effective alternative at his disposal? In cases where we are certain that a person has no or very little chance of becoming morally virtuous or at least morally good at some minimal level, one could argue that we have a duty to morally improve them through technology. Otherwise, that person's life would be somewhat wasted. It's also worth adding that in the comparison described above, it's not only about effectiveness in terms

We can put forth the following arguments in response to the above criticisms. First, perhaps not necessarily our moral intuitions, especially concerning the way virtue is acquired, are unworthy of at least initial trust. Since I have written elsewhere about the credibility of moral intuitions in general (Szutta, 2018, Chapter 7) and there is already substantial literature on this topic,⁸ I will limit myself to the argument for the credibility of a specific intuition regarding the significance of the way virtue is acquired. It does not necessarily have the character of an outdated mechanism (or a mental habit) that only applied to times when the time-consuming development of character through autonomous participation in appropriate practices was the only effective method of achieving virtue. Perhaps this intuition also relates to our understanding of a meaningful life.

Maybe an essential element of a meaningful life is, at least to some significant extent, to be the author of oneself, to create oneself through autonomous decisions. Let me propose another short experiment to support this interpretation of the defended intuition. Imagine that at the moment of your birth, your parents requested the implantation of a small chip in your head, which, by appropriately stimulating your brain, would motivate you to follow a specific life pattern planned by your parents. Let's assume that this plan involves having certain interests at school, choosing the right studies, selecting the right life partner, etc. Assume also that by making the "right" choices, you would appeal to the objective reasons you've recognised, but they would serve more as justifications for choices predetermined by your parents rather than

of achieving the desired outcome but also about efficiency measured in terms of time. When comparing traditional virtue acquisition with the method of artificial moral enhancement, we should also take into account that the traditional method is more time-consuming, and along the way to achieving virtue, there may be many moral failures, which, assuming we have an effective method of artificial enhancement, could be avoided.

⁸ An interesting example of the latest defences of moral intuitions is Bengson et al., 2020.

as causes or reasons determining your choices. If you were to discover such a chip and its function, wouldn't your life lose at least some important part of its value to you?

Thus, doesn't the admiration for human achievements simultaneously imply a kind of affirmation that human life contains an essential, meaningful element? And does this element not consist in the fact that what we achieve in life, we do so, at least to some extent, also through our own efforts and decisions?

In conclusion, let me propose an additional, auxiliary argument aimed only at those who accept the existence of a personal God, who created humans and desires their moral development. If God is good, omnipotent, and omniscient, and if it were good for us, He would have made us morally perfect from the very start, without the need for the uncertain and painful birth of our moral character, without regressions, and the risk of failure. If, despite His goodness, omniscience, and omnipotence, He did not create us as perfect, then perhaps He deemed there is a better alternative: that we have a part in creating ourselves, that we are co-authors of ourselves. Accepting God's existence, we have another reason to support the belief that the way we become morally good people matters because it matters whether we can at least partially call the outcome of our lives, who we become, our self-determination.

Conclusion

The thought experiment and the discussion above point to a rather neglected reason against introducing moral improvement through artificial intervention in the human body. By implementing the disposition for effective moral good through this means, we deprive individuals of the opportunity for autonomous self-determination. However, we should not consider this reason finally determining the impermissibility of artificial moral enhancement. It has a *prima facie* character. The final decision on the issue discussed here requires considerations of many other reasons, such as whether the price of giving up (or depriving) a

chance for self-creation is worth it to achieve a society in which the vast majority if not all, act morally well. This text should be considered a proposal of a reason to be included in the broader discussion of the permissibility and value of moral enhancement.⁹

Bibliography

Aristotle (2014). *Nicomachean Ethics*, transl. R. Crisp. Cambridge University Press. Original work published ca. 300 B.C.E.

Bengson, J., Cuneo, T., Shafer-Landau, R. (2020). Trusting moral intuitions. *Nous*, 54(4), 956–984. <https://doi.org/10.1111/nous.12291>

Buchanan, A. (2009). Moral status and human enhancement. *Philosophy & Public Affairs*, 37(4), 346–381. <https://www.jstor.org/stable/40468461>

Eberl, J.T. (2018). Can prudence be enhanced? *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, 43(5), 506–526. <https://doi.org/10.1093/jmp/jhy021>

Fabiano, J. (2018). *Probing the Risks of Moral Enhancement* [PhD thesis]. University of Oxford.

Harris, J. (2011). Moral enhancement and freedom. *Bioethics*, 25(2), 102–111. <https://doi.org/10.1111/j.1467-8519.2010.01854.x>

Hauskeller, M. (2013). The “Little Alex” Problem. *The Philosophers’ Magazine*, (62), 74–78. <https://doi.org/10.5840/tpm20136299>

Herce, R. (2019). Is human enhancement possible if it comes from the outside? *Scientia et Fides*, 7(2), 165–170.

Lavazza, A., Reichlin, M. (2019). Introduction: moral enhancement. *Topoi*, 38(1), 1–5. <https://doi.org/10.1007/s11245-019-09638-5>

Persson, I., Savulescu, J. (2013). Getting moral enhancement right: the desirability of moral bioenhancement. *Bioethics*, 27(3), 124–131. <https://doi.org/10.1111/j.1467-8519.2011.01907.x>

⁹ For critical comments made to the earlier versions of this paper, I would like to express my gratitude to Krzesimir Cholewa, Michał Dominów, Paweł Homel, Jakub Nowicki, Stanisław Kamiński, Paweł Sikora, Natasza Szutta, Błażej Szymichowski, James Tartagila, and Emilia Wrońska.

Savulescu, J., Persson, I. (2012). Moral enhancement, freedom and the god machine. *The Monist*, 95(3), 399–421. <https://doi.org/10.5840/monist201295321>

Schaefer, G., Schaefer, G.O. (2014). *Moral enhancement and moral disagreement* [PhD thesis]. Oxford University, UK.

Specker, J., Focquaert, F., Raus, K., Sterckx, S., Schermer, M. (2014). The ethical desirability of moral bioenhancement: a review of reasons. *BMC Medical Ethics*, 15(1), 1–17. <https://doi.org/10.1186/1472-6939-15-67>

Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics*, 9(3/4), 331–352. <http://www.jstor.org/stable/25115831>

Szutta, A. (2018). *Intuicje moralne. O poznaniu dobra i zła* [Moral intuitions. About the cognition of the good and evil]. Lublin: Wydawnictwo Academicum. <https://doi.org/10.52097/acapress.9788362475612>

Tartaglia, J. (2020). *Philosophy in a Technological World: Gods and Titans*. Bloomsbury Publishing.