

# Moral enhancement, freedom, and what we (should) value in moral behaviour

David DeGrazia

## Correspondence to

Dr David DeGrazia;  
ddd@gwu.edu

Received 9 May 2012  
Revised 12 September 2012  
Accepted 16 September 2012  
Published Online First  
25 January 2013

## ABSTRACT

The enhancement of human traits has received academic attention for decades, but only recently has moral enhancement using biomedical means – *moral bioenhancement* (MB) – entered the discussion. After explaining why we ought to take the possibility of MB seriously, the paper considers the shape and content of moral improvement, addressing at some length a challenge presented by reasonable moral pluralism. The discussion then proceeds to this question: Assuming MB were safe, effective, and universally available, would it be morally desirable? In particular, would it pose an unacceptable threat to human freedom? After defending a negative answer to the latter question – which requires an investigation into the nature and value of human freedom – and arguing that there is nothing inherently wrong with MB, the paper closes with reflections on what we should value in moral behaviour.

## INTRODUCTION

The improvement of human traits or capacities can be a source of great value.<sup>i</sup> Sometimes we value an improvement intrinsically. One might celebrate an improved ability to sing on key, for example, as valuable in itself irrespective of further benefits that might accrue from the increased mastery. In many cases, an improvement is valued instrumentally. Thus, a swimmer might appreciate increased lung capacity as a means to swimming faster in races. Of course, many improvements are valued instrumentally and intrinsically. In the literature we will engage, the relevant improvements are usually referred to as *enhancements*.

The enhancement of human traits has received sustained academic attention for over two decades. Much of the attention has focused on enhancements in biomedical contexts. Thus, many articles and books have appeared on such topics as cosmetic surgery, doping in sports, the pharmacological enhancement of mood and cognition, as well as the prospective genetic enhancement of human traits.<sup>ii</sup>

In these discussions, enhancement has often been conceptualised by way of contrast with treatment or therapy: Enhancements are understood as

interventions designed to improve human form or function without responding to genuine medical need.<sup>3</sup> In this conceptual scheme, a genetic intervention that aims to reduce the effects of muscular dystrophy would count as treatment whereas a similar intervention that is intended to increase muscularity in a healthy individual would count as enhancement. Obviating the need to distinguish enhancement from treatment in questionable cases, and furnishing a conception that is not restricted to biomedical contexts, an alternative approach makes no reference to treatment, medical need or even the much-contested concept of normalcy. In keeping with this conceptual alternative, I will here define ‘enhancement’ as *any deliberate intervention that aims to improve an existing capacity, select for a desired capacity, or create a new capacity in a human being*.<sup>iii</sup> This covers such non-biomedical enhancements as education, music lessons and specialised vocational training as well as biomedical enhancements (including medical therapies that aim to improve capacities). Notably, among the biomedical interventions, enhancement in this sense includes embryo selection, not just interventions on a given individual.

In recent years a literature has emerged on the topic of *moral* enhancement: interventions that are intended to improve our moral capacities such as our capacities for sympathy and fairness.<sup>5–7</sup> While commentators agree that many traditional means of moral enhancement such as explicit instruction and consciousness-raising groups are legitimate and important, they disagree over whether biomedical and perhaps other non-traditional means of moral enhancement are desirable. The present paper will enter this discussion. Before doing so, it will be helpful to identify examples of non-traditional means of moral enhancement—a few already attempted, others merely prospective—and then to state more precisely the questions to be addressed.

Here are some examples of non-traditional means of moral enhancement:

- ▶ Glucose as a means of increasing resistance to temptation to do something wrong or to stop trying to do what one should<sup>8</sup>
- ▶ Selective serotonin reuptake inhibitors as a means to being less inclined to assault people<sup>9</sup>
- ▶ Propranolol as a means of decreasing unconscious racial bias<sup>10</sup>

<sup>i</sup>This paper has benefited from feedback following presentations at the Carnegie Council of New York and the Department of Philosophy, George Washington University. Special thanks to Tom Douglas, Ingmar Persson, David Wasserman, Eric Sidel, Tad Zawidzki and two anonymous reviewers for written comments and to Jonathan Haidt for pressing me on moral pluralism. Thanks also to Guy Kahane for bibliographical suggestions and Jacqueline Drayer for research assistance.

<sup>ii</sup>For two excellent anthologies covering a wide range of topics, see references 1 and 2.

<sup>iii</sup>Cf. reference 4, p 23 although Buchanan makes no reference to selection for desired capacities. Like the more conventional definitions of ‘enhancement’, this one focuses on interventions that *aim at* certain improvements or desired results whether or not they actually come about.



Editor's choice  
Scan to access more  
free content



- ▶ <http://dx.doi.org/10.1136/medethics-2012-101092>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101094>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101096>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101153>



CrossMark

To cite: DeGrazia D. *J Med Ethics* 2014;**40**:361–368.

