



Revisiting Moral Bioenhancement and Autonomy

J. Y. Lee

Received: 9 December 2020 / Accepted: 16 May 2021
© The Author(s), under exclusive licence to Springer Nature
B.V. 2021

Abstract Some have claimed that moral bioenhancement undermines freedom and authenticity – thereby making moral bioenhancement problematic or undesirable – whereas others have said that moral bioenhancement does not undermine freedom and authenticity – thereby salvaging its ethical permissibility. These debates are characterized by a couple of features. First, a positive relationship is assumed to hold between these agency-related concepts and the ethical permissibility of moral bioenhancement. Second, these debates are centered around individualistic conceptions of agency, like free choice and authenticity, which hail from an atomistic tradition of autonomy. My view is that emphasizing individualistic conceptions of autonomy do not provide particularly strong foundations on which to argue about the issue of the permissibility of moral bioenhancement. This is because individualistic autonomy is not the kind of agency-related consideration we *ought* to value. Instead, I propose that we investigate the relationship between moral bioenhancement and a more *relational* kind of autonomy. Focusing on this latter relationship, on my view, clarifies the potential for moral bioenhancement to support or enhance people’s autonomy.

Keywords Enhancement · Moral Enhancement Moral Enhancement · Moral Bioenhancement · Neuroethics · Relational Autonomy · Autonomy

Introduction / Published online: 22 June 2021

Morally educating agents, or socializing them to internalize certain moral principles with the aim to morally improve them, is neither novel nor particularly controversial as a practice. Call this *traditional moral enhancement* (TME). Moral enhancement is of interest to us because, as Thomas Douglas has pointed out, “according to every plausible moral theory, people often have bad or suboptimally good motives” [1]. In the past decade or so, *moral bioenhancement* (MBE) has been identified and distinguished as a potential type of moral enhancement that utilizes *biomedical* interventions “employed for directly controlling some aspect of human neurocognitive functioning that is viewed as instrumental to moral thought and/or behaviour” [2]. Such methods thereby go *beyond* traditional methods in their use of biomedical technologies. The term enhancement, furthermore, also implies an *improvement* of morality in those who possess moral faculties [3], rather than *treating* the absence of morality in those who do not possess such faculties (e.g. sociopaths).

Much of the discussion on MBE to date has revolved around Ingmar Persson and Julian Savulescu’s well-known work, ‘Unfit for the Future: The Need for Moral Enhancement.’ In this book, Persson and Savulescu argue that rapid technological progress has outpaced the evolution of human moral psychology, the latter of which is suited to “small and close-knit societies” [4]. The novel and potentially destructive technological tools we have on hand leaves all of us in a position whereby “it is enough if very few people are... deranged enough...to use this power for all of us to run a significantly greater risk of death and suffering” [4].

J. Y. Lee (✉)
University of Copenhagen, Copenhagen, Denmark
e-mail: ji.young.lee@sund.ku.dk

Offsetting such drastic risk from technological progress would require drastic intervention to match. Persson and Savulescu suggest that MBE would be the kind of intervention up to the task. They believe that the “motivational internalization of moral doctrines...could be sped up by means which the scientific exploration of the genetic and neurobiological bases of our behaviour might put into our hands” [4]. The biomedical interventions that might facilitate individual moral enhancement include “psychopharmaceuticals, deep brain stimulation (DBS), and genetic selection and engineering” [5].

