

- Ransford, H. E. (1970). 'Skin Color, Life Chances, and Anti-White Attitudes', *Social Problems*, 18(2): 164–79.
- Rodgers, W. M., and Spriggs, W. E. (1996). 'What Does the AFQT Really Measure? Race, Wages, Schooling, and the AFQT Score', *Review of Black Political Economy*, 24(4): 13–46.
- Santos, A., Meyer-Lindenberg, A., and Deruelle, C. (2010). 'Absence of Racial, but Not Gender, Stereotyping in Williams Syndrome Children', *Current Biology*, 20(7): R307–8.
- Savulescu, J. (1994). 'Rational Desires and the Limitation of Life-Sustaining Treatment', *Bioethics*, 8:191–222.
- Savulescu, J., and Persson, I. (2008). 'The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity', *Journal of Applied Philosophy*, 25(3): 162–7.
- Savulescu, J., and Persson, I. (2011). 'The Turn for Ultimate Harm: A Reply to Fenton', *Journal of Medical Ethics*, 37(7): 441–4.
- Silber, M. et al. (1987). 'The Effect of Oral Contraceptive Pills on Levels of Oxytocin in Plasma and on Cognitive Functions', *Contraception*, 36:641–50.
- Smith, G. (2011). 'Steve Jobs Doomed Himself by Shunning Conventional Medicine until Too Late, Claims Harvard Expert', *Mail Online*, 14 October, <http://www.dailymail.co.uk/news/article-2049019/Steve-Jobs-dead-Apple-CEO-shunned-conventional-cancer-medicine.html> (accessed 7 December 2012).
- Stock, S., Karlsson, R., and von Schoultz, B. (1994). 'Serum Profiles of Oxytocin during Oral Contraceptive Treatment', *Gynecological Endocrinology*, 8(2): 121–6.
- Terbeck, S. et al. (2012). 'Beta-Adrenergic Blockade Reduces Implicit Negative Racial Bias', *Psychopharmacology*, 222(3): 419–24.
- Terbeck, S. et al. (Under Review). 'Emotion in Moral Decision-Making: Beta Adrenergic Blockade Increases Deontological Moral Judgments'.
- Thaler, R. H., and Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Tse, W.S., and Bond, A. J. (2002). 'Serotonergic Intervention Affects Both Social Dominance and Affiliative Behaviour', *Psychopharmacology*, 161:324–30.
- Watson, G. (1989). 'Free Agency', in J. Christman (ed), *The Inner Citadel: Essays on Individual Autonomy*. New York: Oxford University Press, 109–22.
- Wood, R. M. et al. (2006). 'Effects of Tryptophan Depletion on the Performance of an Iterated Prisoner's Dilemma Game in Healthy Adults', *Neuropsychopharmacology*, 31(5): 1075–84.
- Young, R. (1986). *Personal Autonomy: Beyond Negative and Positive Liberty*. London: Croom Helm.

3.2

Commentary

(Im)Moral Technology? Thought Experiments and the Future of 'Mind Control'

Robert Sparrow

Philosophers love thought experiments. A good thought experiment progresses intellectual enquiry by clearing away extraneous details and exposing the philosophical essence of a problem. By clarifying the issues at stake, thought experiments help distinguish different philosophical positions and reveal connections between different philosophical problems.

In some areas of philosophy, inventing a good thought experiment is its own reward: philosophers have made careers out of formulating problems that have kept other philosophers busy trying to solve them. However, in applied ethics, where the ultimate goal must be to contribute to solving real-world problems in all their messy complexity, arguments involving thought experiments are both especially tempting and especially problematic. They are especially tempting because it is hard to see how we can solve the messy complex issues we face in the real world if we cannot first develop our intuitions and principles in deliberately simplified test cases. They are especially problematic because of the difficulties involved in moving from real-world dilemmas to thought experiments and back again.

In order for an argument involving a thought experiment to progress debate in applied ethics three things must be true. First, the thought experiment must accurately represent and illuminate a pressing ethical dilemma. Second—and most obviously—the central claims of the argument regarding the thought experiment must be plausible. Third, it must be possible to apply or develop the arguments established with reference to the thought experiment to the real-world cases the experiment is intended to illuminate.

