Brassington, Iain. "Good Is as Good Does? The Case of 'Moral Enhancement'." *Bioscience and the Good Life*. London: Bloomsbury Academic, 2013. 137–164. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781472544315.ch-007>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 13:18 UTC.

Copyright © Iain Brassington 2013. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

Good Is as Good Does? The Case of ‘Moral Enhancement’

The possibility of ‘moral enhancement’

Over the past couple of chapters, I’ve tried to outline a position that is sceptical of the ability of bioscientific advances to make all-things-considered good enough lives enormously better – but none of this has amounted to hostility. However, in respect of one potential and putative enhancement, I think that there is room to add to a general scepticism a more positive reason to reject methods to alter human capacities: this is in relation to so-called ‘moral enhancement’.

Very broadly, the term ‘moral enhancement’ capitalizes on our increased understanding of the neurobiological factors that determine behaviour, and puts forward the idea that there could be interventions – most likely, by altering the brain or some other determinant of behaviour – that would make a person more moral: that is, more disposed to do the right thing. In this vein, Tom Douglas’ formula for moral enhancement is that

[a] person morally enhances herself if she alters herself in any way that may reasonably be expected to result in her having morally better future motives, taken in sum, than she would otherwise have had.¹

This account may well invite criticism – but it’ll do for now.

Of course, there is a great deal to be said for moral enhancement taken in the broadest sense: parents teaching good behaviour to their children are engaged in what they at least hope will be a form of moral enhancement – and it’s expected that they ought to be. And one aspect of being a moral agent at all is the expectation that we should seek to restrain our baser instincts and make ever finer distinctions when they are there to be made; this, too, is a kind of moral enhancement. For example: imagine that Bob has been brought up to dislike and distrust members of a particular ethnic group, and absorbs these attitudes with his mother’s milk. Bob leaves home and goes to university, where he finds himself forced to associate with someone from that ethnic group, and it begins to dawn on Bob that this person really isn’t as bad as all that; by the end of his first term, Bob has been forced to re-evaluate his attitudes. But it’s hard to shake years of conditioning; and so while he is now more willing to accept the baselessness of his

old attitudes rationally, there is still a slight residue – an involuntary half-second reluctance to trust members of that group, say. However, Bob knows that this residual feeling is baseless and unjust, and tries to suppress it – to teach himself to react differently. This is, I'd contend, a kind of moral self-enhancement; and it would be not only admirable, but quite possibly mandatory.

So it would be absurd to present a case that moral enhancement is necessarily wrong. But there are still problems, as we shall see over the coming pages; and they might well be quite big.

Strategies for moral enhancement

A paper by Tom Douglas in the *Journal of Applied Philosophy* from 2008, which is fast becoming the *locus classicus* in respect of moral enhancement, considers the case for intervening to ensure that people have better motives than they otherwise might have had. In Douglas' scenario, the intervention is by means of a pill, but we could imagine other methods such as direct brain stimulation of one sort or another; the point is that a person who previously had a propensity to behave in *this* way could undergo some kind of procedure that would increase the likelihood of his behaving in some other – implicitly more desirable – way. More, Douglas claims that there are some 'counter-moral emotions'² that would always be (though maybe 'provide' would be a better word) uncontroversially bad motives for action, and the attenuation of which would reduce their ability to interfere with good motives.

One example that he gives of a counter-moral emotion – both bad in itself, and likely to interfere with good motives that a person may have – is a strong aversion to certain racial groups. Another is a strong impulse towards aggression, which may occasionally be in order, but is likely in most cases to be a morally bad motive. But these examples are meant to be illustrative more than definitive; what counts is that

there are some emotions such that a reduction in the degree to which an agent experiences those emotions would, under some circumstances, constitute a moral enhancement.³

Finally, he hypothesizes, inducing a reduction in the degree to which a person experiences those emotions is not impossible in principle. So, if we can control the expression of (say) a gene that is strongly linked to aggression, or make some difference to an agent's neurochemical balance such as to make him more likely to behave in a desirable way – it's widely believed that oxytocin, for example, can modulate behaviour to make it much more pro-social – then we have at our disposal a method of making people better. The details need not bother us all that much: what counts is the basic idea of intervening directly to make a person 'better behaved'.

One immediate worry here might be that we're embarking on some kind of *Clockwork Orange* scenario in which a rich moral life is displaced by conditioned good behaviour, with the implication that 'enhanced' goodness is no such thing. I'll return to a version of this worry later; but the objection may be a little simplistic in these terms. Importantly, Douglas talks a lot in the paper about a person enhancing *himself* – and, indeed, one possible use for moral enhancement of the sort that appears to be envisaged here is to help people like Bob in the example above control the emotional responses to certain situations that they would prefer not to have. It would still, in the end, be up to him to try to cultivate the right attitudes – it's just that the task would be easier; and, to reprise the Real Tomato Argument from Chapter 5, taking a pill need not mean that the results were void of moral merit.

Douglas' emphasis is on interventions on currently existing people; but we could also intervene to ensure that people have more desirable moral dispositions by making more positive decisions about who gets to be born in the first place. In a 2006 paper, Julian Savulescu et al. consider the relative merits of eugenic selection for desirable behaviour.⁴ Like Douglas', their argument relies on the premise that at least a significant factor in the determination of moral behaviour can be naturalized – that is, that it can be attributed to an empirically scrutable fact about a person's physical make-up. They begin their argument by pointing out that there is a credible genetic factor in at least some morally problematic behaviour. For example, there is evidence that a genetic abnormality that leads to low levels of the enzyme monoamine oxidase A (MAOA) can be correlated to anti-social behaviour – and that this suggests that it ought to be possible to do something to reduce the levels of such behaviour. They conclude that there is a sound moral reason to select against embryos that carry the low MAOA mutation.

Though their preferred strategy is to select embryos rather than attempting to alter particular genes or the expression of particular neurotransmitters within a given embryo or person – they prefer genetic selection to genetic enhancement in this case – I think that it is still fair for the sake of this chapter's concerns to place embryo selection of this sort at least broadly within the scope of arguments for moral enhancement. They are still talking about interventions that draw on bioscientific insights and are expressly intended to improve behaviour; this is close enough to moral enhancement to make no real difference for my purposes.

Such enhancement – of the sort envisaged by both Douglas and Savulescu et al. – could be presented as serving entirely the interests of the society in which the embryo would come to birth, since it is reasonably safe to assume that society has an interest in at least most of its members being less aggressive. This impression might be furthered by the way that Ingmar Persson and Julian Savulescu talk about the need for moral enhancement in their 2012 book *Unfit for the Future*. They argue that the big problems that face humanity in the

decades and centuries to come and threaten the very existence of the species, or (at best) the welfare of its members – notably, climate change, but also other threats that arise from the fact that it is very much easier to kill great numbers of people by action or inaction than it is to save fewer – are problems because we have a systematic bias in favour of those who are spatially and temporally close to us, and because we have a bias to treat moral responsibility narrowly, in terms of causation.⁵ They are problems that – it is hypothesized – will not be solved by technology alone, not least because there is currently little incentive to develop it. Understanding the neurological and genetic reasons for our behaviour, though, would mean that we would be in a much better position to do something about it, and to amplify and solidify ‘the internalisation of moral doctrines’ by tailoring our genes and neurology towards the ends required by morality properly and widely understood:

We call moral enhancement by such means *moral bioenhancement*; possible examples of moral bioenhancement would be drug treatment and genetic engineering.⁶

What is meant by moral enhancement here is not reducible to making people more pro-social (although increased sympathy would come in to it), but about making them better at being good. Oxytocin is one suggested method; but the precise method isn’t one that is all that important.

