

This is a pre-copyedited, author-produced version of an article accepted for publication in the Journal of Medicine and Philosophy 43(5), 547-567, following peer review. Please do not cite this version.

MORAL ENHANCEMENT, SELF-GOVERNANCE, AND RESISTANCE

Pei-hua Huang

Monash University

Abstract. John Harris recently argues that the moral bioenhancement proposed by Persson and Savulescu can damage moral agency by depriving the recipients of their *freedom to fall* (freedom to make wrongful choices) and therefore should not be pursued. The link Harris makes between moral agency and the freedom to fall, however, implies that all forms of moral enhancement, including moral education, that aim to make the enhancement recipients less likely to “fall” are detrimental to moral agency. In this paper, I present a new moral agency-based critique against the moral bioenhancement program envisaged by Persson and Savulescu. I argue that the irresistible influences exerted by the bioenhancement program harms our capabilities for conducting accurate self-reflection and forming decisions that truly express our will, which subsequently undermine our moral agency.

Keywords: Moral Enhancement, Autonomy, Moral Agency, Freedom, Resistance

I. CONDITIONS OF BEING A MORAL AGENT

In his 2011 paper *Moral Enhancement and Freedom*, John Harris argues that the attempt to use biochemicals, such as oxytocin and selective serotonin reuptake inhibitors (SSRIs), to enhance people's morality could be seriously detrimental to their moral agency. His argument is as follows. Being morally good is not just about doing good deeds. Rather, for an agent to be a moral agent, she must have *choices*—she must have opportunities to decide between morally wrongful acts and morally good acts. An agent who never has any choice but to perform good deeds would not qualify as a moral agent. According to Harris, this is why God made us “sufficient to have stood, though free to fall.”¹ On the religious view, our fallibility does not indicate that God is not omnipotent, and hence *unable* to make us infallible. Rather, we are morally fallible because God granted us this critical freedom, the freedom to fall, in order to ensure that we are morally responsible for our conduct.

According to its advocates, moral bioenhancement will help us become better moral agents by motivating us to act in a reliably morally upright way (Douglas 2008; Persson and Savulescu 2012). In its most extreme form, as depicted by Persson and Savulescu (2012), the behavior of recipients of moral bioenhancement might be transformed, such that when these people are about to harm others their motivation to act in this way is guaranteed to be overridden. In short, the morally bio-enhanced might end up never performing wrongful acts.

Harris thinks that this seemingly desirable result would undermine our moral agency. He thinks moral agency requires the possibility of the agent choosing to act wrongly, because if she is *always* compelled to perform morally right acts, then there is no way to attribute moral credit to

¹ “Sufficient to have stood, though free to fall,” is from John Milton's famous work, *Paradise Lost* (Milton 2007). This phrase is the kernel of Milton's sense of free will. The reformation debate about free will and predestination was framed by Erasmus' (2013) *On Free Will*, and Luther's (2013) *The Bondage of the Will*, among others.

her (Harris 2011, 104). When no one errs, morality becomes an obsolete term; morality is only meaningful when we truly have the freedom to do “immoral things” (Harris 2011, 105).²

There is an important and appealing idea that Harris’s argument is built around, which concerns the deep conceptual links between notions of autonomy, choice, responsibility, and moral agency.³ His view about the relation between these concepts is partly inherited from certain theological notions concerning freedom and evil. Some theistic philosophers have argued that the existence of evil is compatible with the existence of an omnipotent and perfectly good God, because such a being would want to create moral agents, which would require allowing them the genuine freedom to act wrongly and cause suffering (e.g. Plantinga 1977). From this perspective, the freedom to fall is more important than any other freedom we have. Moral agents must have the freedom to contemplate terrible moral wrongs and ultimately *decide* whether they’d like to, or refuse to, realize what they are contemplating (Harris 2011, 103; 2014, 248, 250). Apart from its theological sources, this understanding of the relationship between agency and choice is arguably tied to the very notion of achievement. In general, in order for a person to receive credit for any kind of substantial achievement, she must have been required to exert effort in order to secure that achievement (see Bradford 2015). Thus credit for success is only fully deserved when failure was a plausible outcome.⁴

² Note that Harris gives the freedom to fall argument another interpretation in later papers (See e.g. Harris 2013, 2014, 2014, 2016). There, he argues that the freedom to fall argument is important because some seemingly “wrongful acts” could bring the morally best result in certain circumstances. Making the seemingly morally wrongful impossible, then, renders the morally better outcome inaccessible. I will discuss this interpretation and its implications later in Section III.

³ See also Harris (2014).

⁴ See also Harris (2013), where he argues that “Ethics is for bad guys! The good don’t need ethics, but the truly good in this sense are so few and far between that few of us ever encounter them. Ethics is for those occasions on which altruism fails; or for those people who fail to think and feel, and/or who are not disposed to do as they should!” (p. 171) Thanks to an anonymous reviewer for drawing my attention to this point.

The core of Harris’s concerns—that moral bioenhancement could (at least if it was realized in some of the ways suggested by Persson and Savulescu) undermine people’s moral agency by changing the way that they make moral choices—seems reasonable to me. However, I will argue that his focus on the freedom to fall doesn’t precisely capture *why* moral bioenhancement threatens our moral agency. His analysis in this 2011 work implicitly—and, I will argue, implausibly—suggests that *any* measures which attempt to reduce the likelihood of an agent “falling” pose a threat to her moral agency. Even traditional means of moral education, as advocated by Harris himself, start to seem threatening to our moral agency. In short, his view runs the risk of collapsing the notion of “sinner” and “chooser” into one. In the latter part of this paper I present a view that reformulates the core concerns underlying Harris’s argument, while taking the focus off the freedom to fall as such. Instead, my view emphasizes the importance of self-governance as an *internal* capacity of the agent as opposed to something grounded in the *external* options available for the agent to choose.

II. ALTERNATIVE POSSIBILITIES AND MORAL AGENCY

In this section I argue that what makes an agent morally responsible for her conduct is not given by the *options* among which she chooses but rather the *authorship* of those choices. An idea that is central to Harris’s view is that if agents only have choices that are morally desirable, it makes little sense for us to give them moral credit for their acting in a morally desirable way. However, in many philosophical discussions around moral responsibility and the relevance of alternative choices, the favored view is contrary to Harris’s position. For example, Harry Frankfurt (1969) argues that even when someone has no alternative choices over certain tasks or decisions, she may still be responsible for choosing the only option. Consider the following case, modified from Frankfurt’s argument:

SALARY: Today Sarah received her monthly salary. Part of her salary was paid with a special check that could be donated to a charity or cashed out in casino chips (her company has corporate ties with both organizations). As someone who always puts other people's welfare before personal leisure, Sarah decides to donate her check to the charity, and she does so as soon as she's paid. However, unbeknownst to Sarah, if she had chosen to spend her check at the casino her choice would have been overridden, because there was a secret anti-gambling team following her around, ready to use a special hypnotic device which would have guided her away from the casino (and people targeted by the anti-gambling team's hypnotic device would believe that they themselves made the decision). Since Sarah didn't intend to go to the casino at all, her choice was not a result of the use of the device. This act only would have been a result of the device, if Sarah had chosen otherwise.

