

Commentary: The Implementation Ethics of Moral Enhancement

NICHOLAS AGAR

Vojin Rakic makes a valuable contribution to the debate about moral enhancement.¹ Before I comment on the specifics of his proposal, I will make some general observations about how philosophers have engaged with the possibility of making humans morally better.

There has been a shift in the focus of the debate of moral enhancement since its inception with Tom Douglas's 2008 paper "Moral Enhancement."² Here Douglas presents moral enhancement as a thought experiment. He is explicit about the purpose of this thought experiment. "My aim is to present this kind of moral enhancement as a counterexample to the view that biomedical enhancement is always morally impermissible."³ Douglas seeks to refute what he calls the bioconservative thesis: "Even if it were technically possible and legally permissible for people to engage in biomedical enhancement, it would not be morally permissible for them to do so."⁴ He offers moral enhancement as a clearer counterexample to this thesis than the cases of cognitive enhancement or the enhancement of physical capacities that have been the focus of most of the debate about human enhancement.

Since the publication of Douglas's article, there has been a shift from moral enhancement as a thought experiment to moral enhancement as practical advice. No advice to actually perform moral enhancement follows from Douglas's rebuttal of the bioconservative thesis. The aim of much of the later writing is practical—moral enhancement is offered as a response to challenges we face. According to Ingmar Persson and Julian Savulescu, we must enhance ourselves morally in order to avoid "ultimate harm"—an event that would make "worthwhile life forever impossible on this planet."⁵ They do not offer moral enhancement to prevent ultimate harm as part of a thought experiment designed to counter global rejection of human enhancement. It goes without saying that Persson and Savulescu reject the bioconservative thesis, but they are more concerned about the actions we take than about the philosophical theses we endorse.

The assessment of moral enhancement as a thought experiment differs from assessing moral enhancement as practical advice. Suppose you are an opponent of enhancement who believes that moral enhancements are not counterexamples to the bioconservative thesis. In your response to Douglas, you describe some human modifications that both satisfy the criteria for moral enhancement and seem morally impermissible. Douglas wins if he successfully describes a single case of moral enhancement that is clearly morally permissible. He has his counterexample to the bioconservative thesis even if the vast majority of moral enhancements are patently immoral. The goal of refuting the bioconservative thesis is achieved even if we insist that no moral enhancement that anyone is likely to attempt will be permissible.

The assessment of moral enhancement as practical advice differs from deciding whether moral enhancement could be a counterexample to the bioconservative thesis.

You do not demonstrate that moral enhancement is an effective response to the threat of human extinction by presenting a thought experiment in which humans morally enhance and thereby avoid extinction. We need additional information if we are to accept moral enhancement as useful practical advice. For example, it is possible to describe a thought experiment in which detonating nuclear warheads on Antarctica solves the problem of anthropogenic climate change. We can coherently conjecture that large quantities of fallout ejected into the upper atmosphere will end up removing carbon. Perhaps this imaginative exercise successfully demonstrates that there are no distinctively philosophical objections to this proposal to reverse anthropogenic climate change by nuking Antarctica. But the absence of a convincing in principle argument against this measure does not mean that nuking Antarctica should be offered as a practical response to anthropogenic climate change. By analogous reasoning, philosophers can imagine that people with strengthened empathy will take steps to reduce the likelihood of human extinction. But defenders of moral enhancement should not present a thought experiment in which strengthened empathy produces better moral outcomes and consider their work complete. For moral enhancement to be useful practical advice, we need evidence for this connection. We should not cherry-pick the empirical literature for cases most favorable to moral enhancement. Suppose an opponent offers reasons for thinking that this connection will, in fact, not obtain. It is not enough to reply by insisting on the superficial coherence of the thought experiment.

Allocating Intellectual Labor to the Consideration of Moral Enhancement as Practical Advice

I have suggested that the philosophical purposes of those who present moral enhancement as a thought experiment differ from those who present moral enhancement as practical advice. There are connections between the two goals. Promoting moral enhancement as a practical response to ultimate harm will require responses to the bioconservative thesis. If all human enhancements are morally impermissible, then it follows that it is wrong to seek to prevent ultimate harm by means of moral enhancement—we should not offer moral enhancement as practical advice.