However, such interventions may turn out to serve the interests of the enhanced person and person-to-be just as much as the people around whom they live:

Given that compulsive violent or criminal behaviours are typically punished with incarceration as well as social rejection, we can assume that these behaviours typically harm a person who expresses them, as well as their victims. Selections which constrain a person from making such harmful choices may therefore be justified.⁷

As such, there is a case to be made for it being the sort of thing that may make the enhancee’s life better, at least indirectly. And so the case for bioscientific moral enhancement seems to be fairly straightforward: making people morally better is desirable just insofar as it is morally desirable to be morally good and to promote the good; it is desirable because those of us who have inclinations that we would rather not have might thereby have a strategy to cultivate our own virtue more easily; it is desirable because, if it is possible to reduce the frequency and likelihood of morally undesirable behaviour, that will probably mean fewer criminal and harmful actions, which means that there is a good chance that less harm will accrue to fewer victims; and it is desirable because fewer people will have to suffer incarceration, social resentment, and the other harms that follow from the misfortune of having a disposition towards aggression.

The argument from freedom

There is a range of problems that might be raised for proponents of moral enhancement – and possibly with moral enhancement itself. John Harris has argued against moral enhancement on a couple of grounds. One of these is that avoiding the potential problems of human behavioural blundering requires not so much *moral* as *cognitive* enhancement – or, perhaps, that cognitive enhancement will allow us to avoid moral disaster much more effectively than ‘moral enhancement’ as outlined above would. I’ll return to this theme in a little while. Another objection, though, is that moral enhancement programmes are antagonistic to other desirable characteristics – notably freedom.

It’s a concern with freedom that brings Harris to see moral enhancement as a source of positive suspicion. For him, citing Milton, ‘sufficiency to stand is worthless, literally morally bankrupt, without freedom to fall’⁸; he claims that

ethical expertise is not “being better at being good”, rather it is being better at knowing the good and understanding what is likely to conduce to the good. 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 along with virtue. There is no virtue in doing what you must.⁹

and that

the sorts of traits that seem to lead us to wickedness or immorality are also the very same one required not only for virtue but for any sort of moral life at all.¹⁰

Harris does not seem to be denying that the targets that people like Douglas or Savulescu or Persson have in mind are anything other than worth pursuing – or even urgent. But what is at stake is the specifically *moral* aspect of achieving them. Grant that the manipulations envisaged will make people behave in the manner that morality would dictate: there would be no actual *morality* on display, since agency itself would be diminished; we would have moved towards the mere playing-out of a pre-determined set of actions that would not indicate any particular praiseworthiness on the part of the people carrying them out. Those people’s ‘moral’ capacities would not be enhanced at all – we would not see any improvement in morality richly understood. (The distinction is rather similar to that drawn by Kant between acting *from* duty, and acting merely *in accordance with* it.¹¹)

Are these worries sufficient to undermine arguments for moral enhancement? In the 2006 paper mentioned above, Savulescu et al. appear to be prepared at least to acknowledge their coherence: they admit that appeals to the importance of an ‘open future’ may generate an argument against their behavioural genetic selection – although they also (rightly, in my opinion) note that open future arguments are not powerful.¹² But in *Unfit for the Future*, Persson and Savulescu

take issue with the idea that being biologically determined to have certain dispositions makes any difference to freedom. For example, suppose it is true that women are, by nature, likely to be more altruistic and less aggressive than men because of something to do with biology. (Cordelia Fine points out that the evidence backing up this claim is, at best, shaky¹³ – but we can let that pass for the sake of the argument.) Well, even so, this does not mean that women are less free than men. If moral behaviour is straightforwardly neurophysically¹⁴ determined, then all moral enhancement would do is tip the balance in favour of us doing praiseworthy things. We would be no more or less free than otherwise. Conversely, if moral behaviour is not neurophysically determined,

[t]hen moral bioenhancement cannot be fully effective because its effectiveness is limited by the indeterministic freedom that we possess. So, irrespective of whether causal determinism or indeterminism reigns in the field of human action, moral bioenhancement will not curtail human freedom and responsibility. Biomedical manipulation cannot chance the basic laws of our behaviour by making us more (or less) causally determined . . .¹⁵

While this might answer some of Harris' qualms, his view of freedom still seems to be one in which freedom is magnified the more options that one has; hence to have one's brain manipulated in such a way as to make it much less likely that (say) one will become violent is to reduce freedom. But this prompts an important question: Is it true that the scope of someone's freedom really is related – at least directly – to the number of options open to her? If it is not, removing undesirable options would not necessarily be a constraint on freedom or moral action richly understood; and I'll argue in the coming paragraphs that reducing the number of options is *not* a constraint on morality or moral action richly understood – and that the 'freedom to fall' is therefore inessential to moral freedom *per se*. In fact, I think that having fewer options can be, morally, a good thing: and in a sense, it's what moral education is all about.

Freedom and options

Suppose you are walking along a quiet street one night, and notice that ahead of you is someone rather substantially the worse for wear from drink. Strictly speaking, one of the things that you could do in this situation would be to attack and rob that person. There is nothing really stopping you behaving in this way. However, things can be *formal* possibilities without being *live* possibilities. A reasonably decent person would not see the option of attacking the stranger as a live possibility; it's simply not the kind of thing that appears in the suite of plausible things that could happen. This is not just because of worries about getting caught and preferring not to be punished: pertinent as those worries might be from one perspective, the truly morally developed person doesn't

even have to make that kind of calculation. Attacking people on the street is, in everyday parlance, unthinkable. To borrow a phrase from Bernard Williams, someone who has to think about whether attacking strangers on the street is permissible, weighing up the reasons to act against those not to, has had – and even if he decides not to act – ‘one thought too many’¹⁶; and this is what I mean by the idea that there are certain formal possibilities for action that are not, in a morally well-constituted subject, real or live possibilities.¹⁷

Though writing in a different context, John McDowell makes the point that in moral upbringing, ‘one learns not to behave in conformity with rules of conduct, but to see situations in a special light, as constituting reasons for acting’¹⁸ – or, as in the example above, not acting. Thus, in an important passage, he claims that

the attractions of whatever wickedness might bring do not constitute some reason for wickedness, which is, however, overridden by the reasons against it; rather, given that they are achieved by wickedness, those attractive outcomes do not count as reasons at all.¹⁹

Of course, street attacks – and worse – do happen; but while we condemn them and the sort of person who carries them out, I think it reasonable to say that it would not be unjust to think that a person who has to work out that and why she should not attack people on the street is morally lacking in some important way too. This is as much as to say as that a person with more live options for action is not always more morally praiseworthy than someone with fewer; having fewer options might, in fact, be a sign of being a complete moral being. David DeGrazia echoes this thought – satisfyingly enough, as part of a defence of moral bioenhancement – when he points out that ‘[a]bility to act otherwise, or freedom to fall, is not a necessary condition of free action’.²⁰

Neither is it obvious how the inability to see certain logically possible courses of action as live options makes us less free, any more than the inability to see red as green does. But even if, for the sake of the argument, we *are* less free in some sense, it does not obviously follow from that either that a diminution of freedom for the sake of better behaviour is undesirable (DeGrazia again: ‘Presumably, freedom has significant value. It doesn’t follow that its value overrides everything else that matters’²¹) or that our action has any less moral merit or that we are in any way less whole moral agents. After all: not attacking people in the street does not really attract a great deal of praise in any case – that’s simply a matter of behaving in a way that’s minimally acceptable, and it’s strange to think that our integrity as agents depends on there being a realistic chance of us doing what we ought not do and would not think of doing. I have any moral merit at all, it’s more bankrolled than bankrupted by my utter disinclination to go raping and pillaging my way around Manchester.