In SALARY, although Sarah thought she could have chosen differently (namely cashing out the check in the casino), she did not actually have this choice. Viewing this scenario from Harris's perspective, Sarah could *only* donate to the charity and therefore she does not deserve credit for her donation to the charity. In this highly unusual scenario Sarah would not be able to indulge in personal leisure, no matter what. In a sense, she couldn't *fall*. She could only do what she morally ought to do. However, this doesn't mean that we should disregard how Sarah reached the conclusion that the special check should go to the local charity. All roads lead to Rome, but *which* road we take can make a great difference. In SALARY, the road Sarah took was that she *herself* put the welfare of others ahead of her personal enjoyment, and owing to this moral commitment, she decided, *by herself*, that the check ought to be donated to the charity. The anti-gambling team had nothing to do with the formation of this decision. When emphasizing the causal history of how

Sarah reached her decision, it seems strange to deny that the decision was produced by her and that she's entitled to receive the moral credit.

Although the circumstances depicted in these sorts of Frankfurt-style cases are unusual in real life, they highlight that philosophers in debates about free will and moral responsibility sometimes overlook situations where one would, even given other options, freely or genuinely choose the only option she has. Whether or not an agent is morally responsible for the choice she has made therefore *cannot* be determined just by examining if the agent has other alternatives.

There is some disagreement about exactly why agents under such circumstances should still be deemed responsible for their actions. Frankfurt (1969, 1971) associates moral responsibility with the *will* of agents, arguing that so long as the agent identifies an act with her own willful intentions, then she should bear the moral responsibility for that act. Conversely, according to Fischer's (1982) *actual-sequence* approach, identification is not the key, but the whole causal story. Sometimes, an agent's initial intention may be later superseded by external forces. Under this sort of circumstances, even if the agent eventually decides to accept this change, the agent is not responsible for the result brought by the change—the actual sequence of the whole causal event shows that the real cause is the external forces, not the agent. In any case, whatever theory one endorses to explain why Sarah should be considered responsible for her choice in a case like SALARY, there are strong intuitive reasons for this conclusion.

In response to this, one might point out that in a case like SALARY, it's only by accident that Sarah chooses to do the act that won't subsequently be overridden by the anti-gambling team. One might think that the very fact that the anti-gambling team exists and has the ability to override at least some of Sarah's potential choices is deeply at odds with human autonomy.⁵ On one hand,

⁵ Thanks to an anonymous referee for drawing my attention to this concern.

it's clearly correct to say that external conditions usually play an important role in personal freedom and our moral lives more generally. There are certain understandings of freedom that particularly emphasize on this factor. For instance, under the conception of freedom as non-domination as defended by Philip Pettit (1997), the existence of an external agent who can arbitrarily override a person's choices is fundamentally inconsistent with that person being free. If the person's enjoyment of her freedom is possible only because her superior chooses to let her act as she pleases, then the freedom she is enjoying can be taken away whenever her superior likes, and that—so Pettit and others insist—cannot be counted as *real* freedom.⁶

I do not want to argue against this sort of conception of freedom, or deny more generally that overriding people's choices or removing their options is, in a certain sense, disrespectful to their autonomy. Nevertheless, although these external conditions potentially have some relevance when it comes to judgements about moral responsibility, they are not the *sole* ingredient for personal freedom or moral responsibility. Internal conditions—such as whether the agent under consideration possesses all the important capacities she needs for making her own choice, and whether the agent's choice is actually the product of her exercise of these capacities—are, at least sometimes, more critical for assessing whether or not the agent is a moral agent.⁷ What I want to point out, in short, is that in a case like SALARY, we have good reason to think that the actual causal history matters more when it comes to moral appraisal. We have good reason to think that Sarah is still morally responsible and praiseworthy for her decision to donate the check, despite lacking

⁶ Sparrow (2014) raises this concern. According to Sparrow, moral enhancement operating in a way similar to the anti-gambling team depicted here should not be pursued because such program will inevitably introduce a problematic inequality. In SALARY, for example, people would only be allowed to enjoy their freedom to the extent sanctioned by the anti-gambling team. The very fact that their freedom was “granted” by the team showed that they were in fact dominated and governed. Therefore, despite the possibility of enjoying a certain amount of freedom, the inequality between the team and the ordinary workers makes the moral enhancement program problematic – the former dominates, the latter is dominated.

⁷ Thanks to Justin Oakley and Robert Sparrow for raising this point to me.

the freedom to fall. If the notion of autonomy is grounded in a conception of self-governance, then it seems plausible for us to focus more on the actual causal history of how the decision is shaped.

III. NORMATIVE IMPLICATIONS: INAPPLICABILITY OF MORAL ENHANCEMENT

In the previous section, I argued that under certain circumstances, agents without the freedom to fall can still be responsible for the decisions they make. In this section, I challenge the importance of the freedom to fall from another angle—its wider normative implications.

To begin with, the importance of the freedom to fall, as suggested by Harris, lies in having the freedom to behave wrongly. It is this freedom that gives morality its significance.⁸ However, given a certain understanding of this premise, it potentially has the implication that absolutely *any* attempt to reduce the likelihood of an agent's choosing to do morally wrong deeds poses a threat to that agent's moral agency. This seems absurd because it seems to lead to the conclusion that moral enhancement shouldn't be pursued *in any form*, including the most widely used and least controversial form: traditional moral education. Most of us, including Harris, are not moral nihilists or opponents of education, and believe that moral education is, generally speaking, a good thing. But given those commitments, *prima facie*, it seems that we should deny that reducing the likelihood of someone performing wrong deeds is necessarily a threat to that person's agency, and hence we should deny that the freedom to fall is an essential condition for someone to be a moral agent.

Here is a potential objection to this line of argument. To suggest that moral education “reduces” the student's freedom to fall, and hence threatens her moral agency, is arguably to conflate

⁸ See Harris (2013).

two things. One external condition on the exercise of a person's moral agency is the number of options available to her. Another distinct external condition on the exercise of a person's moral agency is the number of options that she considers to be viable, feasible, or eligible *for her* based on her moral outlook and understanding.⁹ The difference between moral bioenhancement and traditional moral education, so one might argue, is that bioenhancement reduces the number of options available to the agent in any form, whereas moral education only reduces the number of options that the agents considers to be viable *for her* based on her moral outlook. In short, moral education does not threaten the freedom to fall in the way with which Harris was primarily concerned with, and it therefore isn't morally problematic in the way I suggested above.

