

# The Trouble With Moral Enhancement

INMACULADA DE MELO-MARTÍN

## Abstract

Proponents of moral enhancement believe that we should pursue and apply biotechnological means to morally enhance human beings, as failing to do so is likely to lead to humanity's demise. Unsurprisingly, these proposals have generated a substantial amount of debate about the moral permissibility of using such interventions. Here I put aside concerns about the permissibility of moral enhancement and focus on the conceptual and evidentiary grounds for the moral enhancement project. I argue that such grounds are quite precarious.

## 1. Introduction

At a time when racist, sexist, and xenophobic attitudes are again becoming normalised, falsehoods are publicly presented as reality, the US – one of the most polluting countries in the world – is rolling back on initiatives aimed at mitigating climate change and promoting the use of fossil fuels, and wars all over the world are prematurely ending the lives of thousands of people and displacing thousands more, the idea of finding some quick technological fix to morally enhance human beings appears more appealing than ever. Alas, the moral enhancement project,<sup>1</sup> the latest fad in the bioethics literature, betrays, like most technological fixes, a problematic understanding of the nature of the problems it ostensibly attempts to solve. Part of the overall process to remake humanity by various technological means, the moral enhancement project calls for the use of biomedical technologies, from drugs to genetic interventions, to make human beings more moral.<sup>2</sup> Its proponents warn that a failure to pursue this project will likely bring about the annihilation of the planet

<sup>1</sup> Proponents and critics of using various biomedical technologies to enhance humanity's moral sense use "enhancement" and "bioenhancement" interchangeably. I use here simply the term "enhancement".

<sup>2</sup> See for instance, T. Douglas, 'Moral Enhancement', *J Appl Philos.* 25:3 (2008), 228–45; I. Persson and J. Savulescu, 'The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity', *Journal of Applied Philosophy*, 25:3 (2008), 162–67; I. Persson and J. Savulescu, *Unfit for the Future: The Need for Moral Enhancement* (Oxford: Oxford University Press, 2012); D. DeGrazia,

and of the human species with it.<sup>3</sup> Noticeably, the fact that human beings find themselves in this precarious situation is, according to proponents of moral enhancement, the result not only of our less than reliable moral sense, but also of our increased technological power. This makes all the more puzzling the insistence that yet more technology can save us from our folly.

This perplexing fascination with technological solutions, particularly biomedical interventions, as a way to solve complex social, political, and ethical problems is not the only concern that the moral enhancement project raises. Indeed, proposals for human enhancement in general and moral enhancement in particular, are actually quite confusing. Its most staunch proponents have upheld inconsistent claims. For example, they defend both the necessity of moral enhancement as the only way to avoid our planet's demise,<sup>4</sup> and its simple advisability alongside other, more traditional, means of improving people's moral sense and addressing complex social problems.<sup>5</sup> They similarly have defended the claim that biomedical interventions able to enhance people's moral capacities are within reach,<sup>6</sup> and that their development is in its infancy.<sup>7</sup> They repeatedly remind us that evolutionary processes have led to the development of a human moral sense that is less than stellar,<sup>8</sup> but nonetheless trust the human ability to develop the right kind of moral bioenhancements.<sup>9</sup> The fact that proponents often misunderstand and misinterpret the criticisms presented against the

---

'Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour', *J Med Ethics* 40:6 (2014), 361–68.

<sup>3</sup> For the most forceful defense of this claim see Persson and Savulescu, 'The Perils of Cognitive Enhancement'.

<sup>4</sup> Persson and Savulescu, 'The Perils of Cognitive Enhancement'; I. Persson and J. Savulescu, 'The Turn for Ultimate Harm: A Reply to Fenton', *J Med Ethics* 37:7 (2011), 441–444; I. Persson and J. Savulescu, 'Getting Moral Enhancement Right: The Desirability of Moral Bioenhancement', *Bioethics* 27:3 (2013), 124–31.

<sup>5</sup> Persson and Savulescu, *Unfit for the Future*; I. Persson and J. Savulescu, 'The Art of Misunderstanding Moral Bioenhancement', *Cambridge Quarterly of Healthcare Ethics* 24:1 (2015), 48–57.

<sup>6</sup> Persson and Savulescu, 'The Perils of Cognitive Enhancement'.

<sup>7</sup> Persson and Savulescu, *Unfit for the Future*.