We should note here, however, that there is a lack of consensus on the finer details regarding what *constitutes* MBE, what exactly it *targets*, how it should be *implemented*, whether it is *safe*, and whether it would really *work*. While defining MBE is itself a disputed exercise [6], some have argued that we might take a ‘minimalist approach’ and treat MBE as any intervention that targets other-regarding capacities that are “to some degree ‘prosocial’ in the broad sense of the word” [7]. To this end specific interventions aimed at targeting certain traits have been suggested, such as the use of oxytocin, serotonin or SSRI’s, β -blockers, or psychedelic drugs [7]. For instance, SSRI’s are believed to boost people’s motivation to cooperate, whereas oxytocin is believed to enhance people’s generosity [8]. Others have argued that MBE interventions should target second-order capacities such as deliberative reasoning capacities, or ‘procedural’ qualities such as “logical competence, conceptual understanding, empirical competence” [9], rather than ‘first-order’ emotions such as empathy. And while Persson and Savulescu have recommended mandatory universal MBE, others like R.B. Gibson have suggested that we might see benefits from MBE even without universal implementation, for example if “enough of the population took the enhancement to reach the behavioural equivalent of the herd immunity threshold” [10]. Many scholars, however, wonder whether such interventions would be uniquely effective, and whether resources are not better invested in traditional methods [11]. One might say, for instance, that traditional methods which permit people to retain their responsiveness to “moral reasons” [12] would suffice to counteract the qualities we find morally wanting in human beings.

The uncertainties raised in the previous paragraph provide a rich basis for continued philosophical

debate. In this paper, however, I will focus primarily on the issue of MBE and the question of its relationship to autonomy in particular. My hope is that my perspective, which aims to establish a positive link between MBE and autonomy, will also be able to provide ancillary direction for some of the uncertainties aforementioned. Now plenty of debate regarding the relationship between MBE and autonomy has already taken place. Scholars like John Harris, for instance, have argued that MBE deprives people of the “freedom to fall” [13], by which he means that doing good is no longer a *choice*. There is no virtue, he says, “in doing what you must” [13]. Persson and Savulescu have refuted those claims, demonstrating that fears about freedom of choice or moral deliberation being undermined is “overblown” [14], and that regardless, adherence to moral behaviour justifies such threats to freedom because the latter is “...not the only value, it has to be balanced against...the welfare for all” [15]. Furthermore, Lavazza and Reichlin have said that *procedural* moral enhancement might be seen to mitigate the issue of freedom, given that it “aims to improve decision-making rather than specific views, motivations, or moral provisions” [16]. David DeGrazia has argued that concerns about freedom are parallel for both moral bioenhancement and traditional moral enhancement. Contrary to the idea that MBE is particularly encroaching on freedom relative to TME, he says that acting freely is compatible with acting under some degree of influence in either case: “traditional means often do exert psychological and therefore causal influence, yet unless truly excessive, this influence is compatible with the agent’s acting freely” [17]. If we welcome such influences in the case of traditional moral enhancement, why not also for moral bioenhancement?

Despite such diverging viewpoints about the relationship between MBE and autonomy, this debate is characterized by a couple of common underlying characteristics. First – with the exception of those who believe that the value of autonomy can be *trumped* by the value of moral behaviour – a positive relationship is assumed to hold between autonomy and the ethical permissibility of MBE. Second, the kinds of autonomy¹ discussed in these debates, while

¹ I use ‘autonomy’ in a very broad sense here – autonomy might otherwise be referred to as the relevant freedom and agency concepts on which the permissibility of MBE hinges.

various, are surprisingly *individualistic*. These individualistic conceptions of autonomy subsume conceptions of individual free choice and authenticity that have long been problematized by, for example, feminist theorists in the philosophy of autonomy.

My tack for this paper is thus to encourage the MBE literature to refocus their attention away from overtly individualistic conceptions of autonomy, and towards a more *relational* kind of autonomy which duly highlights the importance of the social conditions at play in the making of autonomy. This is because, as I will demonstrate in the first part of my paper, the kinds of autonomy that are currently given attention in the MBE literature should not be taken to be the most *valuable* kinds. In the second part of my paper, I suggest that we explore the relationship between an alternative, relational conception of autonomy and MBE. I believe this exploration will have novel implications for the MBE and autonomy debate, as it will help us make sense of the more *positive* possibilities for MBE to enhance or support autonomy. I conclude my paper by reflecting that my view would help highlight the idea that MBE is not something that we might implement at the *cost* or *sacrifice* of autonomy, but rather as a positive *supplement* to it.

Mapping the Debate On Moral Bioenhancement and Autonomy

In this first section, I will critically survey what I believe are common characteristics that comprise the bulk of the debate in the literature on moral bioenhancement and autonomy. My view is that many of the types of autonomy discussed in this literature should not be touted as the most *valuable* types. This section will motivate why, in the second section of my paper, I wish to turn my attention toward a *relational* account of autonomy as the item of focus instead.