However, the unwelcome normative implications of Harris's freedom to fall argument still arise. It's true that the number of *available* options and the number of *feasible* options are different, and that conflating the two won't help clarify the subject matter in this debate. However, neither moral bioenhancement *nor* moral education strictly speaking reduce the total number of options available for an agent to choose between. What they do is change the range of options that an agent would regard as viable or feasible for her. If it is an unacceptable violation of the agent's autonomy to use moral bioenhancement to reduce the range of options that she regards as viable or feasible for her, then, other things being equal, it should also be an unacceptable violation of the agent's autonomy to use traditional moral education to reduce the range of options that she regards as viable or feasible for her.

One may argue that this isn't how education works. Education isn't about reducing the range of options that an agent regards as viable, it's about supplying the agent with skills or knowledge. But there are reasons for us to think characterizing education in this way is incorrect.

⁹ Thanks to an anonymous referee for drawing my attention to this concern.

Consider Harris's own view that moral education should be favored because the most dangerous harms one could ever do to others are most likely to result from negligence and idiocy (Harris 2011, 110).¹⁰ For Harris, we often err because we don't really *know* what we are doing. The purpose of moral education, then, is to help us *know* what is right so that we don't do what is wrong out of ignorance.

This line of thought goes back to Plato's argument about *akrasia* (weakness of will) in the Protagoras, where he argues that there is no such thing as weakness of will but rather just a lack of knowledge. If an agent truly *knows* that doing *p* is morally better than doing *q*, then she will not consider doing *q* at all. Doing *q*, under this situation, just isn't an option for her. It's like the multiple choice quiz that has four different "choices" for the question "5+7=?". From the perspective of the mathematical novice, it seems that *there are* options. However, from the perspective of those who understand addition and the meaning of all the symbols, there is *no real* option other than selecting "12". Of course, moral questions are much more complex than this simple question. However, this only means that to *truly understand* what we ought to do is much more difficult than learning elementary mathematics. But still, to the extent that moral education is successful, the number of options the agent considers as real options will be reduced. If Plato is correct, then the person who has received an effective moral education may know that in a certain sense she *could* cheat to gain an unfair advantage, but for her there is simply no question of whether she ever *would* cheat. This is much like the position of the person who has received the kind of moral bioenhancement treatment that Persson and Savulescu and others promote.

¹⁰ Harris uses the word "idiocy," despite its problematic associations with the stigmatization of cognitive impairment. Because of this, it might be better to use a word like "stupidity", but I will follow Harris's terminology for the purpose of this discussion.

The ideal outcome of moral education would be to make those morally concerning “options” disappear from our mind. Confucius reportedly told his students that after he reached the age of seventy, following what his heart desired always resulted in performing right action rather than transgression.¹¹ Unless we want to say that moral saints like Confucius are not moral agents, and that such moral status is undesirable, we should abandon the idea that the freedom to fall is the kernel of our moral agency.

The morally perfect status I’ve just described could raise a further concern. The person who is unable to ever perform any kind of wrong action—whether due to their education or being a recipient of bioenhancement—may find themselves unable to perform an act which, in the particular circumstances, would in fact be the best act for them to perform. Sometimes, due to unfortunate circumstances, the least morally bad option for an agent is to lie, cheat, steal, kill, or do some other kind of act that is usually understood as paradigmatically vicious or wrongful.¹² Of course it’s true that people sometimes find themselves in such situations. Most of us think that lying is generally vicious, but that it would be praiseworthy (or even obligatory) to lie in order to, say, help innocent Jewish people escape Nazi Germany. However, *this* sort of freedom to fall is not the kind of freedom I am discussing in this paper. What I am interested in is the freedom to do things that are morally wrong *in the given moral context*. This “falling” doesn’t refer to any particular act that was considered vicious in *most situations*, but rather to an act that is truly morally wrong *in the situation at hand*. The ideal moral agent that I’ve described above would have highly refined practical wisdom; other things being equal, such an agent would, if morally required in a particular situation, be disposed to perform acts that are conventionally recognized as vicious.

¹¹ See *The Analects*, Book II Wei Chang.

¹² The discussion of this sort of case can also be found in Chan and Harris (2011) and many of Harris’s works, e.g. Harris (2011, 2013, 2014, 2014, 2016)

What this challenge really achieves is a reminder that moral scenarios are often highly complex. Hence if “moral enhancement” programs are not designed in a context sensitive way, they may actually lead people to become morally inferior.¹³ It is also somehow strange to hold the view that people who are in a situation where they need to lie in order to help others are doing something morally correct while arguing that they are nevertheless “acting wrongly.” After all, if we think the moral valence of “lying” could sometimes be positive, then how could we say “lying” is “falling”? What we should say is that no act is intrinsically vicious and that the moral valence of an act depends on the context.¹⁴

The challenge I launched in Sections II and III is a *reductio ad absurdum* against Harris’s view of what moral agency requires. If the essence of our moral agency is the freedom to fall—*i.e.* the ability to make wrongful choices—then, as it stands, it seems that every moral development that removes our ability to make a wrongful choice diminishes our moral agency, because it reduces the extent to which we are “free to fall.” But that seems absurd. The kind of freedom that is essential to our moral agency can’t be about having a maximally wide range of options to act wrongfully. This means that Harris cannot retain both his freedom to fall argument and his support of moral education and cognitive enhancement and at the same time.

IV. A NEW MORAL AGENCY-BASED CRITIQUE OF MORAL BIOENHANCEMENT

The difficulties of Harris’s argument indicate that abandoning the idea of freedom to fall as a way of thinking about moral agency may be sensible. Nevertheless, this does not mean that the kind of moral bioenhancement program proposed by Persson and Savulescu poses no threat to our freedom

¹³ See Harris (2014) and Huang and Tsu (forthcoming).

¹⁴ For more discussions of moral particularism and moral enhancement, see Pugh (2017) and Huang and Tsu (forthcoming).

and our moral agency. It only shows that an adequate expression of concern about that threat may need to appeal to a different conception of autonomy.