- ▶ Deep-brain stimulation (eg, electrical stimulation of the amygdala) as a means to reducing aggression<sup>11</sup>
- ▶ Neurofeedback to increase sympathy and/or treat antisocial personality disorder (psychopathy)<sup>12 13</sup>
- ▶ Selection of embryos that contain a gene coding for a greater disposition to altruism<sup>iv</sup>
- ▶ Genetic interventions to gametes, embryos or postnatal human beings as a means to the same end
- ▶ Embryo selection or genetic engineering as a means of avoiding or neutralising genes associated with antisocial personality disorder<sup>v</sup>
- ▶ Either of these means as a way of securing a stronger predisposition to fairness
- ▶ An artificial chromosome that includes multiple genes coding for stronger predispositions to a variety of moral virtues<sup>vi</sup>

For the remainder of this paper, I will assume that the interventions under consideration are more ‘high tech’ than the administration of glucose. I will refer to such interventions collectively as *moral bioenhancement* or *MB* for short.

The specific questions to be addressed in what follows are these: (1) Why might one take the possibility of MB seriously? (2) In view of pluralism about moral values, what sorts of changes, if any, may we confidently and responsibly count as moral improvements? (3) Given answers to the previous questions—and assuming MB were safe, effective and universally accessible—would it be morally desirable? In particular, would it pose an unacceptable risk to human freedom? (4) What do our reflections suggest about what we should value in moral behaviour?

### MOTIVATING MORAL BIOENHANCEMENT

Why might one take the possibility of MB seriously? We already have traditional means of moral enhancement. Attempting to develop and deploy new, high tech means may introduce new risks, including that of great clumsiness (‘taking an axe to fix a watch’). Moreover, it is hard to imagine an implementation scheme that would be effective in reaching enough people, or the right people, for a moral improvement in a given population while adequately respecting autonomy, specifically the prerogative to decline particular enhancements for oneself or one’s dependents. Nevertheless, MB deserves consideration because the status quo of moral behaviour is deeply problematic and traditional means of moral enhancement may prove inadequate to achieve needed improvements—notwithstanding the phenomenon of moral progress.

The status quo is deeply problematic because there is such an abundance of immoral behaviour, with devastating consequences, and serious risk of worse to come. Consider examples. In the 1990s, genocides occurred in Rwanda and Bosnia as world powers looked on.<sup>vii</sup> Slavery still exists in some parts of the world. Forced prostitution and participation in pornography, often involving children, is a reality on nearly every continent. Around the world, violence and other forms of oppression are committed against

girls and women, religious and ethnic minorities, and others who are considered ‘outside’ the group committing the violence. Meanwhile, ‘(e)ach day, some 50,000 human beings ... die from starvation, diarrhea, pneumonia, tuberculosis, malaria, measles, perinatal and maternal conditions, and other poverty-related causes.’ (p 30).<sup>19</sup> An estimated 850 million people are chronically undernourished, over a billion lack access to safe water, 2.6 billion lack access to basic sanitation while a billion have inadequate shelter and 2 billion lack electricity.<sup>viii</sup> In response to such pervasive severe poverty—and despite being causally implicated in much of it—leaders of developed countries fail to provide the modest levels of foreign assistance that could contribute greatly towards solving these problems. For example, the USA devotes just one-tenth of 1% of its Gross Domestic Product to foreign assistance, much of it to countries wealthy enough to reciprocate rather than to the poorest countries.<sup>ix</sup> At the same time, the USA commits far greater sums to subsidies and tax breaks for those who don’t need them. Nor, with few exceptions, do individuals contribute substantially to prevent these preventable deaths and subhuman living conditions.

In addition to these harms and injustices, there is the threat of *truly massive harm*: nuclear annihilation, pandemics caused by the intentional spread of smallpox, the ebola virus, or biological weapons, global economic ruin caused by cyber terrorism, eventual devastation by the effects of global climate change, or the like. Persson and Savulescu plausibly argue that the risk of truly catastrophic harm increases as our technologies and technical know-how develop.<sup>6</sup> It is increasingly possible for a small number of individuals to acquire the technical capability of inflicting terrible harm.

These examples and reflections support the claim that the moral status quo is deeply troubling. Some would maintain that it is intolerable. Again, we already possess means of enhancing moral capacities. These include explicit moral instruction, mentoring, socialisation, carefully designed public policies, consciousness-raising groups, literature and other media that encourage moral reflection, and individual efforts at improvement. But these means have been, at best, modestly effective: hence the troubling status quo with a real prospect of far worse to come. Moreover, growth in technology is extremely fast while positive change in our psychology is slow. Because the moral status quo is so problematic, and traditional means of moral enhancement may be incapable of yielding satisfactory improvement, the possibility of MB deserves our open-minded consideration. I trust that most of those who are more impressed than I am by the moral progress that has been achieved in recent decades, and who have more confidence than I have in traditional means of moral enhancement, can accept my modest call for open-minded consideration of MB.

### ON THE SHAPE AND CONTENT OF MORAL ENHANCEMENT

In order to consider the possibility of MB intelligently, and not just open-mindedly, we must ask what a moral improvement, regardless of its means, would consist in.

#### An anatomy of moral improvement

We may distinguish three relevant kinds of improvement:

1. *Motivational improvement*: better motives, character traits and overall motivation to do what is right.

<sup>iv</sup>See reference 14 for evidence that a particular gene involved in dopamine production is significantly implicated in the disposition to altruism.

<sup>v</sup>Twin studies focusing on aggression and antisocial behaviour suggest significant heritability of these traits.<sup>15</sup>

<sup>vi</sup>I discuss this possibility in reference 16, chapter 3. My thought-experiment about moral enhancement via an artificial chromosome drew several key ideas from reference 17, *Epilogue*.