In their paper, 'Autonomy and the Ethics of Biological Behaviour Modification', Savulescu, Douglas, and Persson (in this collection) are discussing the ethics of a technology exist for improving moral motivation and behaviour, which does not yet exist and—as I will argue below—will most likely never exist. At the heart of their argument sits the imagined case of a 'moral technology' that magically prevents people from developing intentions to commit seriously immoral actions. It is not too much of a

stretch, then, to characterize their paper as a thought experiment in service of a thought experiment. In this commentary, I will argue that there are serious reasons to question the extent to which their argument meets each of the challenges involved in the use of thought experiments in applied ethics, outlined above. While Savulescu et al. succeed in showing how behavioural modification might be compatible with freedom and autonomy—and perhaps justifiable even if it were not—in the fantastic case they consider, there is little we can conclude from this about any technology of ‘moral bioenhancement’ in the foreseeable future. Indeed, there is a real danger that their argument will license attempts to manipulate behaviour through drugs and brain implants, which raise profound moral issues that they barely mention.

Can we really make people ‘more moral’ through biomedical interventions?

Savulescu and his co-authors begin by pointing out that various drugs and other biological manipulations are capable of influencing human behaviour. In so far as human minds supervene on neurological systems with a significant chemical component, this is hardly surprising. What is more tendentious is whether or not any of the interventions they discuss are accurately characterized as affecting ‘moral motivation and behaviour’, let alone as ‘moral bioenhancement’. Timely application of a sedative gas might prevent someone from getting up to mischief, but we would hardly want to characterize this as a ‘moral enhancement’. At the very least, then, ‘moral bioenhancement’ would require modifying both behaviour *and* motivation. Moreover, because—as Savulescu et al. admit—both behaviour and motivations that may be virtuous in one person and/or circumstance may be vicious in another, any biological manipulations touted as making people more ‘moral’ will need to be extraordinarily finely tuned. It is, I think, not incidental to the rhetorical—if not the logical—force of their argument that Savulescu, Douglas, and Persson rely heavily here on a small number of controversial studies involving ‘sexy’ but still determinedly ‘medical’ drugs, such as oxytocin, serotonin, and propranolol, where they might equally well have pointed to the potential for mundane—though regrettably, for enthusiasts for enhancement, illegal—drugs to be used to modify motivation and behaviour: cannabis for tolerance, ecstasy for beneficence, and alcohol for ‘Dutch’ courage. Of course, thinking about these more familiar cases very quickly reveals how unlikely it is that any chemical manipulation is going to reliably improve people’s *moral* reasoning.

Indeed, as Savulescu, Douglas, and Persson themselves later note, ‘one acts morally when one does the right thing, and for the right reason(s)’ (Savulescu, Douglas, and Persson in this collection). I suspect that Savulescu et al. think of acting for the right reasons on an ‘externalist’ model, wherein it is only required that individuals act as would be required by correct moral reasoning in their particular situation whether they actually reason this way themselves or not. It is worth noting, however, that on some accounts of moral action—not least some Kantian accounts—in order to act morally, agents must themselves consciously embrace—if not rehearse—the reasons for their actions. It is hard to see how any drug could alter our *beliefs* in such a way as to track the reasons we

have to act morally. More importantly, the idea that our motivations should track the reasons we have to act is internal to the idea that we are acting *for those reasons*. Such ‘reason tracking’ also requires appropriate sensitivity to counterfactuals. If we praise an individual for helping someone in distress, the judgement that this was morally admirable depends upon the thought that they should not have felt compelled to help them otherwise. It would be a good drug, indeed, that made us feel love only for what is worthy of love and brave only in the service of a just cause.

Ultimately, Savulescu, Douglas, and Persson concede that ‘the technology to biologically influence moral motivation and behaviour is still in its infancy, or even pre-embryonic stages’, yet they insist that ‘it seems likely that science will afford ever more powerful interventions’ and that ‘it would be bold to rule ... out’ the development of this technology (Savulescu, Douglas, and Persson in this collection). The matter of how and why it has become the case that bioethicists feel compelled to discuss the ethics of every hypothetical technology that cannot be shown to be impossible is worthy of an essay in its own right. However, it is clear from this admission that their argument fails the first test that I suggested was necessary for arguments involving thought experiments in applied ethics—it does not illuminate a pressing moral dilemma.

Autonomy and the ethics of ‘moral technology’

Savulescu, Douglas, and Persson then go on to consider the genuinely interesting philosophical question of whether moral bioenhancement would be wrong by virtue of restricting the ‘freedom to do wrong’ and thereby reducing personal autonomy—a criticism that has been put forward (somewhat bizarrely, given his enthusiastic advocacy for other forms of enhancement) by John Harris (2011).