A Positive Relationship Between Autonomy and the Ethical Permissibility of Moral Bioenhancement

As pointed out in the introduction, scholars like John Harris believe that there is an inverse relationship between MBE and autonomy which thereby explains why MBE is undesirable. The underlying assumption here is that MBE *ought* to be compatible with autonomy if MBE is to be permissible. As

I will demonstrate in "[Individualistic Conceptions of Autonomy](#)", many scholars are preoccupied with *denying* that MBE is incompatible with autonomy, in order to defend the ethical viability of MBE. With the exception of those who might argue that partial losses of autonomy may be justified "in exchange for moral neuroenhancement" [16], it is evident in these discussions, too, that scholars take for granted the idea that there is a positive relationship between certain *kinds* of autonomy and the ethical permissibility of MBE. Furthermore, proponents of MBE generally do not go so far as to say that undermining autonomy *tout court* would in general be a good thing – only that it is a value that might be pushed aside in favour of *other* values. It is clear, then, that a positive relationship between autonomy and the ethical *permissibility* of MBE is presumed.

That the ethical permissibility of MBE hinges on considerations or questions relevant to autonomy seems quite plausible. While the positive relationship between autonomy and the ethical permissibility of MBE is indeed correct, something that is less certain is whether the *types* of autonomy frequently discussed in the existing literature is all that valuable to promote. If the sort of autonomy we are referring to is *not* all that valuable, then the positive relationship between autonomy and the ethical permissibility would be weakened, or at the very least overstated. While the debate does well to be shaped by the *link* between autonomy and MBE, it mistakenly assumes that it is the link to *individualistic* conceptions of autonomy that are important. Instead, as I will claim later in "[Moral Bioenhancement and Autonomy Revisited](#)", it is *relational* autonomy that is important.

Individualistic Conceptions of Autonomy

Attempts to characterize the nature of autonomy in the philosophical tradition have tended to "specifications of two broad categories of conditions: competency and authenticity" [18]. Competency regards the requisite capacities an agent must have to make autonomous decisions, like deliberative and rational faculties. Authenticity conditions go a step further by specifying and proposing to distinguish *which* exercises of such capacities can be attributed to the agent's true self. Following this tradition, the MBE literature largely values two major kinds of autonomy:

autonomy as a kind of *free choice* and autonomy as *authenticity*.

Free choice can be considered a kind of minimal requirement for individual agency, involving not only individual competency but the capacity to make *voluntary* and *uncoerced* decisions from a sufficient range of possibilities. We might think free choice plausibly necessitates *procedural independence* in decision-making, which according to Gerald Dworkin is about “distinguishing those influences such as hypnotic suggestion, manipulation, coercive persuasion, subliminal influence, and so forth, and doing so in a non-ad hoc fashion” [19]. *Authenticity*, meanwhile, would refer to choices that are taken to issue truly from one’s own will, and which reflect the *person* one is. An example of a well-known view that endeavours to capture this desideratum is Harry Frankfurt’s hierarchical account of freedom of the will. On this account, desires that count as expressive of the agent’s own will are ones that are consonant with second-order *volitions* – the latter of which constitute the *want to want* the desire in question to “be his will” [20].

For simplicity’s sake, let us call this pair of concepts – free choice and authenticity – *individualistic* conceptions of autonomy, given their emphasis on *individual* powers and *individual* identity. Free choice and authenticity may sometimes come apart. That I succumb to a temptation to eat cake, for instance, might be a free choice on my part – if it was a voluntary action which nobody else forced me to do. Yet, if I do not *want* to want the desire to eat a cake, and if I do not identify with the self that succumbs to such temptations, I would arguably not have made the most *authentic* choice here – à la Frankfurt – despite it being a *free* one.