<sup>vii</sup>For a detailed depiction of massive moral horrors in the twentieth century, including these genocides, see reference 18

<sup>viii</sup>See reference 19, p 12. Pogge cites reference 20, p 24 and reference 21, p 49.

<sup>ix</sup>See reference 22, p 388. Pogge cites reference 23, p 202.

2. *Improved insight*: better understanding—accessible when decisions are needed—of what is right.
3. *Behavioural improvement*: greater conformity to appropriate moral norms and therefore a higher frequency of right action.

This classification may be oversimplified. For one thing, ‘internalists’ in moral philosophy maintain that it is impossible to make a judgment about what is right without having at least some motivation to comply with the judgment, suggesting that moral insight and moral motivation are not entirely separable. (For the record, I reject internalism.) Moreover, some philosophers and psychologists believe that emotions play a role in informing us about what’s right, again suggesting that insight and motivation are not completely separable.<sup>24</sup> This may be so. Despite possible oversimplification, I believe the suggested anatomy of moral improvement will facilitate discussion of the targets of MB and of what we value in moral behaviour.

Motivational improvement and improved insight are conducive to behavioural improvement. Other things being equal, either type of improvement will tend to bring about better behaviour. If someone’s moral motivation remains constant as she gains moral insight, she will become more likely to do what is right (assuming she has any motivation to do so) because she’s more likely to know what is right and therefore what to do. Conversely, if someone’s moral insight remains constant as his moral motivation improves, he will become more likely to do what is right (assuming he has any moral insight at all) because he will be more motivated to do it. A fortiori, the conjunction of motivational improvement and improved insight will be conducive to behavioural improvement. Let us now consider each kind of improvement in greater depth.

Improved insight is highly *cognitive* and may come about through cognitive enhancement. Knowing what is right, after all, is a kind of knowledge. Here one is reminded of Aristotle’s distinction between (1) ‘moral virtues’ (by our modern lights, a combination of moral and prudential virtues) such as courage, temperance, and justice and (2) intellectual virtues such as theoretical wisdom and practical wisdom (Bks. II-IV)<sup>25</sup>. Moral insight as construed here is closely related to practical wisdom as Aristotle understood it: an intellectual strength that is indispensable for making good choices.

By contrast, motivational improvement is commonly regarded as more *conative* (pertaining to caring) or *affective* (pertaining to emotions). This sort of improvement seems more clearly a type of *moral* improvement than improved insight is. We tend to esteem a high degree of moral motivation. Kant even thought that the good will—the will to do everything in one’s power to do what is right—was the only thing in the world that was unconditionally valuable (chap. 1)<sup>26</sup>. (But Kant also thought that the good will was guided by reason and not by affect or emotion, that reason could generate motivation on its own.) Someone who acts with a perfectly good will in this very demanding sense leaves no room for motivational improvement. Often we act with a relatively good will. Some theorists, however, construe motivational excellence less in terms of a good will than in terms of particular moral virtues such as sympathy and fairness.

Regarding the third kind of moral improvement, I make the following assumption: *Behavioural improvement is highly desirable in the interest of making the world a better place and securing better lives for human beings and other sentient beings*. It follows that motivational improvement and improved insight, given their conduciveness to behavioural improvement, are at least instrumentally valuable. One or both may also be intrinsically valuable.

Note, finally, that the concept of moral improvement can be understood to apply to an already-existing individual or to choices about which individuals will come into being and be brought to term. Thus, we might characterise the selection of an embryo possessing a gene coding for a disposition to fairness rather than an embryo lacking this gene, or of an embryo lacking a gene associated with antisocial personality disorder rather than an embryo possessing this gene, as an attempt at moral improvement. Here the intended improvement is applied not to a given individual, but to a choice among potential persons or ‘across possible persons.’ The improvement is trans-individual rather than intraindividual. A large number of choices of this kind might be regarded as an improvement in some relevant population.

### Addressing a challenge about the content of moral improvement

Before we consider the desirability of MB, let us address a sceptical challenge that stems from moral pluralism. People have different views about what a moral life would look like. Some think that morality requires behaviour in accordance with politically conservative values: a premium on individual responsibility and avoidance of ‘free lunches,’ adherence to time-honoured social expectations and moral rules, strong partiality towards members of one’s own community, acceptance of a government that mostly leaves free enterprise and family life to their own devices, and possibly religion. Others think that morality requires politically progressive values: compassion for those whose basic needs are not met, acceptance of a government that helps to meet those needs, an impartial and inclusive view of humanity (and maybe animals), rejection of traditions and mores that seem regressive, and a fully voluntary relationship to religion. Among ethical theorists, some are deontologists with significant non-consequentialist commitments, some are consequentialist, and some turn their attention away from right action towards virtue, politics or gender relations. Nor are these differences purely theoretical, for they often translate into different normative positions on such matters as abortion, the death penalty, affirmative action, drug use, euthanasia, warfare, animal usage, and world hunger.

Amid this diversity of views about what morality requires, how can we determine what sort of change should count as a moral improvement? It is not enough to speak of greater conformity with appropriate moral norms, or the insight or motivation that conduces towards such conformity, if we have no idea which moral norms *are* appropriate. Without knowing the criteria of right (or morally best) action, according to the present challenge, we can’t know what counts as greater insight into what is right or what motivational factors would help us do what’s right.