We can nevertheless make second-order observations about the allocation of intellectual labor to the debate about moral enhancement. Suppose you are an advocate of moral enhancement whose purpose is practical. You will need to respond to bioconservatives who claim that any human enhancement is immoral. But your work is not done with the rebuttal of the bioconservative thesis. If the goals of morally enhancing humans to avoid extinction or making a better society are viable, then we will need to attend to a variety of what I will call *implementation problems*. These address obstacles to the realization of moral enhancement. Which technologies will be used? Does what we know about human moral psychology make a given proposal difficult or impossible? We will need to determine which are the appropriate means of administering moral enhancement. Are there reasons to prefer pharmacological interventions to interventions that modify subjects' genetic material?

It might seem that implementation problems should be a focus only once we have established the in principle permissibility of moral enhancement. But implementation problems are relevant to the overall assessment of moral enhancement.

Imagine we considered all the conceivable ways of implementing moral enhancement. These included a wide variety of pharmacological, genetic, and cybernetic means of making humans morally better. Suppose that we found, in each case, that there were either insuperable technological obstacles or that moral enhancement could be implemented only in a way that we deemed morally unacceptable. Such a finding should lead us to give up on moral enhancement as a practical plan. It would constitute grounds for refusing to move from moral enhancement as a thought experiment to moral enhancement as practical advice. Note that this is a discovery that we could make only once we had put serious thought into the challenge of implementing moral enhancement. It suggests that philosophers of enhancement should reject a linear order of business in which we attempt to establish the in principle morality of moral enhancement before we address the specifics of its implementation.

I offer a conjecture about the disproportionate claim on philosophical attention of moral enhancement as a thought experiment. Philosophers are subject to a bias toward stereotypically philosophical problems, problems that straightforwardly satisfy the criteria for problems we are accustomed to dealing with. We are, as an academic specialty, less interested in problems that are a worse fit for the template of problems that philosophers typically confront. This bias should be a concern for those whose interest in moral enhancement is practical. It leads to an overemphasis on stereotypically philosophical problems and a comparative neglect of other problems that will need to be addressed if the goal to improve society by moral enhancement is to be realized.

One example of a stereotypically philosophical debate about moral enhancement is the debate about the God machine, an imagined brain implant that achieves morally better behavior by intervening to suppress immoral motivations.⁶ Philosophers have no difficulty in recognizing the debate generated by this thought experiment as philosophical. If someone avoids morally wrong actions due to the intervention of the machine, then are they truly morally good? The born-again bioconservative philosopher John Harris complains about the withdrawal of the “freedom to fall.”⁷ Advocates of moral enhancement challenge this conclusion. My concern here is not about the specifics of that debate. Rather, I express a second-order concern. Those who seek to implement moral enhancement should lament the disproportionate claim on philosophical attention to this stereotypically philosophical problem.

Even opponents of moral enhancement should be concerned about the current allocation of philosophical labor. They may be confident about their philosophical rejection of moral enhancement. Suppose, however, that at some time in the future there is a resolution to proceed with moral enhancement, their philosophical objections notwithstanding. Opponents will be left with few objections to the ways in which advocates choose to implement moral enhancement. Both advocates and opponents of moral enhancement should regret the disproportionate claim on philosophical labor of stereotypically philosophical approaches.

In my paper “Moral Bioenhancement is Dangerous,” I focus on a problem that arises in the implementation of moral enhancement. I observe that current proposals to morally improve us by strengthening inputs into moral decision-making such as empathy are likely to lead to moral decline rather than moral improvement.⁸ Consider the observation made by Persson and Savulescu (2012) and elsewhere that strengthened empathy may lead us to demonstrate a stronger preference for

in-group rather than out-group members.⁹ Persson and Savulescu table this as a problem requiring attention before we proceed with any plan to morally improve by strengthening empathy. This underestimates the gravity of the problem. I argue that it is an especially difficult challenge for any attempt to morally enhance by boosting selected inputs into moral judgment. Our considered moral judgments emerge from a dynamic interplay of reasons and emotions. For example, when we assess the ethics of the death penalty, we find ourselves negotiating a variety of different moral concerns and interests—the perceived wrongness of deliberately taking human life, the appropriateness of expressing outrage against those who perpetrate the worst crimes, conjectures about the deterrent effect of executing convicted criminals, and so on. What we present as our considered moral judgments tend to arise from a judicious balancing of all of these often contrary, frequently competing moral motivations. Advocates of reflective equilibrium offer this as the way to justify moral claims. But even if we reject this approach to justifying moral claims, we can at least accept it as an accurate description of the psychological process that typically guides us toward our considered moral judgments. When we selectively strengthen one of these inputs we are likely to end up with judgments that conflict with our current considered assessments. For example, I advanced the conjecture that strengthening the motivational power of reason may lead us to do despicable things to our kin when directed to do so by utilitarian reasoning that we find difficult to refute.¹⁰