Let's briefly review Persson and Savulescu's proposal (2008, 2012). Unlike Harris, they think that most moral tragedies do not result from negligence and idiocy, but rather, primarily, from people's bad motivations: (i) a lack of motivation to do the things we know we *ought to* do (e.g. failing to take measures to curb catastrophic climate change), or (ii) a malicious motivation to do very harmful things (e.g. use weapons of mass destruction or engage in terroristic violence). They argue that because many studies show that our moral shortcomings are deeply linked to our neurological wiring, biochemical interventions could provide solutions for these motivational problems. For instance, the reason most people living in the developed world aren't motivated to change their lifestyle to ease the rise of global temperatures might be that the sheer number of victims are beyond their ability to empathize with (Dunbar 1992, 2008). Simon Baron-Cohen (2004, 2012) also offers evidence to support the view that there are neurological factors behind females' stronger altruistic tendencies; the difference seems to be largely determined by the different levels of oxytocin in males and females. Crockett et al. (2008) also found that people's responsiveness towards unfairness and injustice is linked to the levels of serotonin in their bodies.¹⁵

¹⁵ However, it must be noted that whether this biochemical approach would succeed is unclear. Indeed, many studies suggest that non-biochemical enhancement programs have their limits. For instance, a study by Dunbar (2012) found that even with a powerful tool like Facebook at our disposal, data of users' interactions with their connections show that most people can still only maintain around 150 social relationships at a time, regardless of the number of Facebook connections they have. This study suggests that non-biological means aiming at enhancing our ability to maintain social relationships have very little effect. Yet this study does not suggest that therefore using a biochemical approach would be more fruitful. We don't know enough about the operation in the brain of the biochemicals that might potentially to enhance our abilities. De Dreu et al.'s studies (2011; 2015), for instance, found that the type of altruism promoted by oxytocin seems to be "parochial" and "ethnocentric". Recent research has also cast doubt on whether the effect found in the oxytocin studies was real. Walum, Waldman, and Young (2016) point out that studies of intranasal oxytocin are generally unable to provide sound statistical conclusions due to the limited sample size. Lane *et al.* (2015) also announce that their attempt to reproduce a well-known experiment on the effect of oxytocin failed. For the purpose of this paper, I will not discuss these findings in greater details. My aim in this paper is, after all, to answer the hypothetical question "If we *can* use biological means to modulate people's moral dispositions, are we morally permissible to do that?" The empirical issue should be left to psychologists and neuroscientists.

If we could have a compulsory moral bioenhancement program that directly targets these root causes of our moral failings, then presumably we could curb climate change or reduce the risks of terrorist attacks. Increasing the general population’s oxytocin or serotonin levels, for instance, may be helpful for achieving these goals because the aforementioned studies show that “manipulations of biology can have moral effects” (Persson and Savulescu 2012, 121). Once we have a group of experts who can design a compulsory moral bioenhancement program, wrongful motivations may disappear, and “violence” and “conflict” may become words that only exist in dictionary.¹⁶

The very *motivation* for this proposal, however, shows that Persson and Savulescu’s program would, by superseding our role as self-governed beings, be harmful to our moral agency.¹⁷ The Frankfurt-style argument I presented in Section II shows not just that Harris’s freedom to fall argument does not work, it also shows that whether an agent is morally responsible for her acts depends, in a crucial way, on the expression of her *will*. This suggests that if the agent’s decision-making is “contaminated” by outside forces—in a way that she cannot (*i*) properly *reflect* on, and (*ii*) *decide* whether or not she would like to engage with—then her authorship of the act is undermined. The moral bioenhancement program proposed by Persson and Savulescu, because its success relies on altering our morally inferior motivations to those morally superior ones, is therefore very likely to damage our capacities for self-governance.

¹⁶ However, even if these biomedical interventions are truly effective in altering people’s behavior, whether these large-scale problems could be solved in this bottom-up approach is still unclear. See Handfield, Huang, and Simpson (2016).

¹⁷ Here I use self-governance and autonomy interchangeably.

Self-Reflection and Moral Agency

Although philosophers have not reached a consensus about the exact conditions for being self-governed, there is widespread support for the idea that self-reflection is at least an important element of being a self-governed agent. According to one way of understanding the importance of self-reflection, moral agents are agents who can reflect upon their own views of how they would like to be treated and can then deliberate about a set of moral rules based on such reflection, in order to govern their own conduct (Kant 1998; Darwall 2006). Self-reflection also plays a major role in the hierarchical approach to autonomy advocated by Frankfurt (1971). He argues that to see if an agent's desire to do x is an autonomous desire, we must check if it is coherent with the agent's other higher-level desires, e.g. the desire to desire to do x . If the agent herself would want to examine whether her desire to do x is truly an autonomous desire, then she must, again, reflect upon her own other higher-order desires. In recent discussions of relational autonomy, some feminist philosophers also argue that part of what makes an oppressive society problematic is that it impairs the essential abilities required in order for one to be autonomous, with self-reflection being one of these impaired abilities (Wolf 2000; Barclay 2000).

It is not difficult, then, for us to see why self-reflection is important to moral agency. To be a moral agent who can be responsible for her acts, person must have the ability to reflect upon her reason for performing those acts and distinguish her own reason from other influences which might be alien to her. It is this particular ability that distinguishes most humans from most non-human animals.¹⁸ For instance, primates like bonobos exhibit altruistic behaviors towards other members of their group (and even to non-members like birds), and they also exhibit "tit-for-tat" revenge

¹⁸ I use "most" here because I do believe some *homo sapiens* do not have this ability. By this I mean, I agree with philosophers like Singer that there is a difference between human and person. "Human" is merely a biological category that denotes members of *homo sapiens*; "person" denotes individuals capable of moral reasoning. However, I remain agnostic about whether this difference will imply that mere humans deserve less respect than persons.

behaviors when they are treated unfairly. Yet it is difficult for us to conclude from these findings that they are capable of moral reasoning and should be seen as moral agents, because we don't know whether their altruistic and tit-for-tat tendencies are purely expressions of their biological instincts. Due to communicative limitations in our ability to interpret animal behavior, we might never know whether bonobos are able to conduct self-reflection on their own reasons for why they "decide" to help or punish others.¹⁹

As well as the bare ability to reflect, another important element of moral agency is that the agent can *accurately* reflect upon her reasons and discern her own reasons from other influences that might turn out to be alien to her. Agents under the influence of hypnotism, psychosis, or brain-washing usually are not regarded as fully responsible for what they have done, because they are unable to accurately reflect on their decision-making process for the relevant acts. At best, they can only be *causally* responsible for what they have done.²⁰

Now we can return to Persson and Savulescu's proposal and examine whether it can work without harming our self-governance in these areas. Consider the following scenario:

EXPERIMENT: Andy is keen to participate in psychology experiments. To help advance human understanding about our complex minds in the most realistic setting, he authorizes all the labs in his city to conduct experiments on him without informing him. He also grants scientists access to his personal diary where he writes down his reflections about his acts in order to make his subjective understanding accessible to the scientists. A group of scientists decides to run a month-long experiment on Andy to test whether a newly discovered substance, *auxilio*, can increase people's altruistic tendencies. The group secretly uses him as a subject in a randomized controlled trial. At

¹⁹ For more discussion, see De Waal (2009). Although this is largely about how evolution can explain the origin of morality, it provides many interesting primate cases which suggest that some primates might be able to do moral reasoning and therefore should be seen as moral agents and therefore as persons.