There isn’t anything necessarily objectionable about the claim that free choice and authenticity may be *generally* valuable to people, on account of their liberty and identity-conferring qualities. But it is surprising that such individualistic conceptions of autonomy are treated as the gold standard against the value of which MBE is measured. These conceptions seem to me to rather *obscure* the potentially positive relationship between MBE and a more relational kind of autonomy. I will say more about what this relational conception of autonomy might consist of in “[Moral Bioenhancement and Autonomy Revisited](#)”, but for this section I will pick out what I take to be limiting about the individualistic emphases made of free choice and authenticity in the MBE literature.

Free Choice

Free choice, as we speak of it here, requires a kind of *procedural independence* which in the autonomy literature is standardly said to exclude coercive interference and other forms of undue subjection to forces outside one’s control. Further, as Gary Watson pointed out, a person is “free to the extent that he is able to do or get what he wants” [21], and restricted in their freedom when they are limited in the range of things they are able to do. For someone like John Harris, free choice is a desideratum of autonomy which MBE happens to be inconsistent with – and therefore MBE is said to be objectionable. He says that “autonomy surely requires not only the possibility of falling but the freedom to choose to fall” [13]. It is implied here that *free* choice must be the basis for our moral (and immoral) choices. The kind of freedom that is taken to *matter*, then, is an agent’s individual choice – what the agent can positively do, so to speak – thereby reinforcing the idea that autonomy is about the agent’s personal powers to *do* something, such as to be immoral, if they so wished.

The problem with losing this freedom of choice via MBE, according to Michael Hauskeller, is that we would essentially become subject to being controlled by an outside will. This, on Hauskeller’s view, would take away the value of *being* moral, since “instead of being ends in ourselves, we thereby become means to the end of morality...” [22]. Scholars like Robert Sparrow have claimed, also, that proponents of MBE like Persson and Savulescu “[neglect] the political dimensions of freedom” [23]. Whereas traditional moral enhancement methods like education “acknowledges a fundamental moral equality between educator and educated” [23], thanks to its dialogical, reasons-responsive mode of interaction, biomedical interventions would instrumentally “reshape the agency of others” [23] in a way that reflects “profound inequality” by circumventing that possibility [23]. This would, too, constitute a violation of autonomy as free choice, as our agency would be rendered a mere means to morality. Massimo Reichlin’s work also reflects this critique – his claim is that MBE interventions which alter “subjects’ behaviour in ways that do not involve any conscious reflection or evaluation of evidence by the subject” [24] undermines moral agency, which “requires acting for reasons.” These various violations of free choice take up much of the

controversy around MBE, and are taken to explain the ethical objectionability of MBE from the perspective of autonomy.

But is this necessarily the case? Persson and Savulescu have responded that a lack of freedom to do grossly immoral things, for instance, “does not remove the freedom to fall less deeply, by committing less wrongful acts, still less the freedom to act morally” [14, 15]. They emphasize that conceiving of MBE as “amplifying the motivational power of certain moral reasons doesn’t bypass the agent’s deliberation and decision on the basis of these reasons” [25]. Further, Diéguez and Véliz have argued that “...alternative possibilities of action may be restored if the adduced loss is compensated with an improvement in sensitivity and lucidity that can lead to seeing new options and nuances...” [26]. In general, proponents of MBE who wish to challenge the claim that the freedom to fall is necessary for autonomy may point out that “a proper concept of free will does not have to include the freedom to act otherwise, also known as ‘the principle of alternative possibilities’” [26]. Though I am sympathetic to this defense of MBE in the context of autonomy, it is telling that even these responses cater to the individualistic perspective in endeavouring to show that MBE is consistent with individual deliberative capacities, and so forth. Thus, I would add here that we can go a step further by framing autonomy in a less *individualistic* fashion. This would highlight why free choice is not as valuable as its advocates might claim it to be in the MBE context.