<sup>8</sup> Persson and Savulescu, 'The Perils of Cognitive Enhancement'; J. Savulescu and I. Persson, 'Moral Enhancement, Freedom and the God Machine', *Monist* 95:3 (2012), 399–421.

<sup>9</sup> Persson and Savulescu, *Unfit for the Future*; Persson and Savulescu, 'The Art of Misunderstanding Moral Bioenhancement'.

## The Trouble With Moral Enhancement

moral enhancement project also contributes to this confusion.<sup>10</sup> But, as I show below, moral enhancement proposals are also confused: plagued by conceptual problems, beset by category mistakes, and packed with conclusions that rarely follow from the arguments presented. Equally troubling, such proposals are scientifically uninformed and often misleading. As I will discuss later, the misuse of the so-called “scientific evidence” purportedly supporting various claims about moral enhancement is staggering. Moreover, moral enhancement proposals, particularly when situated in the context of human enhancement in general are alarming. Although presented under the guise of liberalism and progressivism, there is an implicit – and sometimes not so implicit – embracing of the status quo and of a simplistic notion of choice that serves to perpetuate current injustices. The same authors who support the need for moral enhancement are all too happy to defend a moral obligation to enhance future offspring with physical and character traits suspiciously consistent with the ideologies of sexists, racists, and ableists.<sup>11</sup> Similarly, they regularly present technological advances as value-neutral, and thus ignore the role that social, political, and ethical values have in the development of biomedical technologies as well as their effect on what we can choose, what we might think is morally permissible, and what we value.<sup>12</sup> Moreover, they uncritically sanction technological interventions even if doing so contributes to and furthers social injustices.<sup>13</sup>

In spite of all of these problems, the enhancement project has commanded an astounding amount of attention.<sup>14</sup> This is a tribute to its

<sup>10</sup> For arguments about such misinterpretation see Michael Hauskeller, ‘The Art of Misunderstanding Critics: The Case of Ingmar Persson and Julian Savulescu’s Defense of Moral Bioenhancement’, *Cambridge Quarterly of Healthcare Ethics* 25:1 (2016), 152–60.

<sup>11</sup> J. Savulescu and G. Kahane, ‘The Moral Obligation to Create Children with the Best Chance of the Best Life’, *Bioethics* 23:5 (2009), 274–90; J. Savulescu, ‘New Breeds of Humans: The Moral Obligation to Enhance’, *Reproductive Biomedicine Online* 10:1 (2005), 36–39.

<sup>12</sup> I. de Melo-Martín, *Rethinking Reprogenetics: Enhancing Ethical Analyses of Reproductive Technologies* (New York: Oxford University Press, 2017).

<sup>13</sup> J. Savulescu and G. Kahane, ‘The Moral Obligation to Create Children with the Best Chance of the Best Life’; J. Savulescu, ‘New Breeds of Humans’. For others also calling attention to this disturbing tendency see, for instance, R. Sparrow, ‘Egalitarianism and Moral Bioenhancement,’ *Am J Bioeth.* 14:4 (2014), 20–28.

<sup>14</sup> The attention is not just the result of the significant amount of criticism that it has received, but also due to a surprising amount of support.

proponents' ability to frame their proposals in ways that makes them appear, on the one hand, business as usual – we have been morally enhancing ourselves all along – and, on the other hand, absolutely indispensable – the only way to save humanity from the destructive technologies we have developed, a cognitive capacity too great for our own good, and a moral brain best suited to the Paleolithic jungle. If to this we add a context where many people embrace scientism and welcome the use of technological fixes to address complex social problems, we have all one needs to make discussions on moral enhancement appear consequential and enhancement proposals to look as if they were desirable.

Is it possible to improve current proposals for moral enhancement, to make them more meaningful, informed, and intelligible? That depends. It is, of course, conceivable that someone could offer new proposals that lack all the problems the present ones have. I confess, however, that I am sceptical of such a possibility. It seems to me that any talk about “morally enhancing human beings” has to rest on some mistake or another: conceptual errors, ambiguities and vagueness, problematic background assumptions, and the like. No doubt, a better understanding of moral philosophy, philosophy of science, and philosophy of technology, as well as attention to feminist perspectives in all of these areas would serve proponents of moral enhancement well, not just regarding moral enhancement, but human enhancement in general. In any case, given the very confusing and confused state of moral enhancement proposals, it seems to me that the best way – perhaps the only way – to make such proposals more meaningful and intelligible is to point out the many problems they have. Space prevents me from attending to all such problems – which will require a book. Thus, in what follows I discuss just some of those problems – and for the most part, I will be able to only scratch their surface.