We might even want to take things a little further, beyond simply denying that having fewer live options is compatible with moral agency properly understood: it is not completely *outré* to wonder whether having fewer live options could itself count as morally enhancing, such that the less quick to anger, the less aggressive, and so on that we are – the less we are slaves to our passions – the freer we are. At least some elements of Kantianism are sensitive to this notion of being free from the passions – but it's a line of thought that goes back as far as Plato. Having certain base impulses purged by moral enhancement of one sort or another may make us *more*, not *less*, free. So the idea that morality properly understood is corroded by attempts to make it less likely that we will choose to behave badly seems to me not to have much merit.

(For the record, it's worth noting the oddity that Harris should be keen on preserving the 'freedom to fall' in this context, but seems less devoted to it in others. In relation of children participating in – or perhaps 'being participated in' would be a better turn of phrase – scientific research, he and Søren Holm suggest that

[o]ne of my main interests is my interest in taking myself seriously as a reflective moral agent, and my interest in being taken seriously by others. Identifying my moral obligations, and acting on them, is not contrary to my interests, but is an integral part of what makes me a moral agent. But what about children?

If children are moral agents, and most of them, except very young infants are, then they have both obligations and rights. It is, furthermore not a requirement for having a moral obligation that a person is able to give a full account of the moral justification for that obligation.

If a parent does not take this into account when making decisions for the child that parent displays one of the following attitudes; either the attitude that the child is not (and need not be) a serious moral agent at all, or the even more problematic attitude that the child is so deeply fallen in moral turpitude that it is not willing to discharge any of its moral obligations. The only way of taking the child seriously as a developing moral agent, interacting with other moral agents in a complex world, is to assume that the child is serious (or would wish to be serious) about discharging its moral obligations. Parents are therefore justified in assuming that the person they are making decisions for is a moral person, who wants to discharge his or her moral obligations.²²

(This appears to boil down to the idea that we take children's moral agency seriously by making certain decisions on their behalf – which seems to be a claim that we can and maybe should deny their freedom to fall *as a part of taking their agency and capacity for virtue seriously*. I do not find this argument at all convincing – and Barry Lyons' dismissal of the idea is quick and cutting, and explodes it in a matter of a few sentences²³ – but, still: it's tempting to wonder just how much freedom to fall is necessary even by Harrissian lights.)

Having said this, I think that Harris is arguing on the right lines; his position can be fortified, and that's a task to which I'll turn my attention now. The first step is to invoke the spirit of Aristotle.

Nicomachean moral enhancement

Aristotle is suspicious of the idea that one's moral characteristics are what they are by nature; he states at the start of Book II of the *Nicomachean Ethics* that moral goodness 'is the result of habit', which means that

none of the moral virtues is engendered in us by nature, since nothing that is what it is by nature can be made to behave differently by habituation.²⁴

Since a person's characteristic patterns of behaviour often can be changed, that means that the disposition to display certain or all virtues must be plastic. For Aristotle,

we become just by performing just acts, temperate by performing temperate ones, brave by performing brave ones . . . It is the way that we behave in our dealings with other people that makes us just or unjust, and the way that we behave in the face of danger, accustoming ourselves to be timid or confident, that makes us brave or cowardly . . . Hence we must give our activities a certain quality, because it is their characteristics that determine the resulting dispositions. So it is a matter of no little importance what sort of habits we form from the earliest age – it makes a vast difference, or rather all the difference in the world.²⁵

Habituation means that we give ourselves a 'second nature' that will determine how we act in everyday life; it is by emulating the best that one improves oneself as their habits and dispositions will become one's own. And in some cases, one doesn't even have to decide to emulate the best, the decision having been made by others:

Legislators make their citizens good by habituation; this is the intention of every legislator, and those who do not carry it out fail of their object.²⁶

The point is that Aristotle seems to have some account of a programme of becoming more virtuous that one could adopt or even impose on others. At least on the face of things, this looks as though it could be read as being a kind of claim about moral enhancement. Moreover, his account does not jar too badly with a number of modern intuitions about ethical formation; I'll come back to them in a little while.

In the Aristotelian picture, there is a meaningful parallel between acquiring the virtues – a disposition towards temperance, justice, and so on – and acquiring any other skills. Thus

the causes or means that bring about any form of excellence are the same as those that destroy it . . . for it is as a result of playing the harp that people become good or bad harpists. The same principle applies to builders and all other craftsmen. Men will become builders as a result of building well, and bad ones as a result of building badly. Otherwise there would be no need of anyone to teach [people]: they would all be *born* either good or bad. Now this holds good also of the virtues.²⁷

There is a slight oddity about the idea that one becomes a good builder by building well – we might ordinarily suppose that it's only good builders who are capable of building particularly well to begin with. But this is not a serious problem for my purposes – we could avoid it by rephrasing things slightly while retaining the gist – and his statement has a certain attraction to it. The point is that one would acquire practical skills by associating with skilled people and doing as they do, gradually becoming habituated to that particular skilled activity; likewise, one would acquire the virtues by associating with virtuous people and doing as they do, gradually becoming habituated to virtue oneself.

So let's accept that there is a mechanism by which one becomes virtuous or skilful, and take the example of a novice builder. He will be less-than-excellent to begin with; but we might reasonably hope that given a decent mentor and time, he will become excellent. If the development of his skills is retarded, he will remain a bad builder. The alternative makes little sense: being a bad builder is not a vocation in its own right; indeed, though one sets out to be a good builder, being a bad builder for a while is probably an integral part of realizing this ambition. Still, this does not mean that a person has any latent skill as a builder that habituation realizes. Before starting to build for the first time, a person is – to paraphrase Wolfgang Pauli – *not even bad*: not a builder at all.

It's tempting to give a similar account in respect of moral development. We tend not to blame children for bad actions, and this is because we think that, at least in some respects, they are not moral agents to begin with: they, too, are not even bad. The first stage of moral training involves the task of making a person – usually a child – capable of bearing moral epithets, just as the first stage of making someone a builder is to make him capable of bearing the appropriate epithets. At the same time, good behavioural patterns are instilled by socialization: getting children to write thank-you letters, for example, helps generate desirable patterns of behaviour that mean that courtesy comes naturally in later life, just as repeatedly building walls means that building them well soon enough becomes a matter of instinct.

The flip side of this is that Aristotle also says that there is a process by which one becomes a bad builder, harpist or whatever; and that it is the same as that by which one becomes a good one. If acquiring practical skills is truly analogous to acquiring the virtues, this would suggest that the manner in which someone becomes vicious is, in the first place, by habituation, and, in the second, like that in which someone becomes a bad builder.

The first of these ideas – that one can become vicious by habituation – seems reasonable enough, and seems to fit in respect of people who turn out to be moral failures; there're plenty of stories that we could tell in which a person seems to have become vicious at least partly because of some kind of malhabitation, and we do tend at least to be sympathetic to the idea that people can be habituated towards evil by a bad upbringing – which seems

to be the correlate of a belief that a good upbringing is a requirement for the formation of an admirable moral character. For example, the news in late 2009 and early 2010 carried the story of two young brothers, aged 10 and 11, tried for a brutal attack on two other boys of about the same age. The boys were found guilty, and such was the court's assessment of the danger they posed to others that they were detained for an indefinite period. However, the court heard how the defendants had been raised in what social workers called a 'toxic' environment at home – they had witnessed domestic violence, pornography, horror films, had access to alcohol, and had been fed cannabis by their mother.²⁸ Though this one rather lurid case is clearly not enough to sustain a whole argument, it is illustrative of the idea that behaviour is in at least some respect associated with upbringing – that a tendency towards violence is not too surprising in someone who is habituated to violence from an early age. This looks to be a good example of the process that would habituate or engender moral virtue being the same as that which would habituate or engender moral vice. The process that can create virtuous dispositions can also destroy them.