This is a significant challenge to any enthusiast of moral enhancement—including traditional moral enhancement. Thus, parents who believe in providing moral instruction to their children face this challenge. One should not inculcate moral values that are wrong, so how can a parent be sure that she or he is justified in providing a particular type of moral instruction? Also facing this challenge are public school teachers who attempt to inculcate in students certain moral virtues such as civility, respect for differences and concern for the poor.

When public schools inculcate values appropriately, they stick to values that are not really debatable among people who genuinely care about morality. So, in a way, does criminal law insofar as it punishes murder, theft and rape but does not punish a failure to contribute to charity or a failure to work hard and

cultivate self-reliance. This suggests a way of determining what should count as an enhancement in considering what MB research the state should support and what kinds of MB, if any, the state might encourage or possibly even require: Stick to improvements that represent *points of overlapping consensus among competing, reasonable moral perspectives*. What do I mean by ‘reasonable’ here? It is very difficult to provide a satisfying, non-question-begging answer, but let me venture a brief reply—while noting, again, that the challenge of determining what counts as a moral improvement confronts traditional means of moral enhancement no less than it confronts MB.

Leading contemporary progressive and conservative visions—ranging from socialism to welfare-state capitalism to moral conservatism to libertarianism—count as reasonable whereas neo-Nazism, apartheid and the Taliban’s worldview (at least as regards women’s status) do not. Consequentialist, deontological, virtue-based and feminist views that accord persons some sort of moral equality qualify as reasonable; Nietzschean elitism according to which only the most powerful and creative are worthy does not. As far as justified state policy goes, the idea is to locate points of overlapping consensus among reasonable visions.

Now, it is generally accepted that parents enjoy greater latitude than the state in raising their children as they see fit, a prerogative grounded in the importance ascribed to family life and the presumptive freedom from interference that it requires in order to thrive. Consistent with this value, it might make sense to permit parents to adopt more debatable visions of morality—among reasonable alternatives—and therefore of what counts as moral enhancement, and to select forms of MB accordingly. This would be a biological extension of the current parental prerogative to provide moral instruction to their children as parents see fit. Competent adults deciding for themselves could also be accorded wider latitude to determine what should count as a moral enhancement.

Public policy decisions regarding MB should be acceptable from any reasonable view of morality. I suggest that this area of overlapping consensus is fairly broad. That, in large part, is because we can agree on many things *that can go wrong* with our moral agency: our motives, our insight or our consequent behaviour.<sup>x</sup> Thus, our public policies might support research into and possibly—if and when some forms of MB are demonstrably safe and effective and the state is prepared to make them universally available—encourage or even require the use of certain MBs that help to reduce or eliminate any of the following moral defects.

- ▶ Antisocial personality disorder, a severe failure of motivation
- ▶ Specific forms of evil such as sadism and intrinsic delight in cheating others, another severe failure of motivation
- ▶ Lesser forms of moral cynicism that make one more likely than a good person to be corrupted, to cheat on taxes, not to bother to contribute what one agrees is one’s fair share, etc—a more ordinary failure of motivation
- ▶ Defective empathy as found in persons with narcissistic personality disorder and in others who are very self-absorbed—a failure of insight
- ▶ Significant prejudice against the interests of those outside one’s group of identification, a failure of insight or motivation

- ▶ An inability to focus on unpleasant realities (eg, starving children, the abuse of women, the worst conditions of factory farms) that all reasonable people can agree are morally problematic—a failure of insight
- ▶ Weak will or susceptibility to temptation, a failure of motivation
- ▶ Impulsivity in relation to violence, a failure of motivation
- ▶ Unwillingness to find common ground when failure to compromise is disadvantageous to all, a failure of motivation
- ▶ Inability to find creative solutions to difficult problems involving competing interests and values, a failure of insight
- ▶ Inability to grasp subtle, complicated details that are of undeniable moral relevance (eg, the ways in which affluent persons benefit economically from the legacies of colonialism and slavery and from current injustices such as treaties with dictators or strongmen who disserve their country-people<sup>xi</sup>), a failure of insight

All reasonable persons can agree that the items on this list represent moral defects. Disagreements on the content of moral improvement do not include (reasonable) disagreement on these matters. Thus, public policy could responsibly consider supporting MB that addresses the aforementioned defects.

Meanwhile, competent adults and parents making decisions for children could go somewhat further in the direction of a specific (reasonable) moral vision in selecting forms of MB. For example, progressive individuals might want to select for themselves or their children greater compassion for the needy than some libertarians would accept as optimal. Conservatives might want to select an extraordinary sense of individual responsibility. Both might want to choose MBs that enhance one’s disposition to fairness, but they might have different understandings of what fairness requires in some domains (eg, economic justice) even as they agree in other domains (eg, adhering to the terms of an explicit contract that was freely entered into). Such selection of MBs in accordance with particular visions of morality is appropriate so long as the vision of morality is reasonable.

These remarks, I suggest, constitute an adequate general reply to the challenge of moral pluralism.