These are concerns about the implementation of moral enhancement. When defenders of moral enhancement respond to these concerns, they must treat them differently from in principle objections. It is certainly possible to imagine that boosting empathy will be magically matched by a corresponding strengthening of the power of reason to reject immoral privileging of members of our in-group. It is possible to imagine that strengthening the motivational power of utilitarianism among moral actors who find it persuasive will be matched by increased concern for the feelings of those harmed in the process of bringing about good outcomes. However, if we are interested in implementing moral enhancement, we must consider whether such alterations are likely to occur. We should not translate a successful response to overly confident rejections of all human enhancements into an actual plan to morally enhance humans, unless we can offer assurance that we will find an appropriate balancing of moral inputs. We can imagine that selective strengthening of inputs into moral evaluation will produce systematic moral improvement, but we should acknowledge they are very likely to make us morally worse.

How Gene Editing Addresses a Problem in Implementing Moral Enhancement

Vojin Rakic's focus on the possibilities of gene editing locates his contribution in the implementation ethics of moral enhancement. Recent advances in gene editing add little to moral enhancement as a thought experiment. Philosophers have always been free to formulate thought experiments in which the genomes of human embryos are modified in whichever way they nominate. In a philosophical thought experiment, the introduction of Albert Einstein genes can produce humans with his intellect. This can work as a thought experiment about genetic enhancement even if geneticists tell us that the suggestion that we might produce additional Einsteins by finding and inserting Einstein genes into human embryos

is absurd. As I suggested above, the consideration of such thought experiments falls short of justifying the suggestion that such genes be located and introduced into human embryos. Rakic's focus promises to right an imbalance in the allocation of philosophical labor. He makes an effective case that gene editing has the potential to address a conflict between moral enhancement and free will.

Rakic's interest in the ethics of implementing moral enhancement is evident in his suggestion that we use the name "police machine" as a substitute for "God machine." Both terms make clear that something has been introduced into subjects' brains that prevents them from performing certain categories of action. In both cases there are legitimate concerns about how the machine interacts with human moral psychologies. But the term "police machine" raises concerns about the motives behind that control that are of great importance to the implementation of moral enhancement. The citizens of societies with low levels of police corruption may not object to measures to limit immoral behavior. But citizens of societies in which there is more general concern about the actions of police have legitimate grounds for unease about the source of this moral authority.

Hypothesizing that moral enhancement is enacted by a morally perfect being makes sense for those who present moral enhancement as a thought experiment. But those with an interest in how moral enhancement is likely to be implemented have questions better expressed by Rakic's choice of term. The label "police machine" directs attention at the agencies that will bring moral enhancement about. It does not morally prejudge the issue. Those of us fortunate enough to live in stable liberal democracies with low levels of police corruption welcome most of the interventions of the police and so might be more open to the introduction of police machines than would citizens of Somalia, the nation listed by the organization Transparency International as having the world's highest perceived level of public sector corruption in 2017.¹¹ Somalis should be more concerned about their leaders' plans to introduce moral enhancers into the drinking water than should the citizens of New Zealand, the country listed in 2017 as having the lowest perceived level of corruption.

Rakic presents a choice about how to implement moral enhancement. One option would be to make it compulsory. This seems to be an intolerable intrusion in subjects' freedom. The other option is to make it voluntary. This raises the problem that those most in need of moral enhancement are the least likely to avail themselves of it. Those who seek it are less likely to need it. Rakic's solution to this particular problem is to use gene editing to introduce moral improvement into the genomes of future people. This would be involuntary moral enhancement, not chosen by its recipients. Because the enhancement occurs before there is a person, there are no existing interests for it to conflict with. In this scenario, people come into existence morally enhanced and presumably form desires and projects that are responsive to these enhancements.