²⁰ See Vincent (2013) for discussion on the distinction between different types of responsibility.

no point does Andy find out about his involvement in the study, and thus he attributes all of his choices to his own deliberation. The experiment finds that Andy did in fact behave more altruistically under the (secret) influence of *auxilio*.

In this case, Andy's choices seem morally better when under the influence of *auxilio*, but whether we could say that Andy is in fact the *author* of these choices is uncertain. Because of the protocol of the experiment and how *auxilio* functions, Andy is unable to tell whether his desire to help is truly his own. This means that unless the research team approaches him and tells him about *auxilio*, he won't know that some of his behavior is influenced by this treatment. Unfortunately, the biochemicals for the moral bioenhancement proposed by Persson and Savulescu, such as oxytocin and SSRIs, work much like the *auxilio* in this scenario—they can alter people's behavior without engaging with people's self-reflection. Using these substances for such purposes, then, runs a serious risk of damaging people's autonomy.

The importance of proper self-reflection is not just a theoretical concern for self-governance. It plays an important role in real-life rulings as well. Consider the case of Terrence Lewis Martin, a former Minister of Parliament in Tasmania. Mr. Martin was charged and found guilty of having sexual intercourse with a 12 year-old girl and producing child exploitation material (Martin 2011). At a later point, however, it became clear that his hypersexuality was induced by the dopamine replacement treatment (DRT) he received for treating symptoms of his Parkinson's disease. While DRT has been found to cause hypersexual desire to its recipients, patients are often unable to tell whether the development of the hypersexuality is an authentic desire. If the patients are not warned of this potential side-effect, they may begin to identify with their abnormal sexual activities. Failing to reflect upon the real causal story behind the seemingly authentic desire and their decision making, makes some patients "choose" not to tell their doctors. Some of them feel ashamed and

embarrassed about *who they really are*, which exacerbates the overall situation (Carter, Ambermoon, and Hall 2011). In the case of Mr. Martin, the court eventually ruled that his 10-month prison sentence should be suspended. The underlying rationale for this ruling seems to be that because Mr. Martin's behavior was changed in a such way that he could not accurately reflect upon his decision making, he should have reduced accountability for his hypersexual activities.²¹ Although the case of the DRT and the moral bioenhancement program proposed by Persson and Savulescu produce different effects, the fact that both *bypass* rational scrutiny is equally worrisome.

The Making of Decisions and Freedom of Resistance

The ruling in Mr. Martin's case, however, likely was not made just because he couldn't conduct accurate self-reflection. More important is that due to an influence he was unaware of, he could not make his own decisions about what he ought to do. It's fair to think that should Mr. Martin have known how his behavior could be shaped in this way, he would have asked the doctors not to prescribe him the DRT, or cut down the DRT once he developed signs of hypersexuality. However, in reality, Mr. Martin was not in a position to make this decision. Falsely identifying with the influence introduced by DRT led to a tragedy Mr. Martin would not have wanted to cause. The same issue can be seen with Andy in EXPERIMENT. Since the experiment was run secretly, he could not accurately self-reflect on his behavior and decide whether or not he would like to keep behaving the way he did when he was on *auxilio*. If the team decides not to cease dosing him with *auxilio*, then Andy might end up being a puppet of *auxilio*. In other words, in both cases, Mr. Martin's and Andy's ability to form their own decisions was undermined by their lack of knowledge of the influencing factors.

²¹ For more discussions on DRT and moral responsibility, see Vincent (2012).

However, not *all* external influences undermine our decision-making ability in the way that *auxilio* and DRT do. In fact, many feminist philosophers of autonomy have argued against the traditional view that autonomous decisions must be made in a way that is *completely* free from external influences.²² A brief review of how we become *who we are* reveals that we are not truly “self-made” or “self-created.” Instead, who we are is at least partly, if not wholly, made up by our cultural background, our language, socioeconomic situation, and other environmental factors.²³ Such a view does not argue that all our decisions are the product of collective influences outside us and that there is therefore no room for us to say we are the authors of our decisions. Rather, what this view implies is that we must develop a finer account to distinguish different types of influence—one that can explain why the influence exerted by genuine moral education *strengthens* rather than undermines our ability to make autonomous decisions, whereas the influence exerted by oppressive norms *harms* our agency.

Employing the concept of resistibility to reformulate the issue is, I argue, the best way to clarify why many of us believe that we can still be autonomous, even while acknowledging the inescapability of some external cultural influences. To be self-governed, we do not need to be entirely free from social influences and our own personal causal histories. What we need is the possibility to *rationally resist* these influences. This requires the exercise of our capacities for self-reflection, but it also requires external social conditions that shield individuals from certain forms of intellectual coercion, oppressive influences, and ideological indoctrination. Our upbringings inevitably influence much of the shaping of our worldview. However, under appropriate social conditions, we are still able to critically reflect upon those influences, and once we reach maturity,

²² Kant (1998) and Rawls (1971) are, arguably, two of the most famous thinkers who endorse this sort of individualistic conception of autonomy.

²³ Some authors in the relational autonomy literature doubt that it is truly possible for us to be self-governed; for further discussions, see Mackenzie and Stoljar (2000).

decide if revising our existing worldview is a more sensible decision. Call this the *freedom of resistance*.²⁴

This conception of freedom explains why many argue that women’s ability to make “autonomous decisions” can be harmed by oppressive social norms. Such norms are often imposed in a way that discourages women from using their rational capacities to scrutinize the ideology behind norms and accurately reflecting on whether they truly wish to live in accordance with those norms. For instance, problematic norms could be obscured by *prima facie* sophisticated theories (e.g. using the language of social Darwinism or fatalism to justify unfair working conditions) or protected by hiding facts that could undermine the norms (e.g. the fact that women can live independently and pursue their own aspirations). Under these conditions, the freedom of resistance is seriously damaged. Many puzzling cases of autonomy discussed by feminist philosophers, such as Stockholm syndrome, may be seen as instances of the impairment of this freedom.