## WOULD MORAL BIOENHANCEMENT BE DESIRABLE?

### Setting up the questions

Suppose, somewhat fancifully, that research into particular types of MB demonstrated that they were quite safe and effective. That would eliminate some leading reasons to doubt the value of at least some types of MB. Suppose also that the relevant types of MB were not very expensive and the state made them available to anyone who could not afford them. That would address concerns about distributive justice. Let’s also suppose that efforts to improve the efficacy of traditional means of moral enhancement have proved disappointing. Suppose, further, that there are no realistic grounds to fear long term effects of these forms of MB on society (eg, the emergence of a dangerous gulf between the morally enhanced and the unenhanced). I make all these idealising assumptions in order to isolate our question: Would MB, against the background of those assumptions, be morally desirable just in terms of what it involves? In effect, the question is whether there is anything *inherently* wrong with MB. My thesis is that there is nothing inherently wrong with MB and that, against the background of

<sup>x</sup>Cf. reference 5, p 231.

<sup>xi</sup>For an informative discussion of this particular point, see reference 19, pp 30–51.

our idealising assumptions, the relevant types of MB would be desirable. I also contend that in the real world right now, in which those idealising assumptions are not true, the state should support research into MB and allow private companies (properly regulated) to do the same. We should at least learn more about the potential and the risks of MB.

Many scholars doubt what I assert: that there is nothing inherently wrong with MB. Some doubt this on the basis of a conviction that there is something inherently wrong with biomedical enhancement technologies in general. Chief among their objections are the charges that (1) biomedical enhancement is unnatural, (2) use of biomedical enhancements evinces an insufficient appreciation for human 'giftedness', and (3) biomedical enhancements pose a threat to personal identity. Elsewhere I have attempted to neutralise these objections.<sup>xii</sup> Here I will address a set of concerns that are directed at MB in particular and appeal to the nature and value of *human freedom*.

Articulating these concerns, John Harris writes:

The space between knowing the good and doing the good is a region entirely inhabited by freedom. Knowledge of the good is sufficiency to have stood, but freedom to fall is all. Without the freedom to fall, good cannot be a choice; and freedom disappears and along with it virtue. (p. 104)<sup>7</sup>

Importantly, this statement pertains to motivational improvement but not improved insight (which assists in 'knowing the good'). Because Harris is not addressing improved insight, and because it is quite implausible to hold that the latter poses a threat to freedom, *I will construe Harris' argument and similar arguments as directed entirely at motivation-based MB*—though I will hereafter omit the qualifier, 'motivation-based.' (Certainly, these arguments do not apply to embryo selection, which cannot alter—and therefore cannot affect the freedom of—particular individuals).

According to Harris, MB would eliminate the freedom to choose to do wrong—to 'fall'—and therefore the freedom to choose rightly. With freedom, he thinks, goes virtue. His position assumes that virtue is intrinsically, not just instrumentally, valuable. Perhaps that is correct.

What, more precisely, is the objection? Harris' discussion seems to suggest that MB would *entirely* eliminate freedom. This seems exaggerated, so let's construe the objection as asserting that MB would eliminate *some significant measure* of our freedom and thereby of virtue—and that this loss would be unacceptable.

How valuable is freedom? Harris asserts that 'sufficiency to stand is worthless, literally morally bankrupt, without freedom to fall. I, like so many others, would not wish to sacrifice freedom for survival' (p 110).<sup>7</sup> He holds that moral instruction and socialisation are the best means of moral improvement. These traditional means of moral enhancement, the thinking goes, do not significantly diminish an agent's freedom whereas MB would do so. This assumes, plausibly, that the traditional means can improve motivation and not merely insight.

Harris' discussion provokes two questions. (1) To what extent would MB threaten our freedom, including the freedom to act wrongly? Answering this question requires an exploration of the nature of freedom. (2) How valuable is freedom, including the freedom to do wrong? This question prompts us to consider what we value, or should value, in moral behaviour.

## On the nature of freedom

What is human freedom? According to one conception, each human agent is an 'unmoved mover' who is capable, in acting freely, of injecting causality into the natural order of causes and events without her action being determined by that causal order. As Chisholm puts it, '(i)n doing what we do, we cause certain events to happen, and nothing—or no one—causes us to cause those events to happen.'<sup>29</sup> <sup>xiii</sup> A free agent can sometimes exercise her will in such a way that is independent of causal influences that bear upon her. She may be hungry and want to eat, for example, but may decide to fast for some reason that she takes to be compelling—perhaps a belief that she should save the only available food for a hungrier person. Importantly, according to this view, if she freely chooses not to eat, her decision is *not* causally determined by psychological forces such as a desire to appear admirable, a need to shore up her self-image or even an ambition to live up to a moral ideal. On this view, reasons to act in a particular way and causes of behaviour are distinct sorts of things. The free agent sometimes manages, in effect, to step out of the causal order and act on the basis of a reason that does not cause her choice and action. She is thus an 'unmoved mover'.

This classical conception of human freedom accommodates the widely held intuition that we cannot be free if anything causes, or determines, us to choose and act as we do. It therefore embraces 'incompatibilism'—the thesis that freedom and universal causal determinism are incompatible—plus the belief that freedom is a reality. (Other incompatibilists are 'hard determinists', holding that there is no such thing as genuine freedom). Many philosophers have doubted that this view is even coherent.<sup>xiv</sup> After all, if the agent's choice is not causally determined by *anything*, as the present view requires of free action, it is debatable whether the choice is more than a random occurrence. Put another way, it is unclear whether the action can be properly said to belong to the agent, to be *his*, as an action must be in order to manifest his freedom. But we can set such doubts aside. For if freedom involves such radical independence from the causal order, then MB, as part of the causal order, can't touch it. In that case, MB can't possibly constitute a threat to freedom (p 5).<sup>31</sup>

Suppose instead that we are free in a sense that is compatible with universal causal determinism (as 'compatibilists' claim). In this more Humean spirit, we might say, roughly, that we are free when we determine our actions through our own will, or when we are able to do what we want.<sup>32</sup> I abstain from eating after dinner, say, not because someone or something is blocking my access to food, nor because anorexia dominates my eating choices, but because I want to maintain a healthful diet and see the avoidance of after-dinner snacks as conducive to this goal. I want to abstain and, in the absence of external or internal constraints, I do what I want. Unless there is a very special story, such as one about hypnosis, it appears that we may judge that I act freely in this case. Is there any reason to think that MB would threaten freedom under this conception?