## **2. What is Wrong With Moral Enhancement Proposals?**

Although a significant amount of the debate over moral enhancement has focussed on whether biomedical interventions to enhance humanity's moral capacities are or are not permissible, such a question presupposes that talk of moral enhancement is meaningful. Alas, it is not clear that such is the case.<sup>15</sup> So, how do proponents

<sup>15</sup> For arguments about the meaningless of moral enhancement proposals see I. de Melo-Martín and A. Salles, ‘Moral Bioenhancement: Much Ado About Nothing?’, *Bioethics* 29:4 (2015), 223–32.

## The Trouble With Moral Enhancement

understand “moral enhancement”? They do not all agree on what exactly constitutes moral enhancement and different authors have given diverse definitions of it. For example, Douglas takes moral enhancements to refer to interventions that will expectably leave an individual with more moral motives or behaviour than the person would otherwise have.<sup>16</sup> Persson and Savulescu understand moral enhancements as improvements, by genetic or other biological means, to moral dispositions, such as altruism and a sense of justice, in ways that make it more likely that one will arrive at the correct judgement regarding what is right and more likely to act on that judgement.<sup>17</sup> DeGrazia sees them as interventions that aim at either boosting or selecting an existing moral capacity to act in a particular way, or at creating such a capacity when nonexistent.

None of these definitions is particularly clear, but in any case the *moral* part of the moral enhancement equation is equated with motives, behaviours, or some type of primitive disposition such as altruism and a sense of justice. But as many have pointed out – indeed, as even proponents have recognised<sup>18</sup> – none of these things by themselves would uncontroversially result in a moral person or a moral action. People can do the right thing for the wrong reasons, appropriate motivations can be disregarded, and primitive dispositions are hardly sufficient – and perhaps even necessary – to produce people who are unproblematically more *moral*.<sup>19</sup> Of course, if the biointerventions in question worked – a very big if indeed – we might want to say that the person has been motivationally, behaviourally, or dispositionally enhanced. But that is a far cry from creating a more *moral* person. Indeed, Douglas himself, for instance, explicitly states ‘I will not claim that the morally enhanced person is more moral, has a more

<sup>16</sup> Douglas, ‘Moral Enhancement’.

<sup>17</sup> Persson and Savulescu, ‘The Perils of Cognitive Enhancement’; Persson and Savulescu, ‘Getting Moral Enhancement Right: The Desirability of Moral Bioenhancement’; Savulescu and Persson, ‘Moral Enhancement, Freedom and the God Machine’.

<sup>18</sup> Persson and Savulescu, *Unfit for the Future*.

<sup>19</sup> See, for instance, H. Maibom, ‘Feeling for Others: Empathy, Sympathy, and Morality’, *Inquiry* 52:5 (2009), 483–499; S. Nichols, *Sentimental Rules: On the Natural Foundations of Moral Judgment* (Oxford: Oxford University Press, 2004); S. Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability* (Cambridge, MA: Harvard University Press, 2006); B. Herman, *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press, 1993); M. Weber, ‘The Motive of Duty and the Nature of Emotions: Kantian Reflections on Moral Worth’, *Can J Philos.* 33:2 (2003), 183–202.

moral character, or will necessarily act more morally than her earlier, unenhanced self'. This makes it all the more puzzling that nonetheless he insists in talking about *moral* enhancement. But if enhancing dispositions, behaviour, or motivations cannot be said – not, at least, without various problematic assumptions – to make people more moral, then the so-called *moral* enhancement proposals do nothing of the sort.

But can the interventions proposed actually *enhance* relevant behaviour, dispositions, or motivations? That also seems questionable. Consider for instance moral dispositions of interest to proponents of moral enhancement, such as altruism or empathy, which are conceptualised in fragmentary and disconnected ways. They take these dispositions to be the result of evolutionary pressures, and interventions, such as oxytocin, are aimed at altering the dispositions in question independently of the context in which they will be activated. But, as mentioned earlier, dispositions such as altruism or empathy are not always morally appropriate – imagine, for instance, someone feeling empathy towards a distraught thief who has become incapable of stealing as much as he wants. Thus, even if the interventions were to work, it is not clear that they would actually be enhancing; that is, it is not clear that they would reliably produce a dispositional change that would be judged to be morally better. This is so because the various traits of interest can be used for good or evil. Indeed, Savulescu and Persson explicitly call attention to the problems with altruism<sup>20</sup> – it is, they say, one of those pesky traits that evolved in a context where it paid to be concerned with the in-group and not particularly sympathetic to the out-group.<sup>21</sup>