While the goals of the moral bioenhancement envisaged by Persson and Savulescu differ in important ways from these oppressive norms, the influence exerted by the bioenhancement program nevertheless appears to be similarly detrimental to the freedom of resistance. For example, research conducted by Crockett et al. (2010) found that participants’ decision-making in trolley problem cases could be altered simply by dosing them with SSRIs: under their influence, many participants’ preference for killing one to save five was overridden.²⁵ We could imagine that there

²⁴ Please note that by freedom of resistance, I do not intend to argue that there *must* be alternative options for one to possess this freedom. For example, there is no alternative answer to the question of the simple addition of five plus seven. Rather, the point I want to make with this conception of freedom is that it is vital for us to have the freedom to *rationally* and *critically re-evaluate* the influences that are imposed on us in the formation of our will and preferences. Even if I only have one option available to me, my will may still oppose this choice. Someone’s freedom of resistance is violated not when their range of alternative options is reduced, but only when their ability to govern their own *will* is overridden and they are prevented from inspecting and interrogating the influences that shape their will. This is especially important if we believe in moral pluralism—there could be several equally desirable values, moral principles, and it is morally permissible to prioritize a particular value over others.

²⁵ We don’t have a clear idea about how SSRIs alter our behavior yet. Currently, it is hypothesized that SSRIs can promote harm aversion, or more specifically, the aversion to harm others *directly*. Due to such an aversion, people

could be people who are *secretly* prescribed SSRIs due to their refusing to proactively participate in moral bioenhancement, and whose choices in trolley problem type cases are altered because of this in a way that they don't ever consciously recognize. The worrisome effects on the freedom of resistance brought about by SSRIs are similar to oppressive social norms; they both erode people's autonomy by undermining their ability to conduct accurate self-reflection and make decisions with the freedom of resistance.

Moral Education and the Preservation of Self-Governance

Moral education as a means of improving our moral behavior does not necessarily pose the same kind of threat, nor the same degree of threat, to our self-governance. This is because methods of education at least sometimes allow people to scrutinize and interrogate the moral principles and theories they are being encouraged to adopt. In other words, the educational content delivered by moral educators generally don't *directly* determine the attitudes and behaviors of the educated. Rather, these contents need to be digested via the rational faculties first, at which point one might find them unconvincing or inconsistent and therefore refuse to follow the teaching.

This is part of the reason why moral education is not always effective. Sometimes, it might be that the educators use unpersuasive arguments for an otherwise good moral principle. Sometimes, it might be that the moral system advocated by the educators is too different from the pre-reflective commitments of the people being educated. The recent effective altruism movement is one instance of this. Despite the persuasiveness and clarity of the arguments of people like Peter

under the influence of SSRIs will prefer options that do not require them to directly harm others. In the classical trolley problem, this would be "not to pull the lever and change the course of the trolley." As a result, even though the default course of the trolley will kill five people, and doing nothing, in a sense, indirectly harms the five, people dosed with SSRIs would still prefer the option that does not require them to directly harm others. For more details, see Crockett et al. (2010).

Singer (1972), most people don't choose to start working in Wall Street in order to maximize the money they can donate to the poor. This reflects the more general point that good education doesn't just "implant" ideas in people's heads. Rather, it deeply engages with their rational processes. Good moral education influences an educated person's decision-making by providing her with different moral reasons and arguments thereby deepening her understanding of the moral reasoning behind her decisions. But it does not attempt to bypass her personal reflection or fully determine the content of her will. It allows that her decisions are still *her* decisions.

This feature also suggests that even if moral education is abused by authorities, as any other technologies and regulations can be, the harm will generally be less severe than what might be brought about by the kind of bioenhancement program that Persson and Savulescu propose. After all, we have seen many North Koreans—and people in similar situations of extreme political propaganda—choose to defect. Presumably, such people's education would have tried to indoctrinate them with patriotism and hatred for non-communist countries, and indeed, their freedom of resistance and other important faculties would have been seriously harmed. But their defection suggests that even under the worst abuses of education, some people can still retain a degree of freedom to resist the ideologies imposed on them.²⁶

²⁶ The case of North Koreans' defection is not to argue that an abuse of education can never be successful, nor does it argue that there is a clear boundary between moral education and indoctrination. Indeed, well-fabricated propaganda can seriously harm our ability to make autonomous choices (Brison 2000; DiPaolo and Simpson 2015). DiPaolo and Simpson also point out that methods widely used in indoctrination can also be seen in ordinary educational settings. For instance, most parents *teach* toddlers not to touch hot stove not by providing rational arguments ("touching hot stove will burn you, and you don't want to feel the burn, so don't touch it"), but by deterrence ("DON'T TOUCH IT!"). See DiPaolo and Simpson (2015, 8-10). The point I want to make here is that unless we are total nihilists about ideals of rationality and the purpose of education, a distinction between indoctrination and education is still worth making.

V. POTENTIAL CHALLENGES

In the previous section, I focused on explaining what truly makes Persson and Savulescu's moral bioenhancement dubious. I argued that the bioenhancement is worrisome mainly because of its potentiality to damage our self-governance. I first pointed out that the affective nature of the bioenhancement makes it difficult for the enhanced to correctly reflect on their decision making, which in turn make it difficult for us to attribute them with moral responsibility for what they have done. I then argued that despite the fact that we are influenced by numerous effects, not all of them are the same. Those that may become irresistible and dominate our decision making are particularly worrisome. This particular concern draws the fundamental distinction between moral education and the moral bioenhancement proposed by Persson and Savulescu.

Here, I would like to consider about two more potential challenges that proponents of this particular type of moral bioenhancement may raise against my account.

Status Quo Bias

Status quo bias is a bias that irrationally favors existing conditions, leading to a reluctance or refusal to accept change. Bostrom and Ord (2006) argue that many arguments against new technologies stem from this bias. For instance, in the debate on human enhancement, conservatives often cannot provide good arguments to show that we are already optimal and therefore do not need any improvement. Their reluctance to use new technology to improve human capacity, as Bostrom and Ord argue, only shows that they are irrationally attached to the current conditions. Persson and Savulescu may argue that my argument against their moral bioenhancement program also stems from a status quo bias. They may argue that, indeed, self-governance is important for moral agency, yet we don't have sufficient proof to argue that moral bioenhancement will seriously damage self-

governance. Hence once we boil the argument of resistance down, we would find a status quo bias in disguise.

This is not an implausible argument, but it backfires. Clark (2016) points out that the core of status quo bias is that if someone cannot provide a good rational argument to support her refusal to change, it is very likely that the account merely reflects her bias toward the current situation. However, what if those who propose the change fail to provide a rational argument to support it? Clark argues that proponents of using new technologies to modify human capacities cannot rely on status quo bias to refute their detractors as the new technologies under consideration may be too *new* to us. We are not, therefore, in a position to assess whether they can bring about the results promised by their supporters. Nor are we in a position to confirm the worries of conservatives. Indeed, whether or not Persson and Savulescu's program will completely deprive, or only partially damage, self-governance is currently unclear. The DRT case I cited could at best tell us that some biochemical treatments could seriously alter their recipients' behavior and damage their ability to be self-governing. It is not a direct case against the use of oxytocin nasal spray or SSRIs. However, a similar problem also occurs with Persson and Savulescu's proposal. Currently, we don't know exactly how effectively these chemicals could influence our moral deliberation. We also don't know whether using them long-term will cause changes that leave irreversible harms to our rational faculties. Thus arguments built on status quo bias are, on this topic, uninformative.