It would seem not. Consider examples. As a result of MB, I might become (1) stronger-willed and therefore less vulnerable to weakness of will, (2) less prone to acting on violent impulses, (3) more prone to act altruistically or (4) more inclined to act fairly. Let us assume, reasonably, that the effects of MB are not so overpowering as to cause what amounts to temporary

<sup>xiii</sup>If I understand Kant correctly, it was his view that we *might* be free of naturally determining causes in this way and that morality was intelligible only on this assumption (reference 26, chapter 3).

<sup>xiv</sup>For a classic statement, see reference 30.

<sup>xii</sup>See references 27 and 28, chapter 6, and reference 16, chapters 3 and 4.

paralysis (say, when I feel a violent impulse) or involuntary muscle movements (say, that cause me to help someone). The agent, that is, *behaves as an agent*: performing intentional actions that are under voluntary control. It is compatible with performing such actions that one is influenced by any number of external or internal influences. Otherwise, since we always act under influences, we would never act freely on the present conception of freedom! So suppose that I render help to someone in need in a case where it is inconvenient to do so yet wrong not to help; and MB made me more inclined to help in such a case. The ‘boost’ provided by MB does not rob my behaviour of freedom any more than the caffeine in my tea robs me of any personal credit for writing this paper. Indeed, my choice to drink tea to enhance my ability to think and write is itself an expression of my agency. The same may be said of moral behaviour that is supported by motivations that are in turn supported by MB. I help the person in need because I want to help that fellow, or because I think I ought to, not because I am forced to do so. Moreover, I sought the motivational assist that I received from MB. There is no threat to freedom here.

‘But wait’, one might interject. ‘You are free in helping that fellow only if you could have *not* helped him. One acts freely only if one *could have acted otherwise*.<sup>xv</sup> Because of MB’s influence, this condition might not have been satisfied here’. For several reasons, I disagree.

First, suppose scepticism were warranted about whether I was free—more plausibly, *sufficiently* free—on the grounds that MB made me less likely to choose wrongly. If so, then parallel scepticism would be warranted about whether one acts sufficiently freely under the influence of such traditional means of moral enhancement as moral instruction or appropriate socialisation. One might reply that the two cases are not parallel: Traditional means of moral enhancement appeal to *reasons* for moral conduct whereas MB simply exerts causal influence on my motivations. But this oversimplifies matters. Traditional means often do exert psychological and therefore causal influence, yet unless truly excessive, this influence is compatible with the agent’s acting freely. In some cases, it is true, there may be a difference: A relatively unpressured moral education may appeal to reasons, with little or no psychological pressure, whereas a form of MB may directly (causally) affect one’s motivation. But, as we have seen, acting freely is compatible with acting under a degree of influence. Moreover, in the case of competent adults deciding to take MB, they choose the means to improved motivation—just as they might choose to go to a consciousness-raising group that applies a certain amount of psychological pressure. Even if some significant causal work is done in either case by something outside of one’s agency, one is not a passive subject of this influence but instead actively welcomes it.

On the compatibilist conception of freedom, acting freely requires not perfect independence from causal influence—which is impossible on this conception—but the right sort of relationship between one’s preferences and action as well as the absence of certain sorts of freedom-subverting conditions. Elsewhere, I have defended the following analysis of free action (which I called ‘autonomous action’): ‘*A autonomously performs intentional action X if and only if (1) A does X because she prefers to do X, (2) A has this preference because she (at least dispositionally) identifies*

*with and prefers to have it, and (3) this identification has not resulted primarily from influences that A would, on careful reflection, consider alienating*’. (p 102).<sup>28</sup> I maintain that any action that meets these conditions is genuinely free. Yet there is no reason to suppose that the moral behaviour of those who have undergone MB would necessarily or routinely fail these criteria. In the case in which I render assistance where doing so is inconvenient but morally required, I would not consider the influence of MB on my motivation to be alienating. I welcomed this influence.

What about those whose moral behaviour is motivationally influenced by MB that was ‘imposed’ on them by their parents? They, too, have no special reason to consider the influence alienating. After all, it is helping them—without forcing them—to act as they should. This, like moral instruction that comes with a certain amount of psychological pressure, is within the bounds of appropriate parenting (notwithstanding the fact that some people don’t want to do what they should). Because the children are not choosing the interventions for themselves, there is a greater possibility of alienation than in the case of consenting adults, but no more than in the case of many unremarkable instances of parents’ ‘imposing’ traditional means of moral enhancement on their children.