So far so good: the problems begin when we ask more deeply about the analogy between becoming vicious and becoming – say – a bad builder. The habituation thesis is definitely not a plausible account of the process involved in becoming a bad. If it's true that becoming a bad builder is a step on the way to becoming a good one, then it won't destroy the excellences associated with building well, or make it more difficult to become good – which is quite unlike moral development, where acquiring a vice is antagonistic to the corresponding virtue. Similarly, while it might be possible to desensitize a previously virtuous person to vice, or *vice versa*, it is not obvious that you can desensitize people who possess other practical excellences. For example, a virtuoso harpist may spend a great deal of time in the company of barely competent students, but while exposure to them may mean that she understands their mistakes, that will not destroy her excellence. To lose that excellence would require that she fall out of practice altogether. By contrast, whereas a virtuous person who is rarely presented with any particular requirement to exercise his virtue will not obviously be less good at doing so when the need arises, he may have his moral sentiments coarsened by spending a great deal of time with brigands. This means that the process of becoming a bad person is rather different from the process by which one becomes a bad musician, builder, or whatever.

It is possible to salvage the analogy, provided that we're prepared to accept that there are two distinct sides to Aristotle's idea, which Aristotle has mistakenly run together. One of these sets is 'functional', and relates to the ability to do well; the other is 'cognitive', and relates to the ability to judge well. (The ability to judge well may be required to do well, but it is not the same.) This distinction will turn out to have a role in thinking about biomedical moral enhancement, and in rebuilding something like an argument from freedom.

Rebuilding the argument from freedom

Let's return to the builder example. A good builder will be able to plaster a wall evenly, and will take proper pride in a job done well. He will also be able to judge what kind of plaster to use, how much water is required, and so on, and none of this is required actually to put the stuff on the wall. Similarly, a convincingly good agent will not only have a desire to do the just thing and a horror of doing the unjust; he will also have the entirely separate cognitive skills required to assess what justice demands.

Now: in his account of moral enhancement, Douglas is concerned explicitly about instilling better motives in people, and not so much about processes of moral reasoning. The concern appears to be attributable to the way that understanding of the phrase 'moral motive' has to do with the 'psychological – mental or neural – states or processes that will, given the absence of opposing motives, cause a person to act'.²⁹ To this extent, his paper speaks to the functional aspect of Aristotelian ethical formation – the bit that's analogous to spreading the plaster on the wall – more than the ability to judge what the right thing to do would be.

But we need to be careful with the language here. To be caused to act in a certain way is not the same as acting for a cause. Douglas is talking about the things that cause a given behaviour; the cause in this sense is whatever it is that explains and brings that behaviour about. But this is not the same as talking about the cause that a person may have; to use the word 'cause' in this sense is to talk about whatever it is that invites behaviour. A drone aircraft flying above enemy territory can be caused to drop a bomb; but it does not drop a bomb for the sake of any cause.

Put another way, he is talking about the reasons *why* an agent acts as he does, but not reasons *to* act. And the important thing is that the two things may come apart: a person's mental or neural states explain *why* he likes ice-cream, but do not furnish a reason *to* prefer it; facts about the evolutionary advantages of ascribing agency to observed phenomena may tell us *why* humans believe in deities, but they will not provide a reason *to* believe. In the same sort of way, we can coherently prise apart the concept of a motive from the concept of a motivation. To have a motive is to have a reason to act or a cause in the name of which one might act. To have a motivation, by contrast – to be motivated – is rather to have been the subject of a certain kind of causation. An agent who acts for a cause, or who has a reason or a motive to act, is an agent who has identified an impetus for acting that he has adopted as his own. By contrast, an agent who is caused, or whose actions are for a reason or are motivated by someone, may actually be in the dark about quite why he acted as he did – and may even be less able to answer the why question than an external observer would be. (We can imagine someone in the dock or the therapist's chair who only a long time after the event comes to understand why he acted as he did.

'Cause' in his sense would be an explanation for action, not an invitation to it.)

To this extent, I think that Douglas has perhaps fallen into a linguistic trap. When he talks about moral enhancement in terms of an agent's having

altered herself in a way that may reasonably be expected to result in her having morally better future motives, taken in sum, than she would otherwise have had³⁰

this is something that can be perfectly well be accommodated by something along the lines of the Aristotelian picture of moral formation by habituation that I outlined above. But when he suggests the possibility that this may be achieved by biomedical means, he has stopped talking about moral motives – he is now talking about moral motivations.

How does this help revivify the argument from freedom? Well, for Douglas, it would appear to be sufficient that an actor is caused by some set of mental or neural characteristics to behave in desirable ways. So imagine a variation on the Frankenstein story, in which Victor Frankenstein has the opportunity to instil in his monster certain dispositions and inclinations: to make him, wherever possible, more responsive to desirable moral motives and less likely to behave in morally undesirable ways.

But this would fall foul of the argument from freedom. This is not so much because the monster would have fewer options, as because it is caused to act without acting *for* a cause: freedom attaches to adopting a cause, having a reason, or having a motive. It's for this reason that McDowell's virtuous agent is no less free for being unable to see the other (merely) logically possible courses of action that his *Doppelgänger* on Twin Earth may have chosen. Though there is a sense in which Douglas' agent and McDowell's are alike, inasmuch as that acting in a different manner is unthinkable, there is another sense in which the difference is vast, just because the McDowellian agent has a reason that makes the other reasons to act inert. Moreover, he could, if things turn out in the right way, morally re-habituate himself to some other cause – a point that will be important in a little while. This means that he would be, in the Miltonic sense, sufficient to stand, though free to fall.

It might even be possible to rehabilitate the concept of dignity in this context. Notwithstanding the familiar problems with the concept, if anything stands a chance of eroding an agent's dignity, this does: in effect, we would be trying to make them beholden to our moral preferences. And even if this is what parents do as part of the normal course of raising children anyway, those children are able to adopt the moral shape of their upbringing as a cause of their own.

Already, then, we've got a reason to wonder whether moral enhancement is compatible with the good life. If part of being a fully-fledged human moral agent is to be morally free, it could turn out that moral enhancees are missing out on a significant dimension of the complete human life. And to be caused to

act in accordance with someone else's preferences is no less corrosive of a good life, richly understood.

Does the same apply to attempts that we might make to enhance ourselves? Imagine that someone is told by others that his habitual patterns of behaviour are morally unacceptable, and resolves to alter that behaviour. There does seem to be less of a worry about dignity here, for although our anti-hero is attempting to force himself to meet others' expectations, that attempt is *his*. So a person who tries to 'recondition' himself to be more likely to behave in accordance with what his better angels would prefer, and takes full advantage of the insights provided by brain-science to ensure that he succeeds in altering his behaviour, might have more of a reason to do so – and would not necessarily have to worry about eroding his own freedom or dignity. (Having said this, Harris is not wholly sold on the idea that this would be so straightforward – moral bioenhancement, he worries, 'may well dull [moral consciousness] to the point where the individual is no longer choosing'.³¹ On the basis of what I've been arguing so far, the absence of choice – conscious choice at least – doesn't worry me too much; but the reason why conscious choice is absent does matter; and inasmuch as that Harris is nodding towards the distinction between being caused and having a cause, I am in agreement with him. And self-administered moral enhancement that elevates the importance of being caused in comparison to that of having a cause might quite straightforwardly be incompatible with human flourishing, since it could quite straightforwardly be said to disrupt an aspect of complete human functioning.)

Still, there would be other problems to be overcome all the same. Freedom isn't the only concern about moral enhancement. The big problem in respect of attempting to give someone better motivations is that there is room for reasonable disagreement concerning what 'better' would be.