Levels of "Moral Chemicals" and Freedom

Some may argue that my account implies that people who were born with higher levels of oxytocin are less self-governing than those who were not. This line of thought comes from interpreting my argument for self-governance as a rejection of higher levels of oxytocin. In fact, in *Unfit for the Future*, Persson and Savulescu have propose an argument related to this point. They argue that if

raising the level of oxytocin will bring about a detrimental effect on people's freedom, then it seems to imply that whoever has higher levels of oxytocin is less free than others (Persson and Savulescu 2012, 112). They argue that it is absurd for us to think that females are less free than males just because of their oxytocin levels. Rather, we tend to think that females and males enjoy the same level of freedom. *Therefore*, it would seem that raising people's oxytocin level to make them more altruistic does not make them less free.

Persson and Savulescu are right that it is absurd to think that females with higher oxytocin levels are less free than males who enjoy lower levels of oxytocin. However, the fact that females have higher oxytocin levels does not imply that it is morally permissible for us to raise the oxytocin levels for those who do not have high oxytocin levels.

Consider the case of height. People who are taller than 180 cm do not enjoy less freedom than those who are below 180 cm. In many societies, being taller than 180 cm also means having a better chance at commanding respect as a leader. In short, it is good to be tall, at least in most societies. However, no matter how great being a tall person could be, attempting to *force* someone to grow taller *against her will* is worrisome. It is not that being tall is not a desirable property; what makes this practice concerning is *forcing* someone whose height is below 180 cm to grow taller. Similarly, having higher level of oxytocin does not render one less free or less self-governing. Instead, it is using pharmaceutical interventions (e.g. raising or lowering oxytocin levels) to alter people's moral decision-making in a way that does not engage with their rational faculties that threatens their self-governance.²⁷

²⁷ However, there are some exceptions concerning the levels of oxytocin and self-governance. Abnormally high or abnormally low levels of oxytocin can impair the functioning of a person's rational faculties. This is also true for serotonin and dopamine.

Potential Reconciliation of Freedom to Fall and Freedom of Resistance

The last challenge I consider here is that there might not be a substantial difference between Harris's freedom to fall and my freedom of resistance. The upshot of this challenge is that both accounts seem to stress the importance of having alternatives, and both accounts try to argue for the use of traditional moral education while rejecting Persson and Savulescu's proposal. It's true that there are some similarities between the arguments made with the freedom to fall and those I made with the freedom of resistance. However, my argument is an attempt to show that there is a way to capture much of the spirit of Harris's concerns, *without* appealing to the idea of the freedom to fall. Moral agency doesn't depend on the possibility of acting wrongly, but on the possibility of resisting influences.

Concerning the question of how to view external influences, for instance, the two conceptions of freedom give us different answers. If we try to answer this question with the freedom to fall, then what we should worry about will be whether the influence under consideration excludes the possibility of our acting wrongfully. However, the freedom of resistance focuses more on whether the influence allows critical reflection in the formation of one's will, and on one's response to the influences that shape one's will. Under this view, trying to use moral enhancement to help people become "moral saints" would not be problematic, as long as the influences exerted by the enhancement program could be reflected upon, and more importantly, so long as people were allowed to make the final decision on what sort of moral saints they would like to be. In this case, "the alternatives" that really matter for the freedom of resistance are not about the option to perform wrongful acts. Rather, the alternatives considered here are about different *goods*. To the extent that we believe in some form of moral pluralism, it is important to let people decide how to prioritize among possible moral values, and not to impose upon them the view that there is only

one conception of the good that is worth pursuing, or only one order of moral priorities that is acceptable. In other words, the alternatives are alternatives to do good.²⁸

VI. CONCLUSION

There are different accounts of what makes one moral agent, and we could count the freedom to fall as one. Yet, as I've argued in this paper, in the discussion on moral enhancement, this view is not particularly helpful. I think that what we care about is whether the moral enhancement program employed by a third party can improve enhancement recipients' morality while respecting their freedom to be self-governed.

The analysis and critique I presented here further suggests that what makes some moral bioenhancement programs seem suspicious is not that they employ biochemical intervention. We use many biochemical interventions, such as consuming tea or coffee, to strengthen our self-governance and boost the performance of our rational faculties. Rather, Persson and Savulescu's proposal is suspicious because the influences it proposes bypass our rational scrutiny and further deprive us of the opportunity to make our own moral decisions. A similar test can be run on enhancement programs that do not involve biochemical intervention, instead using "traditional means". We may find some of them (e.g. indoctrination) equally worrisome to enhancement recipients' moral agency. After all, our examination does not stop at examining the means a moral enhancement program employs; what we care about is the eventual effects the program will bring to our moral agency.

²⁸ Thanks to an anonymous referee for pushing me clarify this point.

ACKNOWLEDGEMENT

I thank Robert Simpson for his generous support and invaluable discussions during the drafting of this paper. I also thank Linda Barclay, Lucy Mayne, Andrew McLoughlin, Justin Oakley, Julian Savulescu, Robert Sparrow, three anonymous referees for the Journal of Medicine and Philosophy, and participants of Designing Moral Technologies in Ascona, Switzerland, the 2016 Annual Conference of Society of Applied Philosophy at Queen's University, Belfast, UK, and Morals and Molecules workshop at Monash University, Australia for providing me with feedback on earlier version of this paper.

REFERENCES

- Barclay, L. 2000. Autonomy and the Social Self. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by C. Mackenzie and N. Stoljar. NY: Oxford University Press.
- Baron-Cohen, S. 2004. *The Essential Difference*. UK: Penguin Books.
- . 2012. *The Science of Evil: On Empathy and the Origins of Cruelty*. NY: Basic books.
- Bostrom, N., and T. Ord. 2006. The Reversal Test: Eliminating Status Quo Bias in Applied Ethics. *Ethics* 116 (4):656.
- Bradford, G. 2015. *Achievement*. UK: Oxford University Press.
- Brison, S. 2000. Relational Autonomy and Freedom of Expression. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by C. Mackenzie and N. Stoljar. NY: Oxford University Press.
- Carter, A., P. Ambermoon, and W. D. Hall. 2011. Drug-induced Impulse Control Disorders: A Prospectus for Neuroethical Analysis. *Neuroethics* 4 (2):91-102.