My sceptical interlocutor might insist that in *some* cases of MB an agent would lack ‘freedom to fall.’ This may be true. But it is also true of many everyday behaviours whose moral value we do not question. For example, I am psychologically incapable of stabbing or shooting a person I perceive to be innocent and non-threatening. In the presence of a handy weapon and a potential victim, I rightly abstain from immoral assault—and it seems that I could not act otherwise and commit the assault. But my inability here is not due to squeamishness, for I could probably stab or shoot someone who was trying to kill a loved one or an innocent person if no less violent means of protection were available. My inability to stab or shoot a loved one or innocent person is due to my stable values and preferences. I freely abstain from immoral assault even though I cannot act otherwise. I meet the conditions of autonomy stated above. The same may well be true, depending on details, where I render assistance for moral reasons and am psychologically incapable of acting otherwise. Ability to act otherwise, or freedom to fall, is not a necessary condition of free action.<sup>33 34</sup>

Nevertheless, some conceivable forms of MB *would* pose a threat to freedom. Imagine a computer chip that could be implanted in someone’s brain such that whenever the agent decided to perform a certain kind of immoral action, he would change his mind.<sup>xvi</sup> The implant would reliably, automatically redirect his decision-making. It seems plausible to assert that this sort of MB would make an agent unfree with respect to certain actions (even if the agent welcomed the MB), and possibly in motivation-based ways—say, if it made the agent suddenly lose the desire to perform the action in question. But most of the types of MB under consideration, including the examples enumerated earlier, do not have this apparently freedom-robbing quality.

There is nothing about MB in general that poses a significant, across-the-board threat to freedom when it acts on an agent’s motivations. Nor, as explained earlier, does MB threaten freedom when it acts on an agent’s moral insight. Even more obviously, MB poses no threat to freedom in cases involving embryo selection. Except in extreme cases involving motivational improvement, MB does not pose a significant threat to freedom.

<sup>xv</sup>While the insistence that someone might have acted otherwise than she did may seem to require a denial of universal causal determinism, there are compatibilist interpretations of ‘could have acted otherwise’ (see, eg, reference 30).

<sup>xvi</sup>This idea was stimulated by a thought-experiment in reference 35, section 4.4.

## On the value of freedom

Suppose matters were otherwise and MB did pose a significant, across-the-board threat to freedom. What would follow about the desirability of MB? Presumably, freedom has significant value. It doesn't follow that its value overrides everything else that matters.

One can choose and act *more or less* freely, depending on the extent and source of influences and pressures that bear on one. In some cases, they might be a bit alienating but not so alienating as to make our choice forced or unduly pressured. That is, we act freely—we meet some appropriate threshold to judge that our action is free (enough)—but our degree of freedom is far short of what characterises our actions when we act *most* freely. I see no reason to preclude tradeoffs between the degree of freedom with which we act and other values.

Suppose, contrary to our discussion in the previous subsection, that MB consistently had the effect of substantially reducing our freedom of action. Imagine that there was a reasonable way of quantifying this effect and that, in comparison with the average degree of freedom with which we act, MB reduced our freedom by 25%. Imagine further that, as a result of MB, there were no more wars or starvation and everyone in the world had access to the basic necessities of life. I, for one, would accept this reduction in freedom across the board if it were necessary for such great improvement in people's moral behaviour and such welcome consequences.

Alternatively, suppose that while most forms of MB pose no significant threat to freedom, a particular one does. A brain-implanted computer chip reliably causes the agent to change his mind whenever he forms an intention to rape someone or molest a child. It robs the agent of 100% of his freedom with respect to rape and child molestation without diminishing his freedom to perform other types of action and without producing any deleterious side effects. In my view, the elimination of freedom with respect to these types of action is no great loss. Indeed, in terms of freedom alone, the trade-off might be worthwhile to the agent in question, by enabling him to avoid the loss of freedom imposed by incarceration.

If you agree with the assertions made in either or both of the two previous paragraphs, then you accept the proposition that maximal freedom of action does not automatically deserve priority over other values. Thus, even if MB significantly reduced our freedom, it would not straightforwardly follow that MB was undesirable. It might yet be desirable in view of its good results. This brings us to a fundamental question.

## WHAT SHOULD WE VALUE IN MORAL CONDUCT?

The anatomy of moral improvement discussed earlier presupposes an anatomy of moral *conduct*:

Moral motivation + moral insight → moral behaviour.

That is, being morally well-motivated and having moral insight conduce to moral behaviour.

Our investigation suggests the possibility of another factor that should be taken into account: freedom. According to some critics of MB such as Harris, one might have moral motivation and insight without freedom (or with insufficient freedom); and freedom gives moral behaviour much of its value. This suggests the following analysis:

Freedom + moral motivation + moral insight  
→ genuinely valuable moral behaviour.

In other words, freedom (or sufficient freedom) in combination with strong moral motivation and moral insight conduces to moral behaviour that has special value due to the relationship between that behaviour and the agent's agency.

I have conceded that there is something right about the amended formula. But we should not exaggerate the value of freedom. After all, moral behaviour itself, the end product, is also extremely important—independently of how free it is. Admittedly, I am asserting this claim, rather than arguing for it, in the confidence that most readers will agree with me. In any event, this claim coheres with the idea that the end product of moral behaviour is, as far as human agency goes, what we are mostly after in wanting a better world with better lives for human beings and other sentient creatures. In the absence of a deity who will give us this better world, it is up to us human agents to attain it. Without a substantial improvement in moral behaviour, we are highly unlikely to do so; indeed, there is a good chance that things will get much worse due to our growing destructive power (as discussed in the section entitled 'Motivating Moral Bioenhancement'). So, while freedom matters, it is—to understate the point—not all that matters in connection with our moral agency. What we do and the expected consequences of what we do are also extremely important.