What we think of as choosing freely is, in the first place, determined by the world outside the agent, the world that tolerates various expressions of individuals’ agency. Take Nancy Hirschmann’s point about contextualizing meanings of freedom, for instance. She has said that determining the meaning of freedom depends on “determining the context in which claims of unfreedom are made, such that my evaluation of freedom will depend on my evaluations of other things” [27]. For example, valuing privacy might lend coherence to claims for “husbands’ “freedom” to discipline their wives”, whilst valuing women’s bodily security might lead to a counterclaim that states “governmental interference in the family is justified to protect women’s “freedom” from bodily harm” [27]. We have no real reason, then, to think that agents ought by default to be limitlessly bestowed with the freedom

to enact choices, and to think they are obstructed as and when they come up against outside interferences from the world. It is, after all, the social sphere which in the first place lays out the *value*-imbued realm of possibility within which individual agents must navigate their choices. It is within this social sphere that freedoms taken to be *normatively acceptable* may be negotiated. Our freedom *not* to be grossly harmed by others, for instance, is negotiated in accordance with what our moral community is willing to tolerate, this imposes on us corresponding duties not to violate others’ freedom to be treated the same way. This qualified version of freedom adequately takes into account the agents’ embeddedness in a social sphere and their potential reciprocal duties as a self among other selves. Yet it is no less ethically important than something like free choice, the latter of which only captures a myopic view about what it might mean to be autonomous. When we move beyond the individualistic perspective, we get closer to a more realistic and *meaningful* notion of autonomy as a contractual and relational enterprise. Thus, the narrativization of the *loss* of free choice in the realm of immoral choices may well be transformed into a narrative about freedom that is *gained* as a result of intersubjective cooperation over the boundaries of such individual choices.

Authenticity

The other side of the loss of autonomy that might be seen to take place as a result of MBE is the loss of one’s *authenticity*. Authenticity is not just about having the *option* to choose to fall, but about exercising the option to do what one *identifies with most deeply* – or something of the sort. The idea that authenticity is valuable, and allied with autonomy, is a common view in philosophical literature. Charles Taylor has pointed out, for instance, that the ‘ethic of authenticity’ partly builds on “earlier forms of individualism, such as the individualism of disengaged rationality, pioneered by Descartes...” [28]. And, as Bublitz and Merkel have noted, “most contemporary theories of personal autonomy are at least implicitly based on the idea of authenticity” [29] which is about “ensuring that actions issue from an agent’s own character” [29]. Kraemer has claimed, for instance, that for an emotion to be authentic the agents must “recognize their own feelings really as their own and identify

with them” [30]. And, as mentioned in an earlier paragraph, the Frankfurt-style hierarchical account of desires would posit that the ‘self’ which matters is revealed via an agent’s higher-level volitions. Alternatively, John Christman’s view of authenticity requires that a person “not feel alienated from the aspect of oneself in question upon reflection *given* the conditions under which that factor came about” [31].

Discussion around authenticity is especially pronounced in literature on neuroenhancement. This is because many pharmacological interventions traditionally used to treat psychiatric conditions are seen to have “off-label uses” [32], like *enhancing* mood and memory, that some believe generate “states of mind that would in some sense be alien to people taking them” [32]. Take, for instance, antidepressants like Prozac, which may be considered inauthentic if taken with a view “...not to [treat] some psychological condition but simply to [become] more confident or extroverted” [33]. In this literature, critiques are especially focused on claims to the effect “that pro-attitudes transformed by direct brain interventions such as neuroenhancements derive from mechanisms that are not the agent’s own, hence, the resulting actions are nonautonomous” [29]. The worry is that the “process of critical reflection on our desires, beliefs, intentions, and emotions” [34] which allow us to *identify* with our authentic selves might be alienated via biomedical intervention. Regardless, in the cognitive enhancement literature there are those who believe that there is no inherent tension with technopharmacological interventions and autonomy in this sense of authenticity. One view might be that “psychopharmacological substances can help users to become who they really are” [35]. As Carl Elliott has pointed out, “antidepressants, stimulants, psychedelic drugs...all have been described as a way of getting in touch with the true self” [36]. More strongly, scholars like Schaefer et al. believe that cognitive enhancement “will generally also enhance people’s autonomy” [37] because cognition is taken to have positive effects for agents’ reasoning capacities [37], the latter of which they view as key to autonomy.