But if bio-intervening in particular dispositions, motivations, or behaviours cannot reliably produce a change that by itself could be judged to be morally better, then it is not clear that such interventions can result in an *enhancement*. “Enhancement” after all, connotes *improvement*. If using a particular bio-intervention affecting altruistic dispositions does not necessarily result in an appropriately altruistic person – rather than in a person who is unduly concerned with the interest of in-group members – then one wonders in what sense they have been enhanced. Thus, one can agree that bio-interventions might produce modifications in people’s dispositions, motivations,

<sup>20</sup> See, for instance, Persson and Savulescu, *Unfit for the Future*.

<sup>21</sup> In another contradictory statement, Persson and Savulescu also claim that altruism is *essentially* a moral disposition, unlike the courage or strength of will that they contend can characterise criminals. See Persson and Savulescu, ‘The Perils of Cognitive Enhancement’, 72.

## The Trouble With Moral Enhancement

or behaviours, but such modifications need not result – and often will not, as proponents acknowledge – in an *enhanced*, i.e., improved, bettered, person.

Now, talk of motivational, dispositional, or behavioural modifications seems significantly less radical than “moral enhancement” would suggest. More importantly, insofar as the purported goal of the moral enhancement project is to protect humanity from “ultimate harm”, talking about motivational, dispositional, or behavioural alterations would call into question the effectiveness of that project – terrorists, for instance, could have their empathy “enhanced” with not particularly good results. In fact, the belief that reliably producing the right kind of modifications – making people more appropriately altruistic, less aggressive, or more suitably empathetic – would constitute any protection against “ultimately harm” is patently absurd. Those determined to obliterate the planet are, after all, unlikely to believe that their moral self is in need of fixing, and thus unlikely to voluntarily use any enhancing interventions. Even if one could make such interventions compulsory, the loners likely to cause such ultimate harm will be somewhat difficult to find. And, of course, none of the bio-interventions proposed are powerful enough to deter someone who is completely committed to destroying the planet.

It seems then that by the proponents’ own lights the moral enhancement project fails to constitute *enhancement* and fails to enhance *morality*. Indeed, recognition of these problems has led proponents of moral enhancement to talk about “modulation” and using bio-interventions to enhance second-order moral capacities. It is true that at least some possible conceptualisations of the new types of targets can reasonably be understood as involving morality. However, to the extent that such is the case – that is, to the extent that the bio-interventions in question attempt to create individuals who *are* good (have the right kind of dispositions and motivations) and *do* what is right, for the *right* reasons – the notion that any of the candidates suggested for these interventions (SSRIs, oxytocin, propranolol) would achieve any such thing betrays an astounding lack of understanding of the complexities of human biology, the nature of normativity, and the relevance of language.<sup>22</sup>

Note that the problems mentioned – fatal ones, in my opinion – do not involve appeals to disagreements regarding whether particular

<sup>22</sup> For arguments about some of these complexities see Harris Wiseman, *The Myth of the Moral Brain: The Limits of Moral Enhancement* (Cambridge, Massachusetts: MIT Press, 2016).

motives, dispositions, or behaviours are or not in fact morally right. We can all agree that having certain altruistic or empathetic dispositions or less biased motives is a good thing. Nonetheless we can still contest the belief that enhancing those traits would make anyone more *moral*. Of course, the fact that morality is context-dependent and that legitimate disagreements can exist about whether particular actions, motivations, or dispositions are right or wrong, permissible or impermissible, makes the whole project even more hopeless.

Proponents of the moral enhancement project might object that education, or what they insist on calling “traditional means of enhancement”, involves just this same piecemeal modification of people’s motives, dispositions, and behaviours. But this is simply false. When we educate children to become moral agents we do not – or not usually – target a particular motivation or disposition, but rather target a *person*, situated in a particular context, using a particular language, and bringing with us particular assumptions about what is right and wrong, what we think might be best for them, directed at them, and attending to their particularities. It is the whole package that concerns us when we practice moral education, not a specific motivation or disposition – even if we are trying to modify some motivation or disposition. The development of moral agency requires engagement with other human beings who live in certain social and political contexts, and it is in that way that – hopefully – we help children to become morally competent adults.<sup>23</sup>