At the same time, I have challenged the idea that MB poses any systematic threat to freedom. Except in extreme cases, we can retain our valuable freedom while choosing to use MB as a means to moral improvement in seeking a better world. There is nothing inherently wrong with MB. In view of what is at stake, we should open-mindedly consider this non-traditional means of moral enhancement.

**Competing interests** None.

**Provenance and peer review** Commissioned; externally peer reviewed.

## REFERENCES

- 1 Parens E, ed. *Enhancing human traits*. Washington, DC: Georgetown University Press, 1998.
- 2 Savulescu J, Bostrom N, eds. *Human enhancement*. Oxford: Oxford University Press, 2009.
- 3 Juengst E. What does 'enhancement' mean? In: Parens E, ed. *Enhancing human traits*. Washington, DC: Georgetown University Press, 1998:25–43.
- 4 Buchanan A. *Beyond humanity?* New York: Oxford University Press, 2011.
- 5 Douglas T. Moral enhancement. *J Appl Philos* 2008;25:228–45.
- 6 Persson I, Savulescu J. The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity. *J Appl Philos* 2008;25:162–77.
- 7 Harris J. Moral enhancement and freedom. *Bioethics* 2011;25:102–11.
- 8 Gaillot M, Baumeister RF, DeWall CN, et al. Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *J Pers Soc Psychol* 2007;92:325–36.
- 9 Crockett M, Clark L, Hauser MD, et al. Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *PNAS* 2010;107:17433–8.
- 10 Terbeck S, Kahane G, McTavish S, et al. Propranolol reduces implicit negative racial bias. *Psychopharmacology (Berl)* 2012;222:419–24.
- 11 Franzini A, Marras C, Ferrol P, et al. Stimulation of the posterior hypothalamus for medically intractable impulsive and violent behavior. *Stereotact Funct Neurosurg* 2005;83:63–6.
- 12 Sitaram R, Caria A, Veit R, et al. fMRI brain-computer interface: a tool for neuroscientific research and treatment. *Comput Intell Neurosci* 2007; 2007:25487. doi: 10.1155/2007/25487
- 13 Caria A, Sitaram R, Veit R, et al. Volitional control of anterior insula activity modulates the response to aversive stimuli. A realtime functional magnetic resonance imaging study. *Biol Psychiatr* 2010;68:425–32.
- 14 Reuter M, Clemens F, Walter NT, et al. Investigating the genetic basis of altruism: the role of the COMT Val158Met polymorphism. *Soc Cogn Affect Neurosci* 2011;6:662–8.
- 15 Eley T, Lichtenstein P, Stevenson J. Sex differences in the etiology of aggressive and nonaggressive antisocial behavior: results from two twin studies. *Child Dev* 1999;70:155–68.

- 16 DeGrazia D, *Creation ethics: reproduction, genetics, and quality of life*. New York: Oxford University Press, 2012.
- 17 Silver L. *Remaking eden*. New York: Avon, 1997.
- 18 Glover J. *Humanity*. New Haven: Yale University Press, 2000.
- 19 Pogge T. Severe poverty as a human rights violation. In: Pogge T, ed. *Freedom from poverty as a human right*. Oxford: Oxford University Press, 2007:11–53.
- 20 United Nations Development Report, Human Development Report 2005. New York: UNDP, 2005.
- 21 United Nations Development Report, Human Development Report 1998. New York: Oxford University Press, 1998.
- 22 Pogge T, Two reflections on the first United Nations Millennium Development goal. *J Hum Dev* 2004;5:385–92.
- 23 United Nations Development Report, Human Development Report 2002. New York: Oxford University Press, 2002.
- 24 Damasio A. *Descartes' error: emotion, reason, and the human brain*. New York: Putnam, 1994.
- 25 Aristotle. *Nicomachean Ethics*.
- 26 Kant I, *Grounding of the Metaphysics of Morals*. 1785.
- 27 DeGrazia D, Enhancement technologies and human identity. *J Med Philos* 2005;30:261–83.
- 28 DeGrazia D, *Human identity and bioethics*. Cambridge: Cambridge University Press, 2005.
- 29 Chisholm R, *The lindley lecture*. Lawrence, KS: Department of Philosophy, University of Kansas, 1964:3–15.
- 30 Ayer AJ. *Philosophical essays*. London: Macmillan, 1954, chap. 12: Freedom and Necessity.
- 31 Persson I, Savulescu J, Getting moral enhancement right: the desirability of moral bioenhancement. *Bioethics* 2013;27:124–31.
- 32 Hume D. Enquiry Concerning Human Understanding. Chapter VIII: Of Liberty and Necessity. 1748.
- 33 Frankfurt H, Alternate possibilities and moral responsibility. *J Philos* 1969;66:829–39.
- 34 Wolf S. Asymmetrical freedom. *J Philos* 1980;77:151–66.
- 35 Savulescu J, Douglas T, Persson I. Autonomy and the Ethics of Biological Behaviour Modification (forthcoming in Akira Akabayashi (ed.), *Toward Bioethics in 2050: International Dialogues*).