Whichever account of authenticity we take to be most plausible here, the *value* of authenticity conditions appear to be as overstated as that of *free choice* in debates about enhancement. Though Erik Parens has observed that critics of enhancement technologies often worry that such technologies threaten authenticity

by “[separating] us from what is most our own” [38], authors rarely elaborate on *why* this would be a bad thing, especially in the context of enhancement, except to gesture at the idea that it would make one non-autonomous. Perhaps one might defend the value of authenticity by adding that a coherent identity may be inherently valuable, or that *inauthenticity* might be a form of self-betrayal, or “[imply] disregard for facts (wanting to be someone one is not)” [39]. Though these constitute fair reasons to vouch for a value of authenticity in general, I believe there are stronger reasons not to treat it with normative primacy. Scholars have already indicated that authenticity tends to be unhelpfully “premised on essentialist conceptions of the self,” with technology “as either preventing or enabling discovery of this self” [40]. This focus on the preservation of the individual’s sense of authentic ‘self’ highlights the not-so-flattering *individualistic* slant on the kind of autonomy that is considered valuable in enhancement literature.

To see what upholding the value of authenticity might cost us, let us draw from a different literature for a moment, which reflects concerns I take to be parallel to the issue of authenticity and MBE. A discussion I take to be particularly illuminating is Elizabeth Ashford’s objection to Bernard Williams’ well-known ‘integrity’ critique of utilitarianism. Williams’ view is that a theory of morality so demanding (as he claims utilitarianism to be) as to order agents to abandon their ‘ground projects’ – the projects by which individuals deeply identify and take to be constitutive of what their life is all about – is unacceptable. According to Ashford, however, the agent’s supposed *integrity* – their “current unified self-conception” (which seems analogous to the concept of authenticity) – must not be incompatible with overriding moral obligations. She said that slave owners, for instance, were “able to retain their self-conception as morally decent despite owning slaves, because they were inculcated with the view that slavery was morally justifiable...” [41]. Yet any plausible view of morality ought to hold that such individuals should have questioned “their way of life and the norms to which they subscribed, even at the cost of alienation...” [41]. In short, the aim to preserve an ‘authentic’ self must not be treated as prior to the project of endeavouring to be morally upstanding people, in case the two fail to align.

If we were to rather view the value of autonomy in more relational terms, we might see that threats

to *individual* authenticity allegedly posed by MBE are not particularly worrying overall. A *relational* view would give us the resources to claim that for people's autonomy in general to be advanced, it is not so much authenticity that is crucial but a sort of cooperative arrangement over people's agency. That is, if one's "authentic" self happens to be one that is expressed at the cost of the safety and health of others (for instance), we could argue that authenticity is not valuable, given that it is a condition that undermines *other* people's autonomy. As Pei-Hua Huang claims, it is unclear "why we should praise and try to preserve the inner self" when "our inner self may be that of a psychopath who loves to kill or torture" [42]. Yet this way of framing the issue of autonomy is noticeably absent in the MBE literature.

Hopefully, this section has shown that while free choice and authenticity may be part of someone's autonomous exercise, as an isolated pair of agency-concepts they present only a myopic characterization of human autonomy. The next section, then, will explicate what I have been calling a relational conception of autonomy, and show that we ought rather to value this latter conception as a compelling desideratum in the MBE debate.

Moral Bioenhancement and Autonomy Revisited

In the previous section, I critically surveyed the typical characteristics of moral bioenhancement and autonomy debates, and problematized individualistic framings of autonomy throughout. In this section, I will suggest a novel way to shape the debate about MBE and the question of its permissibility in relation to autonomy. Although we've seen in the previous section that plenty of proponents have defended the *compatibility* of MBE with individualistic conceptions of autonomy like free choice and authenticity, I believe we can make a positive argument about MBE by examining its congruity with *relational* autonomy. Though I will not provide an exhaustive account of relational autonomy itself, I will demonstrate its plausibility as a frame of value-setting when it comes to MBE, and suggest that it has potential to affect other aspects of the MBE debate as well.