Denying that education is not a “traditional means of enhancement” is not simply a matter of semantics. In contesting such an analogy I want to call attention to the fact that the goal of bioenhancement and the goal of moral education are similar only in appearance. It is not just that the means are obviously different, it is that the *ends* are different also. Indeed, this is why the proponents’ strategy of equating the ends of moral bioenhancement with those of moral education are so rhetorically powerful. Once we agree that the end is worth pursuing – and who would not want to have people who are more moral? – then the discussion can proceed to focus on the *means* to achieve that goal. And an emphasis on the means leads effortlessly to a concern with the usual inventory of risks and benefits.<sup>24</sup>

<sup>23</sup> Or what, in another rhetorical move, proponents call “morally enhanced individuals with respect to children”.

<sup>24</sup> And it is indeed talk of risks, rather than uncertainties and ignorance, that abounds in the enhancement debate. The language of risks reduces issues of uncertainty, ambiguity, and ignorance to the more controllable and deterministic processes usually associated with risk evaluations, and



## The Trouble With Moral Enhancement

Proponents believe, of course, that the benefits are unquestionably good, while they take the risks to be either not that serious, or, insofar as they do present some serious risks, they have confidence that we will be able to manage them.<sup>25</sup> In fact, when one has framed the need for moral enhancement as the only way to avoid humanity's demise, there will not be many risks that appear not worth taking. Unsurprisingly, given this framing, criticisms of interventions aimed at enhancing human beings – morally or otherwise – are met with accusations of conservatism and delaying progress. Those who frame the debate have the power to direct the discussion.<sup>26</sup>

So far, I have argued that conceptual problems reveal the moral enhancement project to be nothing of the sort: insofar as the targets of biomedical interventions are particular dispositions, motivations, or behaviours, it is incorrect to talk about *moral* enhancement rather than dispositional, motivational, or behavioural enhancement. And insofar as these are understood as piecemeal characteristics or properties, it is mistaken to talk about *enhancement* rather than simply modifications or alterations.

But the moral enhancement project also rests on shaky scientific grounds. In fact, there are so many problems related to the misuse of scientific evidence that it would be difficult to even mention all of them here. I will thus focus on two that are particularly problematic: the use of scientific claims that proponents present as uncontroversial and the weakness of the scientific evidence purporting to show that moral enhancement is plausible.

Moral enhancement supporters attempt to buttress their proposals by appealing to various sources of scientific evidence.<sup>27</sup> In doing so,

---

obscures the multiple uncertainties in the development and implementation of biomedical interventions.

<sup>25</sup> Such confidence in humanity's ability to manage even serious risks confronting the development and use of bioenhancements constitutes another inconsistency present in moral enhancement proposals. Why – it seems imperative to ask – would we trust members of a species that is on the brink of destroying the planet and everything in it to appropriately manage such risks?

<sup>26</sup> For a discussion of the importance of the framing of the moral enhancement project, see de Melo-Martín and Salles, 'Moral Bioenhancement: Much Ado About Nothing?'

<sup>27</sup> See, for instance, Douglas, 'Moral Enhancement'; Persson and Savulescu, 'Getting Moral Enhancement Right'; Persson and Savulescu, 'The Perils of Cognitive Enhancement'; Savulescu and Persson, 'Moral

they not uncommonly present some such scientific evidence as uncontroversially accepted. In fact, however, many of the scientific claims they present are not only highly contentious but by many accounts simply false. This is the case, for instance, regarding many of the evolutionary psychological claims proponents use to support the moral enhancement project. Many such claims have been discredited for multiple reasons, from problematic assumptions, to incorrect interpretations of the evidence, to inadequate conclusions.<sup>28</sup> Some of the claims proponents make in this regard are plainly ridiculous. For instance, trying to argue that altruism and a sense of justice have a genetic basis, Persson and Savulescu offer the following evidence: '[i]t is plausible to think that in general women have a greater capacity for altruism than men. If this psychological difference tracks gender, this is surely good evidence that it is biologically based'.<sup>29</sup>

Another area of research often cited to support their claims regarding the plausibility of using bio-interventions to enhance people morally involves twin studies. Proponents use such studies to defend their claim that the human sense of fairness and altruism have a genetic basis. This is proven, they argue, because identical twins present striking correlations in what they consider fair and unfair in ultimatum games, while such correlations are lacking in the case of fraternal twins. Identical twins apparently also present striking correlation in respect to altruism.<sup>30</sup> As with evolutionary psychology, many people have criticised twin studies on both methodological and

---

Enhancement, Freedom and the God Machine'; Persson and Savulescu, *Unfit for the Future*; T. Douglas, 'Moral Enhancement Via Direct Emotion Modulation: A Reply to John Harris', *Bioethics* 27:3 (2013), 160–68; DeGrazia, 'Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour'.