- Chan, S., and J. Harris. 2011. Moral Enhancement and Pro-social Behaviour. *Journal of Medical Ethics* 37 (3):130-131.
- Clarke, S. 2016. The Reversal Test, Status Quo Bias and Opposition to Human Cognitive Enhancement. *Canadian Journal of Philosophy* 46 (3):369-386.
- Confucius. 1979. *The Analects*. Translated by D. C. Lau. NY: Penguin Books.
- Crockett, M. J., L. Clark, M. D. Hauser, and T. W. Robbins. 2010. Serotonin Selectively Influences Moral Judgment and Behavior through Effects on Harm Aversion. *Proceedings of the National Academy of Sciences* 107 (40):17433-17438.
- Crockett, M. J., L. Clark, G. Tabibnia, M. D. Lieberman, and T. W. Robbins. 2008. Serotonin Modulates Behavioral Reactions to Unfairness. *Science* 320 (5884):1739-1739.
- Darwall, S. 2006. The Value of Autonomy and Autonomy of the Will. *Ethics* 116 (2):263-284.
- De Dreu, C. K. W., L. L. Greer, G. A. Van Kleef, S. Shalvi, and M. J. J. Handgraaf. 2011. Oxytocin Promotes Human Ethnocentrism. *Proceedings of the National Academy of Sciences* 108 (4):1262-1266.
- De Dreu, C. K. W., H. S. Scholte, F. A. van Winden, and K. R. Ridderinkhof. 2015. Oxytocin Tempers Calculated Greed but not Impulsive Defense in Predator-prey Contests. *Social Cognitive and Affective Neuroscience* 10 (5):721-8.
- De Waal, F. 2009. *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.
- DiPaolo, J., and R. M. Simpson. 2015. Indoctrination Anxiety and the Etiology of Belief. *Synthese* 193 (10):3079-3039.
- Douglas, T. 2008. Moral Enhancement. *Journal of Applied Philosophy* 25 (3):228 - 245.

- Dunbar, R. I. M. 1992. Neocortex Size as a Constraint on Group Size in Primates. *Journal of Human Evolution* 22 (6):469-493.
- . 2008. Mind the Gap: or Why Humans Aren't Just Great Apes. *Proceedings of the British Academy* (154):403-23.
- . 2012. Social cognition on the Internet: testing constraints on social network size. *Philosophical Transactions of the Royal Society of London. Series B* 367 (1599):2192-201.
- Erasmus, D. 2013. On Free Will. In *Desiderius Erasmus and Martin Luther*, edited by E. Winter. UK: Bloomsbury Academic.
- Fischer, J. M. 1982. Responsibility and Control. *The Journal of Philosophy* 79 (1):24-40.
- Frankfurt, H. 1969. Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy* 66 (23):829 - 39.
- . 1971. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68 (1):5-20.
- Handfield, T., P.-H. Huang, and R. M. Simpson. 2016. Climate Change, Cooperation, and Moral Bioenhancement. *Journal of Medical Ethics* 42 (11):742-747.
- Harris, J. 2011. Moral Enhancement and Freedom. *Bioethics* 25 (2):102-111.
- . 2013. 'Ethics is for bad guys!' Putting the 'Moral' into Moral Enhancement. *Bioethics* 27 (3):169-73.
- . 2014. "... How Narrow the Strait!". *Cambridge Quarterly of Healthcare Ethics* 23 (03):247-260.
- . 2014. Taking Liberties with Free Fall. *Journal of Medical Ethics* 40 (6):371-4.
- . 2016. *How to be Good*. Oxford, UK: Oxford University Press.

- Huang, P.-H., and P. S.-H. Tsu. forthcoming. Biomedical Moral Enhancement in the Face of Moral Particularism. *Royal Institute of Philosophy Supplements*.
- Kant, I. 1998. *Groundwork of the Metaphysics of Morals*. Translated by M. Gregor. Edited by K. Ameriks and D. Clarke, *Cambridge Texts in the History of Philosophy*. UK: Cambridge University Press.
- Lane, A., M. Mikolajczak, E. Treinen, D. Samson, O. Corneille, P. de Timary, and O. Luminet. 2015. Failed Replication of Oxytocin Effects on Trust: The Envelope Task Case. *PLoS One* 10 (9):e0137000.
- Luther, M. 2013. The Bondage of the Will. In *Desiderius Erasmus and Martin Luther*, edited by E. Winter. UK: Bloomsbury Academic.
- Mackenzie, C., and N. Stoljar. 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. NY: Oxford University Press.
- Martin, T. 2011. Tasmania v. Martin 2010 TASSC 51(Austl.).
- Milton, J. 2007. *Paradise Lost*. UK: Blackwell Publishing.
- Persson, I., and J. Savulescu. 2008. The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity. *Journal of Applied Philosophy* 25 (3):162-177.
- . 2012. Moral Enhancement, Freedom, and the God machine. *Monist* 95 (3):300 - 421.
- . 2012. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press.
- Pettit, P. 1997. *Republicanism: a Theory of Freedom and Government*. Oxford, UK: Oxford University Press.
- Plantinga, A. 1977. *God, Freedom, and Evil*. Wm. B. Eerdmans Publishing.

- Plato. 1961. Protagoras. In *The Collected Dialogues of Plato*, edited by E. Hamilton and H. Cairns. Princeton: Princeton University Press.
- Pugh, J. 2017. Moral Bio-enhancement, Freedom, Value and the Parity Principle. *Topoi*.
- Rawls, J. 1971. *A Theory of Justice*. Original Edition ed. U.S.: Harvard University Press.
- Singer, P. 1972. Famine, Affluence, and Morality. *Philosophy & Public Affairs* 1 (3):229-243.
- . 1993. *Practical Ethics*. 2nd ed: Cambridge University Press.
- Sparrow, R. 2014. Better Living Through Chemistry? A Reply to Savulescu and Persson on ‘Moral Enhancement’. *Journal of Applied Philosophy* 31 (1):23-32.
- Vincent, N. 2012. Restoring Responsibility: Promoting Justice, Therapy and Reform Through Direct Brain Interventions. *Criminal Law and Philosophy* 8 (1):21-42.
- . 2013. Enhancing Responsibility. In *Neuroscience and Legal Responsibility*, edited by N. Vincent. NY: Oxford University Press.
- Walum, H., I. D. Waldman, and L. J. Young. 2016. Statistical and Methodological Considerations for the Interpretation of Intranasal Oxytocin Studies. *Biol Psychiatry* 79 (3):251-7.
- Wolf, S. 2000. Sanity and the Metaphysics of Responsibility. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by C. Mackenzie and N. Stoljar. NY: Oxford University Press.