The argument from reasonable disagreement

In a good deal of the literature, moral enhancement is taken to imply making people more pro-social in some sense – I count Persson and Savulescu's reinforced altruism across space and time to be a version of increased pro-sociality. But since 'morality' is logically different from 'pro-sociality', there is obviously a need to provide an argument about why moral enhancement should be taken to mean boosting pro-sociality. And it makes sense to wonder whether, in the absence of such an argument, we ought to be wary of 'improving' an agent's dispositions until we are sure of the measure by which we are supposed to assess that improvement. Call this the 'Argument from Reasonable Disagreement', or 'ARD' for short.

We might want to assume that there are certain traits the desirability of which should be minimally controversial (although even here, the existence

of a consensus doesn't foreclose the possibility of reasonable disagreement³²). Douglas' own example of non-racism seems to fit into this category – although we should be wary of the possibility that, if reducing racial aversion this strikes us as being obviously worth doing, this is perhaps a warning that his example is too slight to carry the weight of a fully-fledged argument. But there are other dispositions whose undesirability is not so straightforward, and in such cases, there is a question to ask about what the next move should be. So, when Douglas also suggests that a propensity to violent aggression might also merit 'enhancing' away – and that's a trait that Savulescu et al. would happily eliminate as well³³ – the way does seem open for people like Harris and Chan to point out that we ought to slow down a bit, because there are some situations in which violent aggression might be either called for or, at least, not be wholly out of line. (Douglas admits as much himself.) Harris' preferred example here is of using violence to overpower a would-be hijacker, or someone else intent on causing damage or destruction³⁴: violent aggression, channelled in the right way, might be precisely what's in order. While removing the disposition to violence may reduce the chance that someone will be a hijacker, it may also reduce the chance that people nearby would be able to do much to thwart a person sufficiently unenhanced still to mount a hijack.

Douglas' way of avoiding the pitfalls raised by this point is to rephrase the claim:

[T]here are some emotions such that a reduction in the degree to which an agent experiences those emotions would, under some circumstances, constitute a moral enhancement.³⁵

It's not clear that this is enough. For one thing, the claim now looks horribly formalistic: it doesn't tell us anything much about which emotions and dispositions we would do well to enhance or reduce. Neither is it clear that there would be agreement about at least some of the enhancements up for consideration. For there is, as yet, still controversy about the standard against which we might hope to calibrate our moral assessments.

For natural scientists, though there can be reasonable disagreement and a range of points of view about the way the world is, there is also a fairly clear way to resolve it, which is to design and perform another experiment. Things aren't so easy in respect of morality, though: there may be at least *prima facie* reasonable disagreement about what is good, and further reasonable disagreement about what the criteria for applying a word like 'good' are in the first place. In any given moral dispute, each participant will think his position correct – this much is implicit in his thinking that it's worth disagreeing about to begin with – but it is not a given that either is less rational than the other, or has failed in some aspect of moral perception. On the basis that it either is or isn't permissible to, say, kill a person at their request, it's impossible for both to be correct. Yet it's unclear whose moral positioning needs to be enhanced.

It's certainly not obvious that one of them is insufficiently empathetic and that boosting pro-sociality would make solving the dilemma any easier.

This is not hyperbolic scepticism: moral attitudes change over time and distance. What counts as hypocritical flattery in one place may be required by politeness in another; turning the other cheek may be a virtue for some, and seen by others as a blameable unwillingness to take revenge; bloodthirsty Aztec religious rituals were obviously not considered indefensible by the Aztecs who performed them. Aristotle has no time for a generalized philanthropy. We take it as read that justice is closely related to giving to people or punishing people according to what they deserve – but the *Odyssey* nevertheless speaks of Athene urging Odysseus

to go round collecting scraps from the suitors and so learn to distinguish the good from the bad, though this did not mean that in the end she was to save a single one from destruction³⁶

– which seems to show a flat contradiction of the principle coming from the mouth of a goddess: the goddess of justice, no less! It is possible that all these unfamiliar conventions are, as it happens, trading on a misapprehension of some sort – but until we can come up with a way to demonstrate that once and for all, there mightn't be all that much we can say to settle the matter.

Even within a single life – or within a single day of a single life – a disposition may be held to be desirable at some times, and utterly undesirable in others. To be disinterested and scrupulous about detail is a virtue in a judge, but becomes a vice when she removes her wig and goes home to her family. Were we to seek to enhance the moral motivations of judges, ought we to make them more disinterested? That would mean that we had made them vicious outside office hours.

A nice example of this trap is given in a video on moral enhancement on the website of Humanity+, formerly known as the World Transhumanist Association.³⁷ In this video (the transcript of which is available via his website³⁸), Steve Omohundro talks of moral progress:

If you look at recent history . . . there is a trend toward great moral progress. 200 years ago, slavery was generally accepted. Now, it is viewed as immoral almost everywhere in the world, at least officially, and pressure is put on societies that still allow it. The same is true of torture, though there has been a lot of recent controversy about it. We have the Geneva Convention and the notion of war crimes, the sense that war is bad but there are things within war that are especially bad. We have the establishment of women's rights in many countries, though some are still lagging. The same is true of racial equality. And the animal rights movement is growing rapidly.

. . .

When we think about transhumanism, I think we should start from humanitarianism. That is the notion that the things that most humans view today as precious, like human life, love, happiness, creativity, inspiration, self-realization,

peace, animals, nature, joy, children, art, sexuality, poetry, sharing, caring, growth, contribution, spirituality, family, community, relationships, expression, are truly precious. These things matter because they matter to us. We may not know why these things matter to us, but that does not take away from the fact that they matter to us.

For sure: we have a set of social norms and conventions that we did not have a few centuries ago; and we (presumably) believe that the changes are for the better. But it's fallacious – and a touch arrogant – to suggest that 'the things that most humans view today as precious . . . are truly precious'. We cannot say, at least without further argument, that the reason why we prefer today's conventions is other than that they're the ones to which we're used – even allowing, extraordinarily enough, that our forbears had no time at all for any of the things he thinks 'truly precious'; this is a point nicely captured by John Mackie's suggestion that

people approve of monogamy because they participate in a monogamous way of life rather than [participating in] a monogamous way of life because they approve of monogamy.³⁹

To talk about moral progress in other terms is a bit like looking at Chaucerian English and talking about all the progress we've made in becoming better at spelling.

More particularly, it's hard to escape the worry that the progress mentioned in Omohundro's talk is all towards the kind of things that the wealthy, middle-class, educated people who tend to become transhumanists like. So when he continues to say that

I think that the kind of morality and moral structures we want to create using new technologies should serve to preserve these qualities,

it's tempting to respond with a curt 'Well he would say that, wouldn't he?'. There is no suggestion that the moral characteristics of which these people approve may be merely the transient markings of a particular group in a particular epoch. He may have identified characteristics that are, as a matter of fact, desirable; but the problem is that there's room for serious doubt on that front.⁴⁰ (In fairness, we should remember that Omohundro has several companions in guilt: even Kant starts with the idea that his moral system will simply elucidate and take its lead from the 'common idea of duty and of moral laws'.⁴¹)

The ARD applies equally to Persson and Savulescu's proposal – or any other attempt at dispositional engineering. In one sense, the desire that people should be more sympathetic and just⁴² seems fairly trivial; but once we begin to put meat on the bones of what a just attitude would look like, we open the way for there to be reasonable disagreement. Persson and Savulescu take a particular understanding of justice as their touchstone, and think that we

should be time- and location-blind. But it is possible reasonably to think that one has commitments to some that one does not have to others; and, if this is right, then being dispassionate would be the *wrong* way to go.