<sup>28</sup> For some work criticising many of the claims made in evolutionary psychology see, for instance, Hilary Rose and Steven P. R. Rose, *Alas, Poor Darwin: Arguments Against Evolutionary Psychology* (New York: Harmony Books, 2000); John Dupré, *Human Nature and the Limits of Science* (New York: Oxford University Press, 2001); Cordelia Fine, *Delusions of Gender: How Our Minds, Society, and Neurosexism Create Difference* (New York: W. W. Norton, 2010).

<sup>29</sup> Persson and Savulescu, 'Getting Moral Enhancement Right', 130.

<sup>30</sup> Persson and Savulescu, 'Getting Moral Enhancement Right', 130; Persson and Savulescu, 'The Perils of Cognitive Enhancement', 171; Savulescu and Persson, 'Moral Enhancement, Freedom and the God Machine'.

## The Trouble With Moral Enhancement

substantive grounds.<sup>31</sup> To present results from such studies as accepted evidence of the genetic basis of moral dispositions and emotions is therefore highly problematic.

I hasten to note that questioning such evidence in no way means that no biological basis for moral dispositions or emotions exist. It is, of course, unlikely that such a basis is genetic, if this is understood as proponents seem to: that is, as one that would allow us to tinker with some genes so as to reliably produce particular effects. There is, of course, no such thing as a gene for altruism or for a sense of justice. But one need not deny that modifications of, say, neurotransmitters – which are clearly biological – can produce emotional and dispositional changes. Nonetheless, to interpret evidence for such modifications as evidence for the plausibility of moral enhancement – or more precisely, for the plausibility that some morally relevant aspects can be biomedically modified – involves more than just looking into people's brains or making them play some lab game or another.

Proponents of moral enhancement use different sources of evidence as proof that biomedical interventions can affect various aspects of people's moral life. For instance, they defend their claims by citing, among others, experiments that purport to show the effect of propranolol on implicit bias,<sup>32</sup> neuroimaging studies showing that the amygdala plays a role in race aversion,<sup>33</sup> and experiments attempting to assess the effect that selective serotonin reuptake inhibitors (SSRIs) have on harm aversion.<sup>34</sup>

Let us accept, if only momentarily, that the results of some of these scientific studies regarding enhancement interventions are correct, that

<sup>31</sup> See, for instance, J. Joseph, 'Twin Studies in Psychiatry and Psychology: Science or Pseudoscience?', *Psychiatric Quarterly* **73**:1 (2002), 71–82; P. V. Tishler and V. J. Carey, 'Can Comparison of Mz- and Dz-Twin Concordance Rates Be Used Invariably to Estimate Heritability?', *Twin Res Hum Genet.* **10**:5 (2007), 712–717; K. Richardson and S. Norgate, 'The Equal Environments Assumption of Classical Twin Studies May Not Hold', *Br J Educ Psychol.* **75**:3 (2005), 339–350.

<sup>32</sup> See Douglas, 'Moral Enhancement'; T. Douglas, 'Moral Enhancement via Direct Emotion Modulation'; Savulescu and Persson, 'Moral Enhancement, Freedom and the God Machine'; DeGrazia, 'Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour'.

<sup>33</sup> Douglas, 'Moral Enhancement'.

<sup>34</sup> Savulescu and Persson, 'Moral Enhancement, Freedom and the God Machine'; DeGrazia, 'Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour'.