The nature of ‘freedom of the will’ is one of the most ancient and difficult questions in philosophy. Savulescu and his co-authors attempt to bypass concerns about pharmaceutical interventions restricting freedom by highlighting two related cases where we tend not to think that people lack freedom: the ‘naturally virtuous’; and moral education. There is an obvious tension between their description of the naturally virtuous person as someone for whom it is psychologically or motivationally out of the question to do wrong and their later claim that autonomy requires the vivid imagination of alternatives. Nor is it obvious that a pill that renders someone ‘more open to the experiences and lives of others’ achieves the same results as reading Tolstoy, as they suggest. Someone who reads Tolstoy arguably learns *reasons* to be less judgemental and in doing so develops greater understanding: someone who takes a pill has merely *caused* their sentiments to alter. In so far as moral action requires acting for the right reasons, the person who has learned tolerance from Tolstoy has more and better reasons for action.

In any case, the concern that biological interventions to shape motivation and behaviour threaten freedom and autonomy is not exhausted by arguments in metaphysics or philosophy of mind. The tension between God’s omnipotence and man’s freedom is a matter of politics as much as—or perhaps even more than—metaphysics. In asking whether someone is free we are also asking whether they may appropriately be held responsible for their actions.

Savulescu, Douglas, and Persson concede this when they admit that biomedical interventions to reshape the behaviour or motivations of other people constrain their freedom if they cause them to act or feel differently than they would otherwise have been inclined to do. Yet Savulescu et al. try to temper the opposition between freedom and moral enhancement by discussing the case of a hypothetical—and frankly fantastic—‘moral technology’ which would only intervene to prevent people forming the desire to carry out seriously immoral actions. By reducing the number of seriously immoral actions in societies in which it was introduced without significantly reducing people’s freedom, such a device would, they suggest, constitute a powerful technology of moral enhancement. In most cases, they argue, when people chose to act morally they would also be doing so freely: in just a few cases would individuals’ moral choices result from the coercive power of the magical moral technology.

However, Savulescu and his co-authors here underestimate the tension between the power of some and the freedom of others. This tension is highlighted by the notion of ‘freedom as non-domination’ that has been developed by Philip Pettit (1997) in the course of his explorations of the philosophical foundations of republicanism. Pettit argues convincingly that citizens of a society run by a benevolent dictator are, in an important sense, not free even if the dictator is genuinely benevolent and chooses never to exercise his dictatorial powers. To return to the case that so exercised Milton, if God could have intervened to prevent Man’s fall, but didn’t, then God seems as responsible for the fall as were Adam and Eve. God’s power—and not just God’s exercise of his powers—is incompatible with human freedom. Similarly, given that people who are subject to the magical ‘moral technology’ are *not* free to do anything other than act morally, this suggests that there is an important sense in which they do not act freely even when they choose to act in such a way that the technology does not intervene.

Savulescu, Douglas, and Persson’s argument therefore fails the second test that arguments involving thought experiments must pass: it does not convincingly establish their central claim even in the context of the hypothetical technology they discuss.

Having said that, I do want to acknowledge that Savulescu and his co-authors succeed in establishing that biomedical manipulation of oneself is compatible with autonomy and may even promote it by making it easier for us to realize our higher-order goals. This is a not-uninteresting result and should indeed serve to undermine some of the reflexive hostility that the very idea of moral enhancement currently tends to evoke. Nevertheless, the ‘Ulysses and the Sirens’-type cases that Savulescu, Douglas, and Persson discuss are a special case and leave the larger argument about the ethics of enhancing *other* people untouched.

The real world of ‘mind control’

However, it is when we turn to the third test of arguments involving thought experiments in applied ethics—the extent to which their lessons can be applied to real-world cases—that Savulescu, Douglas, and Persson’s argument is most deficient.

In the course of their discussion, Savulescu, Douglas, and Persson admit that previous efforts at ‘mind control’—which include the drugging of children, military interest in

‘brainwashing’, and the history of attempts to ‘cure’ homosexuality—‘were either used in the service of misguided goals or were performed without adequate protections for those subjected to them’ (Savulescu, Douglas, and Persson in this collection). I can see no reasons—and certainly the authors offer us none—as to why either the motivations of governments or the protections they offer the vulnerable are likely to be better in the future.

In the context of a concern for the real-world prospects of any technology that could reliably alter behaviour or motivation, four things stand out in Savulescu, Douglas, and Persson’s paper.