The point of the ARD is this: it conflates giving someone better motives with giving someone what we *believe* to be better motives; and those beliefs may be false. This is admittedly not necessarily a fatal problem. It is possible to endorse a particular set of values while still being awake to the possibility that a radically different set is possible. In like manner, one may endorse a particular set of spellings while still being awake to the possibility that they will alter over the coming centuries. Hence, though for any two cultures *P* and *Q* with divergent sets of moral beliefs, members of *P* will believe things that the people of *Q* will think mistaken, people in either can quite coherently entertain a first-order belief about right and wrong at the same time as a second-order belief about the possible falsity of the first-order belief. Nevertheless, if we accept the formal moral desirability of giving someone morally better motives, we may be systematically mistaken in our attempts to realize our desires; and we may thereby unwittingly give someone motives that others may quite reasonably disavow.

Would-be moral enhancers may want to bite this bullet, and admit cheerfully enough that all they want, and could reasonably want, is to make people more likely to behave in ways that they prefer. But if this really is what's going on, then the word 'moral' seems to have been eviscerated – or more likely become equivalent to 'moralistic' – and we're back to the worries about freedom, dignity, and shaping others' ends for them. It's one thing to make predictions about what morally enhanced people would look like; another to make claims about the kind of people that we would like to see and call that moral enhancement.

For this reason, it may simply be foolhardy to attempt to instil particular dispositions in people without at the same time giving them the ability to reason morally and recalibrate those motivations should the situation require: we may have got ourselves into a muddle about what would be desirable.

Enhancing moral reasoning

As we've seen, John Harris' attack on non-cognitive approaches to moral enhancement is substantially based on the idea that intervening to improve someone's dispositions makes them less free, and therefore less moral. For him, judgement is key:

One thing we can say with confidence is that ethical expertise is not 'being better at being good', rather it is being better at knowing the good and understanding what is likely to conduce to the good.⁴³

Elsewhere, he makes the much stronger claim that '[b]eing moral is like being scientific' and that the idea that emotionally-motivated behaviour is properly

considered moral 'is simply not credible'.⁴⁴ For Harris, morality is a cognitive process. Similarly, with Sarah Chan, he admits that we could use drugs to 'treat anti-social personality disorders and social dysfunction'. But, he asks,

would this constitute a 'moral enhancement', if the beneficial social effect were to be achieved through suppressing patients' own moral judgement even if the outcome was a world with less harm?⁴⁵

We might end up with 'nicer' people: but this may be moralistic enhancement rather than moral enhancement proper.

One thing that does seem to be clear in this part of the argument is that Harris' suspicion of non-cognitive moral enhancement such as that advocated by Douglas may not actually mean that many hits are landed. It's possible – and, I suspect, likely, that Harris and Douglas are talking past each other: the latter is talking about improving behaviour, and the former about improving thinking about behaviour (hence the point about moral *expertise*). As a result, it is not obvious to me that Harris' arguments really inflict much damage on Douglas'. Still, given the problems that face dispositional engineering as a strategy for moral enhancement, we might wonder whether some kind of cognitive moral enhancement programme could improve our moral reasoning and thereby generate more desirable behaviour.

It's a reasonable question, but it's not obvious that it would. One of the reasons for this is that improving a person's reasoning about what he should do is not necessarily going to be enough to make him do it – it will not remove weakness of will. Further enhancements to avoid this are a possibility – but not guaranteed to be desirable. More importantly, there's no obvious reason to think that morally undesirable behaviour really does have much to do with a failure of reasoning, so that being better at moral reasoning is going to improve behaviour; it might simply allow people more easily to convince themselves that their victims deserved to be treated poorly.

Imagine an antebellum Dixieland slave-owner; call him Zeke. Zeke is not an altogether bad man: he believes that one should not behave arbitrarily badly to one's slaves, and even thinks that he has duties to them: duties to provide them with food and basic medical care, to refrain from violence, and so on. However, Zeke also happens to believe that, whether it be by nature or divine ordinance, black people simply are inferior to whites, and properly (or at least not improperly) kept in bondage. The sympathy he shows to his slaves is, more or less, the same as the sympathy that he shows to his cattle: he does not allow his animals to starve, or to suffer illness or injury, or to be beaten, either. Still: he thinks that horses and cattle are not his moral equals, and that it is permissible to force them to work, sell them, and so on – and the same applies to his slaves. We'd be inclined to say, without much need for argument, that his beliefs about the status of his animals may be true, but the corresponding belief about his slaves is plain false; and this is at the root of his immoral behaviour in respect of them.

This example might be taken to speak in favour of a cognitive account of moral improvement. Harris articulates something like this thought:

The most obvious countermeasure to false beliefs and prejudices is a combination of rationality and education, possibly assisted by various other forms of cognitive enhancement, in addition to courses or sources of education and logic.⁴⁶

It is certainly not a given that Zeke's moral failure would not be remedied by making him more empathetic or pro-social: his problem is not so much a lack of empathy, as a mistake about how it should be directed. But for his beliefs about race, he would be an uncontroversially good (or, at worst, much better) person: and so if we want to make Zeke more virtuous, all we have to do is shift that belief about the moral worth of a portion of humanity. It's not clear that the biosciences could help us here – and so in terms of my wider investigation about the contribution that those sciences might make to the good life, we could turn out to be on a hiding to nothing. But even ignoring that, the problem for the cognitive account is that it is not obvious that that would generate a moral improvement anyway.

A significant part of this is that, if Zeke's underlying assumptions about his slaves are that they are at least somewhat lower than himself in terms of moral importance, and there *is* a fault in his moral reasoning, we need to distinguish between correcting the faulty reasoning, and correcting the false belief on which it builds. Imagine that Zeke teaches some of his slaves to read; correcting Zeke's reasoning might lead him to the conclusion that he is thereby wasting good money treating them better than they really warrant, given their inferior moral status. Harris has faith that making people better at moral reasoning will make them more likely to behave in a morally desirable way; but it's at least possible that it'll have the opposite effect. There seems to be no guarantee from the start that someone like Zeke would, by thinking more clearly about how to treat black people, decide that he ought to treat them better. The problem is with the underlying false belief about the moral status of certain humans – but this might not be the kind of thing that can easily be corrected by cognitive sharpening: falsity is not a sign of irrationality. (The same criticism applies to Persson and Savulescu's plea for generational blindness: however admirable such an attitude, it's not obvious that it's particularly *rational*.)

Indeed, there's reason to be fairly confident that improving someone's reasoning would not necessarily make a person's moral behaviour better, because behaviour is often informed by inbuilt prejudice, and that is not always open to reform by clearer thinking. Such prejudice – be it against members of a particular ethnic group, or one sex, or people with certain disabilities, or whatever – can persist even when a person honestly believes that it is not there; it can persist even when there are good reasons to expect that it

would have no cognitive traction at all. A good example of this is provided by the American political activist Jesse Jackson, who is widely quoted as having said that

there is nothing more painful to me at this stage in my life than to walk down the street and hear footsteps and start thinking about robbery, then look around and see somebody white and feel relieved.⁴⁷

This statement claim is important in respect of the idea that cognitive enhancement can eliminate the basis for undesirable behaviour such as explicit racism, because the idea that racial prejudice is reducible to false beliefs, which would be eliminable by clearer thinking – and that racist behaviour would thereby never appear – seems to wither under its light. After all, an educated black civil-rights activist is rather more likely positively to disavow than hold a conscious belief that black people are more dangerous; and yet a prejudice against black people seems to linger. I'm not, of course, suggesting that Jackson has ever behaved in a racist way; but if even his underlying attitudes can be influenced by underlying and consciously disavowed prejudices, it would not be surprising to find that at least some racist behaviour in others is influenced in the same sort of way. Eliminating false beliefs is not enough to improve behaviour.