is, that they show what they purport to show about the effects of propranolol, SSRIs, and so on. Consider, for instance, proponents' claims that implicit racial biases are indeed attenuated with the use of propranolol. Such results can be taken to constitute evidence for the plausibility of moral enhancement only if one presupposes various contested assumptions, among them that implicit biases express some negative appraisal or attitude, that they have some negative effect on the treatment of racial minorities, and that people who hold implicit biases are somehow morally deficient.<sup>35</sup> That implicit biases express some negative attitude and are thus morally suspect might seem obvious to many, as "bias" is often used normatively. If indeed those biases express aversive or hostile attitudes, as proponents of moral enhancement appear to believe, it would make sense to attempt to correct them and to take such corrections as an indication that at least some morally relevant attitudes can be bettered. But some evidence indicates that implicit bias might be the outcome of anxiety or discomfort rather than hostility.<sup>36</sup> Some research also suggests that sympathy, rather than hostility, might underlie the responses to outgroup faces, and that implicit biases result from anxiety about past oppression and mistreatment associated with some racial groups.<sup>37</sup> Thus, it is not uncontested that implicit biases express some morally negative attitude. Furthermore, the evidence that implicit biases, at least as measured by the Implicit Association Test (IAT), correlate with negative behaviour towards racial minorities is similarly disputed.<sup>38</sup> Moreover, even assuming that implicit biases express negative attitudes, whether people who have such biases are morally deficient – and thus whether they can plausibly be understood as improving some aspect of their

<sup>35</sup> For a more detailed discussion of these assumptions see de Melo-Martín and Salles, 'Moral Bioenhancement: Much Ado About Nothing?'

<sup>36</sup> C. Frantz, A. Cuddy, M. Burnett, H. Ray, and A. Hart, 'A Threat in the Computer: The Race Implicit Association Test as a Stereotype Threat Experience', *Pers Soc Psychol Bull.* **30**:12 (2004), 1611–24.

<sup>37</sup> E. Uhlmann, V. Brescoll, and E. Paluck, 'Are Members of Low Status Groups Perceived as Bad, or Badly Off? Egalitarian Negative Associations and Automatic Prejudice', *J Exp Soc Psychol.* **42**:4 (2006), 491–99; M. Andreychik and M. Gill, 'Do Negative Implicit Associations Indicate Negative Attitudes? Social Explanations Moderate Whether Ostensible "Negative" Associations are Prejudice-Based or Empathy-Based', *J Exp Soc Psychol.* **48**:5 (2012), 1082–93.

<sup>38</sup> F. Oswald, G. Mitchell, H. Blanton, J. Jaccard, and P. Tetlock, 'Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies', *Journal of Personality and Social Psychology Studies*, **105**:2 (2013), 171–192.

## The Trouble With Moral Enhancement

morality when using particular biomedical interventions such as propranolol – depends not on determining that such implicit biases are indeed reduced, but on complex normative judgements about, among other things, the causative role of unconscious motivation and our moral responsibility for unconscious beliefs.

It is true that to the extent that implicit biases have negative implications for the treatment of various racial groups, their elimination or reduction would be a very desirable thing regardless of whether the source of the bias is morally unproblematic or whether people are morally responsible for them. If implicit biases do result in discriminatory practices against out-groups particularly, the world would surely be a better place if people did not have such biases. However, if as mentioned earlier, implicit biases are the result, for instance, of recognising that some groups have been unjustly treated and that they are owed compensation for such injustices, it is at least not obvious that eliminating the source of the implicit biases would constitute a betterment of the individuals in question. Similarly, the issue of whether or not people are responsible for implicit bias is very much of relevance to whether we can say that someone's attitudes have been improved when such biases are eliminated or reduced. Thus, maintaining that, regardless of people's responsibility for, or the source of, implicit biases, those who have them are necessarily morally deficient – as proponents must assume if they want to use evidence that such biases are reduced as support for their project – presupposes substantive normative judgements that need to be defended and critically assessed.

It is thus not clear at all that experiments showing that propranolol reduces implicit bias constitute evidence that peoples' moral attitudes can be improved in some morally relevant sense, and thus that such experiments constitute evidence for the plausibility of moral enhancement interventions. But there are still more problems with the evidence presented by moral enhancement proponents. Why should we take the results of such experiments to be showing that propranolol, for instance, does eliminate or reduce implicit biases, that the ultimatum game provides us with relevant information about some evolved sense of fairness, or that use of SSRIs sheds any light on how human beings appraise harm? In fact, we have very good reasons to be sceptical of putting too much stock in these and similar results. This is for several reasons. First, the experiments involve only a handful of people.<sup>39</sup> Second, the conclusions of these

<sup>39</sup> M. J. Crockett, L. Clark, G. Tabibnia, M. D. Lieberman, and T. W. Robbins, 'Serotonin Modulates Behavioral Reactions to Unfairness', *Science* **320**:5884 (2008), 1739; M. J. Crockett, L. Clark,

studies are grounded on results that come from the highly artificial conditions involve in laboratory experiments. Third, the tests used, i.e., the IAT to measure implicit bias, the ultimatum game, etc., have been tested overwhelmingly on similar populations – the so-called WEIRD (Western, Educated, Industrialised, Rich, and Democratic) populations. Indeed, evidence shows that within the field of psychology, 96% of psychological samples come from countries with only 12% of the world's population.<sup>40</sup> Similarly, within the field of human neuroimaging, 90% of peer-reviewed neuroimaging studies come from Western countries.<sup>41</sup>