Rakic is aware of a challenge to his proposal. We may not have to worry about the objections of embryos that we choose to morally enhance, but we do have to worry about the procreative interests of their prospective parents.¹² Might some prospective parents reject this attempt to morally direct their future children by burdening them with moral motivations that will stand in the way of the achievement of their own self-interested goals? These parents might hope for careers in business for their children. They may view these as handicapped by too much willingness to promote the interests of others. We might end up with an outcome

that is worse than the situation prior to moral enhancement. Rakić's suggestion that we offer incentives for people to morally enhance could exacerbate existing social divisions. The poor will be motivated by incentives to morally enhance their future children. The rich will be comparatively insensitive to these incentives and hence less likely to morally enhance their future children. The result could be a class of morally unenhanced haves increasingly free to exploit the morally enhanced have-nots.

If Rakić's proposal is presented as a contribution to implementation ethics then he has a response to this objection. When we consider introducing a new social policy, we need to do more than consider the possibility of abuse. We must consider the likely magnitude of this abuse. There are reasons to expect that the pattern of uptake of moral enhancement described in the previous paragraph will not occur. People who, perhaps mistakenly, find themselves in no need of moral enhancement are more likely to recognize a need for moral enhancement for their future children.

Many people are subject to a moral blind spot regarding their own performances. People whose behavior we recognize as immoral frequently rush to defend the morality of what they do. This is evident in the response of men whose behavior has been highlighted by the "Me Too" movement against sexual harassment and assault. Few of them come forward with an awareness of their moral misdeeds. They sincerely claim that reported nonconsensual touching was in fact entirely morally appropriate. We can imagine that many of those most in need of moral enhancement would reject it because they are subject to a blind spot about their own moral performances. They judge their behavior to be morally acceptable and confidently reject criticism.

I suspect that this blind spot is less likely to extend to children. Even morally corrupt people do not object to the suggestion that their children should receive moral instruction at school. Although they assess their own conduct and find excuses or justifications for apparent failings, they accept that their children stand in need of moral improvement. Wealth inequality means that the children of the rich tend to receive different and superior educations from those available to the children of the poor. We have yet to see this separation in terms of the moral instruction provided to children. The rich are not yet opting to send their children to schools where they can be given moral instruction appropriate for Nietzschean overmen and overwomen. Consider the family dynamic depicted in the HBO series *The Sopranos*. A gangster father, Tony Soprano, seems concerned about the moral conduct of his offspring even as he plans the execution of his current business confederate, without the merest concern about the moral status of this act.

Rakić's conjecture demands future investigation. Suppose that certain sections of society accept involuntary moral enhancement for their future children while others reject intrusion into their procreative freedom and have the resources to forgo incentives. This could lead to a moral worsening. But, as with any interesting implementation proposal, we cannot be confident of this outcome before we have considered it in the light of what is known about human moral motivations and how gene editing might affect these.

Notes

1. Rakić V. Genome editing for involuntary moral enhancement. In this issue of *Cambridge Quarterly of Healthcare Ethics*.
2. Douglas T. Moral enhancement. *Journal of Applied Philosophy* 2008;25:228–45.

Commentary: *The Implementation Ethics of Moral Enhancement*

3. See note 2, Douglas 2008, at 228.
4. See note 2, Douglas 2008, at 228.
5. Persson I, Savulescu J. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press; 2012, at 46.
6. Persson I, Savulescu J. Moral enhancement, freedom and the God Machine. *The Monist* 2012; 95(3):399–21.
7. Harris J. Moral enhancement and freedom. *Bioethics* 2011;25:102–11.
8. Agar N. Moral bioenhancement is dangerous. *Journal of Medical Ethics* 2015;41:343–5.
9. See note 5, Persson 2012.
10. Agar N. Moral bioenhancement and the utilitarian catastrophe. *Cambridge Quarterly of Healthcare Ethics* 2015;24:37–47.
11. See *Corruption Perceptions Index 2017*; available at https://www.transparency.org/news/feature/corruption_perceptions_index_2017 (last accessed 27 Sept 2018).
12. See Agar N. *Liberal Eugenics: In Defence of Human Enhancement*. Oxford: Blackwell; 2004 for my interpretation of the scope and significance of procreative freedom in respect of choices about our children's genetic makeups.