This is instructive in respect of Zeke. It might be that he has a set of conscious beliefs in respect of black people; and it might be that education (in the sense of learning to think more clearly: the sense that Harris seems to mean in the quotation above) or some other form of cognitive enhancement, would alter them for the better. Indeed, the idea that one can draw some of the venom from undesirable moral attitudes by means of education is intuitively plausible. By such strategies, we might well be able to get people to take on board that their attitudes and behaviours are ill-founded and indefensible. But there persists the worry that such interventions would not be enough: if Jesse Jackson is vulnerable to a kind of cultural conditioning about black people that persists despite having in his own use of the first-personal pronoun plentiful evidence that black people are not particularly threatening, then so would be Zeke; and getting Zeke genuinely to believe in the moral equality of black and white people would not obviously be enough to shift the underlying prejudice if it's not enough for someone like Jackson. This, incidentally, gives us another brickbat to throw at the idea that moral failure is a failure of reasoning: increasing Zeke's moral reasoning abilities might just mean that he is better at coming up for a *post hoc* rationalization of his unexamined-because-unconsciously-held attitudes.

In short, if it turns out that racism has, or can have, its roots in some non-cognitive characteristic, cognitive enhancement is unlikely to be enough to get rid of it. There is, then, at least one situation in which improved cognition will

not lead reliably to improved disposition. I suspect that this is enough to break the back of the idea that improving moral reasoning will make people morally better, and that John Harris' claim that

if [love, altruism and the like] are not felt, or worse not felt for particular groups, racial, national, religious political, or for particular individuals, that is when you need moral philosophy, moral reasoning⁴⁸

is, on this ground, open to a charge of naïveté.

By the same token, if certain undesirable attitudes worm their way into our souls on account of our socialization, then unless the social *milieu* changes, these attitudes will stand a chance of remaining; but, more optimistically, the hope remains that a change in the social *milieu* could be all that's needed to eliminate them. Again, bioscientifically informed procedures may well be *de trop*. Moral enhancement in the broadest sense might still be something that we think worth achieving – but it would be more-or-less of the same stripe as the kind of common courtesy that we instil as a part of growing up. It does not generally require oxytocin or giving kids particular beliefs: it's a matter of acculturation.

And the same could well apply in respect of making people morally better in the round: using bioscientific interventions to improve problematic behaviour could well be, at the very least, like using a monkey-wrench to crack a monkey-nut. Once again, we're back to the idea that it's habituation that might be able to carry most, if not all, of the weight of making people better. Once again, Aristotle seems ahead of us:

Moral goodness is the result of habit, from which it has actually got its name, being only a slight modification of the word '*ethos*'.⁴⁹

It's not beyond the realms of possibility that that's where the most progress is to be made in respect of moral enhancement.

Is moral enhancement desirable anyway?

If proponents of dispositional engineering and cognitive moral enhancement are talking past each other, this might be a sign that their positions are not, in the end, incompatible. I suspect that this might be correct – though, if the ARD is well-targeted, we might do well to hesitate before trying to combine them. However, it's also worth noting another aspect of the ARD that has not been mentioned before. Should we make a mistake about the best dispositions to instil in others, and the things we thought desirable at time t turn out to be not desirable at $t+1$, this would be bad, or at least non-optimal, from our point of view. But what about the point of view of the enhancee?

Go back for a moment to the Frankenstein example, where Dr Frankenstein gave his monster a set of desirable dispositions. All being well, the monster's dispositions and his moral thinking would cohere. But if the ARD is correct, this might not be guaranteed: the monster may decide that the dispositions that he has had instilled are mistaken; and this would seem to bring a significant risk that he would be subject to a series of dispositions that he morally disavows. This may well undermine his quality of life: he would be alienated from his own motivations, and such alienation may be distressing.

What goes for monsters goes for humans: to have the freedom to identify a cause, but to be caused to act otherwise, is the kind of thing that we may reasonably expect to be incompatible with a good life; and if the point of the biosciences (and enhancement in particular) is at least partially to make people's lives better, such an eventuality would appear to undermine the point *ab initio*.

Correlatively, though one might expect all people to be morally good in the broad sense, this is not enough for them to count as leading a good life. As Ottfried Höffe points out, 'a morally good life is no sufficient condition for a totally good life'.⁵⁰ For him, though '[t]he morally good life is a condition for the art of living',⁵¹ and though there may be a tension between the art of living and the moral, and though even 'the moral takes precedence', still, it does so

not so exclusively that it eliminates all rights to the art of living. . . . As a consequence, a philosophical ethics is not satisfied with a philosophy of morals. It recognises happiness as an independent, admittedly subordinate principle.⁵²

The idea that the moral takes precedence over the art of living is not something we should accept as a given; but the final sentence of the quotation does capture an important point. Moral enhancement pursued as an end in its own right ignores the fact that moral goodness is at most a *part* of a life worth living. Moral enhancees' lives have not necessarily been enhanced in the round. If we're interested in the promotion of the good life, that matters.

In other respects, moral enhancement may not be the best solution to a problem faced by a would-be enhanced person. Take, for example, the idea that we could intervene pre-natally to optimise a person's MAOA functioning. Savulescu et al. note that it is not only the people around someone who is neuro-genetically disposed to antisocial behaviour that suffer:

Given that compulsive violent or criminal behaviours are typically punished with incarceration as well as social rejection, we can assume that these behaviours typically harm a person who expresses them, as well as their victims. Selections which constrain a person from making such harmful choices may therefore be justified.⁵³

Still, while there may be harms that accrue from being at the receiving end of the criminal justice system, this could just as easily be taken as a call for the reform of that system as for moral enhancement. Indeed, Savulescu et al. admit the possibility that

[a] deficiency of MAOA results in a build up of neurotransmitters. Once they reach abnormal levels, this is thought to *increase the probability* of a person demonstrating an excessive, even violent, *reaction to stress* – a theory demonstrated in knockout mice.⁵⁴

The emphasis is mine, and it seems important. One reason is that genetic cues tell us about probabilities rather than certainties. A second, and more important reason, is that if the MAOA hypothesis is true, a stressful environment is at least a companion in guilt when it comes to causing antisocial behaviour; people with the mutation in question might benefit from having access to (for example) a pill that would help prevent stress; but even here, this would – it seems to me – count as therapeutic rather than enhancing, and not necessarily *morally* therapeutic at that. And, of course, reducing socio-economic stressors would also, presumably, improve the lives of others living in the same circumstances who have the good fortune not to have abnormal MAO activity.

That there is a possible role for environmental factors in understanding and ameliorating undesirable behaviour is no small point. There is evidence that links antisocial behaviour with exposure to lead pollution, for example.⁵⁵ This may well serve as an indication that to focus on moral enhancement as a solution to social problems is to misidentify the root of those problems: if people's social development is retarded or skewed by exposure to pollutants of one sort or another, they do not need bare enhancement; they need therapy. More importantly, it may be possible to improve behavioural patterns without having to worry too much about manipulating individuals' brains, and so enhancing technologies in the sense understood by many of their proponents may be *de trop* (although, admittedly, this would likely have more of an impact on subsequent generations than on anyone who currently finds himself with a brain affected by pollution). Finally, everyone would benefit from a cleaner environment anyway, irrespective of whether or not they had an elevated propensity for criminal activity: Elise Gould estimates that

[f]or every dollar spent on controlling lead hazards, \$17–\$221 would be returned in health benefits, increased IQ, higher lifetime earnings, tax revenue, reduced spending on special education, and reduced criminal activity.⁵⁶

Behaviour isn't the only aspect of life that'd get better.

Finally, there is no guarantee that moral enhancement would be good for humanity as a whole: there could be vices that are worth keeping. If that is too much of an oxymoron, we could still talk about characteristics that, while undesirable when manifested in individuals, are indirectly beneficial.