But why would these factors be relevant to the conclusions of the studies discussed? First, insofar as the number of people participating in these studies is limited, it calls for caution in the interpretation of the results. After all, small numbers of research subjects make statistical analyses difficult.<sup>42</sup> Second, the fact that these are laboratory studies assessing extremely complex aspects of human psychology and neurobiology, and that the tests have been validated in a very particular population, calls for restraint regarding the generalisation of these results. Indeed, the fact that the overwhelming majority of studies in psychology and neuroimaging use WEIRD people as test subjects and that the tests used in such experiments have been primarily validated with WEIRD people would not be a significant problem if it were the case that human psychology and neurobiology are unaffected by cultural and social factors. But it turns out that when researchers have taken the time to test other populations, these other populations do not respond as the WEIRD populations do.<sup>43</sup> Some studies have shown that differences exist even in brain activation based on socio-economic status and that therefore not even brain imaging studies

---

M. D. Hauser, and T. W. Robbins, 'Serotonin Selectively Influences Moral Judgment and Behavior through Effects on Harm Aversion', *Proceedings of the National Academy of Sciences of the United States of America* **107**:40 (2010): 17433–17438; S. Terbeck, G. Kahane, S. McTavish, J. Savulescu, P. J. Cowen, and M. Hewstone, 'Propranolol Reduces Implicit Negative Racial Bias', *Psychopharmacology* **222**:3 (2012): 419–24.

<sup>40</sup> J. Henrich, S. J. Heine, A. Norenzayan, 'The Weirdest People in the World?' *Behav Brain Sci.* **33**:2/3 (2010), 61–83.

<sup>41</sup> J. Y. Chiao, 'Cultural Neuroscience: A Once and Future Discipline', *Progress in Brain Research* **178** (2009), 287–304.

<sup>42</sup> See, on this point, J. Ioannidis, 'Why Most Published Research Findings Are False', *PLoS Medicine* **2**:8 (2005), e124.

<sup>43</sup> Henrich, Heine, and Norenzayan, 'The Weirdest People in the World?'

## The Trouble With Moral Enhancement

can ignore population-level variations. In fact, the much touted high heritability of IQ, for instance, has been concluded from studies done overwhelmingly with children of high socioeconomic status (SES). But some evidence indicates that for high-SES children, where environmental variability is negligible, the genetic differences account for 70–80% of the variation, with shared environment contributing less than 10%. However, for low-SES children, where there is far more variability in environmental contributions to intelligence, genetic differences account for 0–10% of the variance, with shared environment contributing about 60%.<sup>44</sup> In fact, the findings of studies done with diverse populations suggest that members of WEIRD societies, including young children, are among the least representative populations one could find for generalising about human beings' psychology or neurobiology.<sup>45</sup> Many of these findings involve domains that are associated with fundamental aspects of psychology, motivation, and behaviour, and, if this is so, we have very good reasons to be sceptical about claims that a particular behavioural phenomenon is universal based on sampling from a single subpopulation.

Where do all of these problems leave the moral enhancement project? I believe that it leaves it in a not so good place. As the arguments presented here show, both the conceptual and the evidentiary grounds for moral enhancement are shaky. Let me, however, end this contribution by noting that the fact that biomedical interventions fail to constitute enhancements or that, insofar as they do, they fail to enhance morality, means neither that such interventions could have no effect on various aspects of our moral life, nor that the likely-scarce efficacy of these interventions would prevent their development and implementation. Surely, however, claims about the necessity of these interventions to prevent our demise are not going to be particularly helpful in promoting an informed and critical dialogue.

*Weill Cornell Medicine, Cornell University*  
[imd2001@med.cornell.edu](mailto:imd2001@med.cornell.edu)

<sup>44</sup> E. Turkheimer, A. Haley, M. Waldron, B. D'Onofrio, and I. I. Gottesman, 'Socioeconomic Status Modifies Heritability of IQ in Young Children', *Psychological Science* **14**:6 (2003), 623–28.

<sup>45</sup> Henrich, Heine, and Norenzayan, 'The Weirdest People in the World?'