First, their discussion is introduced and justified by reference to a naive conservative account of social phenomena. It simply isn’t true, for instance, to say that the public health problems they refer to in order to motivate the search for technologies of moral enhancement are the result of ‘lifestyle’ choices. They are the result of modern lifestyles but these lifestyles aren’t chosen by individuals so much as imposed upon them by their environments. Moreover, to the extent that it is possible to speak of some individuals choosing these lifestyles, such claims play no role in the explanation of the public health phenomena which result from the aggregate impact of the behaviour of individuals. There are a number of competing (though not, of course, entirely mutually exclusive) explanations for the increase in levels of obesity in wealthy nations since 1985—increases in the sugar and/or fat content of many foods due to developments in agriculture, food processing, and manufacturing; changes in the economics of food production and consumption; changes in patterns of daily activity as a result of lower prices of consumer electronics; changes in intestinal flora as a consequence of increased use of antibiotics; changes in the nature and distribution of paid employment; the development of an urban infrastructure that mitigates against physical activity, etc.—but the claim that this phenomenon can be explained by the idea that lots of people ‘decided to eat more’ is an obvious non-starter. The main role played by such claims about individuals’ responsibility for their health-care condition is in fact to justify the use of coercive measures to reshape behaviour.

Second, there is a pervasive aura of intellectually dubious and politically dangerous sociobiology surrounding their key claim that meaningful biological manipulation of moral behaviour will be possible. Their claim that women are naturally more inclined to self-sacrifice than men is based upon the research of one—controversial—researcher and elides both the possibility of alternative cultural explanations for such differences and the deeper methodological question of how we identify particular behaviours as selfish or selfless. Similarly, the idea that a large percentage of prisoners have a biological disorder that has a tendency to ‘criminal behaviour’ as one of its diagnostic criteria is both circular and brazenly denies a long history of sociological investigations of the social construction and medicalization of deviance. However, more important than the truth or falsity of these individual claims is the way in which construing social problems as rooted in biology both buttresses the existing social order—the foundations of which are clearly primarily ideological rather than biological—and positively invites governments to undertake the project of drugging and electrocuting the poor, the desperate, and the marginalized. Indeed, there is a telling slippage in a sentence that appears early

in the manuscript. The authors say that they aim to explore 'how ethics could justify the use of radical advances in the neurosciences for the purposes of modifying moral motivation and behaviour' (Savulescu, Douglas, and Persson in this collection). I presume they are interested in the intellectual possibility of an argument for moral enhancement. However, another, more cynical, reading would hear this as an invitation to those who feel that they know better than others what sorts of behaviours are desirable to cloak their interests in a facade of ethical argument. Moreover, the project of 'moral bioenhancement' invites this abuse: it assumes that we know what moral behaviour in various circumstances consists in, where in fact this is, within limits, controversial and should also remain so; it will almost certainly involve the powerful acting on the powerless.

Third, the role played by a rationalist account of autonomy in the final section of the paper also invites abuse. If only rational desires count as autonomous then interventions to thwart or reshape the desires of others that one holds to be irrational will not count as infringing their autonomy. It is all too likely then, that governments who wish to manipulate the behaviour of their citizens will mobilize this argument to insist that they do not violate any rights in doing so.

Fourth, by making crucial sections of their case using the imaginary example of their 'moral technology', the authors massively prejudice the argument in favour of their preferred conclusion that moral bioenhancement might be justified even where it does constrain freedom and autonomy. They may well be right about the ethics of their 'moral technology'. However, there is little, if anything, we can conclude about the ethics of plausibly real-world technologies to reshape motivation and behaviour from this idealized case.

Thus, not only does Savulescu, Douglas, and Persson's discussion do little to illuminate the many pressing moral and political concerns about the prospect of biomedical manipulation of motivation and behaviour (in particular, about who would control this technology and to what ends), it significantly misrepresents the relevant moral landscape. Rather than illuminate the ethics of a relevant real-world ethical dilemma, the thought experiment at the heart of their argument obscures it.

Conclusion

I have argued that the vision of moral bioenhancement that Savulescu, Douglas, and Persson put forward is a fantasy. Moreover, insofar as it neglects the many pressing moral issues that would arise the moment any putative technology of moral enhancement actually became available, it is a dangerous fantasy. However, in concluding I want to acknowledge that this interpretation of the import of Savulescu, Douglas, and Persson's argument relies on a substantive account of the proper role of thought experiments in applied ethics, which, I suspect, Savulescu and his co-authors reject. As a piece of philosophy their argument is indeed thought provoking and has significant merits. As a treatment of the bioethical issues we are likely to face in the next century as a result of progress in the neurosciences, however, it is much less compelling.

References

- Harris, J. (2011). 'Moral Enhancement and Freedom', *Bioethics*, 25(2): 102–11.
- Pettit, P. (1997). *Republicanism: A Theory of Freedom and Government*. Oxford: Clarendon Press.
- Savulescu, J., Douglas, T., and Persson, I. (2014). 'Autonomy and the Ethics of Biological Behaviour Modification', in A. Akabayashi (ed.), *The Future of Bioethics: International Dialogues*. Oxford: Oxford University Press, 91–112.