In his *Fable of the Bees*, Bernard Mandeville prefigures Adam Smith by voicing the idea that the common wealth may be increased by the action of self-interested behaviour: 'Bare Virtue can't make Nations live/In Splendor',⁵⁷ and there may be public virtue grown from the soil of private vice. This

sort of claim might be spun out thus: that a great deal of economic activity is generated by people wanting to maximize their own welfare, and would not take place if people were altruistic, or inclined to treat others' welfare as being as important as their own. Such activity, over time, generates welfare for everyone. In reality, things are much more complicated: though there may be a 'trickledown effect' from private vice, there may be problems related to market failure, widened inequality; and the benefits may take years to dissipate, which is of little comfort to those living in need right now. Still: a good deal of current economic theory tells us – insists, in fact – that, overall, the most stable and efficient way to increase general welfare is to allow people *not* to be altruistic. The trick of the just society is to harness those vices and channel them rather than eliminate them:

So Vice is beneficial found,
When it's by Justice lopt and bound.⁵⁸

The lopping and binding are crucial: but this is not a reason to opt for bioscientific moral enhancement. What actually counts as the just distribution of resources is a matter for reasoned public debate; getting people to behave in a way that justice dictates is what public policy is for.

This is a pattern that could be repeated more or less indefinitely. Indolence may be a vice, but a desire for ease can, if appropriately harnessed also be the spur for labour-saving devices that make life better, for example. Elijah Millgram's essay comparing Hume and Gibbon in their accounts of the fall of Rome points out that – bluntly – Gibbon thought that Rome fell because it became too immoral while Hume thought that it fell for just the opposite reason: too much morality.⁵⁹ On the assumption that Hume thought the fall of Rome to be a loss, this would appear to generate the idea that too much morality – or too much morality of a certain kind – might be bad over all. There's something in this reminiscent of Orson Welles' 'cuckoo clock' speech in *The Third Man*:

After all, it's not that awful - you know what the fellow said . . . In Italy, for thirty years under the Borgias, they had warfare, terror, murder, bloodshed – they produced Michelangelo, Leonardo da Vinci and the Renaissance. In Switzerland they had brotherly love, five hundred years of democracy and peace – and what did that produce . . .? The cuckoo clock.⁶⁰

Historically inaccurate as it may be, the speech captures nicely the idea that not all admirable things are godly to the core. There is a sense in which wholly 'moral' lives within a wholly 'moral' world would be lives that noone would want to live, in a world that no one would want to inhabit.

This is not a defence of vice *per se*: it might well be that there are all kinds of traits that are properly listed as vices that we would rather be without,

and others that are properly listed as virtues that we would rather see more frequently, and that no reasonable people would disagree. But there might be indirect benefits to be had from retaining the possibility of at least some other vices; and – more problematically – it might be that certain undesirable behaviours are coeval with other, desirable, ones.

I would want to be among the first to admit that these qualms about moral enhancement are not definitive. However, I think that they are enough to start to pick away at the idea that moral behaviour can easily be altered for the better by bioscientific means. The standard by which improvement is to be measured is unclear; the benefits to society are unclear; and the enhanced individual, for a range of reasons, might have defensible objections to being enhanced; and, most importantly, even when moral enhancement is possible and desirable, there really isn't any reason to suppose that bioscientific interventions are the wisest method of pursuing it.

Moral enhancement by habituation and education leaves room for nuance, for judgement, and for altering not just the person, but the interplay between him and his environment. Bioscientific interventions, it seems to me, are much too rough-and-ready, and take too much for granted.

Notes

- 1 Douglas, 2008, p. 229.
- 2 Ibid., p. 231.
- 3 Ibid.
- 4 Savulescu et al. 2006, *passim*.
- 5 Persson and Savulescu, 2012, *passim*.
- 6 Ibid., p. 107.
- 7 Savulescu et al. 2006, p. 165.
- 8 Harris, 2011, p. 110.
- 9 Ibid., p. 104.
- 10 Ibid.
- 11 Kant, 1993, 4:398–9.
- 12 Savulescu et al. 2006, p. 169.
- 13 Fine, 2011, *passim*, but esp. chs. 9–11.
- 14 Persson and Savulescu use the word 'causally' rather than 'neurophysically'; I've made this substitution because of the use that I'll be making of the word 'cause' later in the chapter. It doesn't detract from the sense of the argument, though.
- 15 Persson and Savulescu, p. 112.
- 16 'Persons, Character and Morality', in Williams, 1999, p. 18.
- 17 cf. Persson and Savulescu, *op cit*, p. 113.
- 18 'Are Moral Requirements Hypothetical Imperatives?', in McDowell, 1998, p. 85.

- 19 Ibid., pp. 90–1.
- 20 DeGrazia, 2013.
- 21 Ibid.
- 22 Harris and Holm, 2003, pp. 125–6.
- 23 see Lyons, 2012, pp. 371–2.
- 24 Aristotle, 1976, p. 91.
- 25 Ibid., pp. 91–2.
- 26 Ibid., p. 92. It's interesting to note that Aristotle associates ethical formation with the *polis* rather than the home. Maybe he thinks that the importance of the domestic is self-evident – except that elsewhere in his writings he has no compunction about weaving commonplaces explicitly into his investigations.
- 27 Aristotle, 1976, p. 92.
- 28 see Anon, 2009b and Anon, 2010b.
- 29 Douglas, 2008, p. 229.
- 30 Ibid.
- 31 Harris, 2013.
- 32 *contra* Shook, 2012, p. 4; cf. DeGrazia, 2013.
- 33 Savulescu et al. 2006, *passim*.
- 34 see Harris and Chan, 2010, p. E183; Chan and Harris, 2011, p. 131; Harris, 2011, p. 106.
- 35 Douglas, 2008, p. 231.
- 36 Homer, 1991, *VXII*: 360–4.
- 37 <http://www.humanityplus.org/2009/05/steve-omohundro-on-ai-and-the-future-of-human-morality/>.
- 38 Omohundro, 2009.
- 39 Mackie, 1977, p. 36.
- 40 In a similar vein, when Agar (2013) expresses a concern that post-humans might have a higher moral status than mere humans, and come to see mere humans as things that could be sacrificed in the pursuit of post-human interests, Douglas (2013) is confident in his response that the transition to post-humanity 'would, after all, make them less liable to sacrifice' (p. 76). The thought here seems to be that enhanced post-humans would express certain moral traits as a matter of course, and that these traits are predictable. I see no reason why we should be forced to accept this – though I see no particular merit in Agar's worries, either.
- 41 Kant, 1993, 4:389.
- 42 Persson and Savulescu, 2012, p. 108.
- 43 Harris, 2010, p. 104.
- 44 Harris, 2012, p. 297.
- 45 Chan and Harris, 2011, p. 131.
- 46 Harris, 2010, p. 105.
- 47 The source of this quotation appears to be an article in the *Chicago Sun-Times* from the 29 November 1993 called 'Crime: New Frontier – Jesse Jackson Calls It Top Civil-Rights Issue' by Mary A. Johnson. I'm grateful to a pseudonymous comment on the *Feminist Philosophers* blog for this information; however, I have not been able to get access to the original to verify it.

- 48 Harris, 2012, p. 3; slightly modified.
- 49 Aristotle, 1976, p. 91.
- 50 Höffe, 2010, p. 331.
- 51 Ibid; slightly modified.
- 52 Ibid, p. 333.
- 53 Savulescu et al. 2006, p. 165.
- 54 Ibid., p. 159.
- 55 see Nevin, 2007; Mielke and Zahran, 2012.
- 56 Gould, 2009.
- 57 Mandeville, 1997, p. 35.
- 58 Ibid.
- 59 Millgram, 2005, ch. 8, *passim*.
- 60 Greene and Reed, 1968, p. 114.