So what, precisely, is meant by a *relational* account of autonomy, and how does a relational approach adequately resist the more flawed characteristics of

individualistic accounts? Feminist scholars in the past few decades have circulated 'relational autonomy' as a term that refers to a "loosely related collection of views" [43] which attempt to reconceptualize autonomy in ways that feature the role and importance of social relationships – thereby resisting individualistic notions of autonomy. Jennifer Nedelsky has criticized, for instance, the "liberal vision of human beings as self-made and self-making men" [44], of which this traditional concept of *autonomy* is a core part. As Lorraine Code pointed out, autonomy has traditionally overemphasized the value of independence and self-sufficiency, and these "variations on (Kantian) self-transparency and self-determination frame the picture of human selves that operates regulatively within the dominant social-political imaginary of liberal democratic societies" [45]. And, as Marilyn Friedman has noted, many defenders of the relational approach "are avid critics of individualistic interpretations of autonomy" [46]. A key feature that these relational theories share is "the conviction that persons are socially embedded and that agents identities are formed within the context of social relationships and shaped by a complex of intersecting social determinants, such as race, class, gender, and ethnicity" [47].

But what should we want autonomy to *mean* if we were to agree after all that the value of individualistic autonomy is overstated or overrepresented in the MBE literature? Some have claimed that we ought to be skeptical about the purported aims of relational autonomy, because of the ambiguity it generates for "delineation between preferences grounded in productive socialization and those grounded in oppression" [48]. Being socially embedded in structures of oppression may appear to present a challenge when it comes to the relational call to reject the individualism of the more traditional notions of autonomy. Marilyn Friedman pointed out, for instance, that a relational view of autonomy may "[increase] the risk of disruption in social relationships" [49]. It is important to note here, however, that in framing autonomy as a relational phenomenon, many relational theorists' have in mind a goal to "acknowledge that autonomous abilities can be undermined by severely oppressive social forces..." [50]. As Linda Barclay clarified, while "certain forms of social determinism militate against the development of autonomy competency, it is equally true that...our capacity for autonomy is

attributable to our developing and remaining embedded within a framework of social relationships” [51].

With this in mind, I will now discuss why relational autonomy is a valuable way of framing autonomy. The first thing to say here is that relational autonomy – if we are to think of it as a valuable phenomenon worth preserving or promoting in the context of MBE – must specify that negotiation of agency must take place among agents whose social relations to one another are *normatively acceptable*. The aim would be to create a space for individual agents to be *agents* – to act, do, choose, express, and so on – on the supposition that agents *coexist* with other agents similarly endowed with agential powers, rather than on the supposition that agents should maximize only their own agential powers independently of their social embeddedness. The crucial question is this: how are we to be agents in a social realm?

Relational autonomy as a theoretical conception of autonomy tempers the assumption that people must have primarily free choice and authenticity (in the individualistic senses) to be autonomous. The relational conception would be compatible with the suggestion that people’s manifestation of their free and authentic choices must be *adjudicated* by something like normatively agreeable community standards, in order to, for instance, protect and uphold a space for the agency of persons who are minoritized or marginalized by a dominant oppressive group. The point is, then, *not* to attempt just to maximize individual free choice or authenticity *wherever* it may manifest. Rather, we are to emphasize the possibility that agents’ variegated social circumstances can determine how easy or difficult it is to exercise such phenomena. Scholars like Marina Oshana have said that “autonomy is a matter of having a stable social status of a particular type” [52]. This sort of idea, I believe, gets us closer to the kind of autonomy it would be valuable to try to achieve in a society overall. The default assumption about the constitutive conditions of autonomy here notably refers to the outside world, or *external* considerations, rather than features only about the individual agent and what they do and identify with. Oshana has claimed, also, that autonomy requires that “persons [possess] influence and authority of a form and to an extent sufficient for a person to oversee undertaking in those domains that are of import to her agency” [52]. Sufficient authority in this context, crucially, should only be possible in tandem

with adequate arrangement and distribution of certain kinds of social relationships, goods and provisions. In view of my point regarding normatively desirable relationships, these relationships must not reinforce ones that are intuitively normatively *undesirable* (e.g. oppressive relations). This means that being oppressed by others, disrespected, discriminated against, deprived of certain goods, silenced, threatened, subject to harm, made to feel unsafe, and so forth, would undermine relational autonomy, to the extent that they fail to reflect relations which capture normatively acceptable ties. Thus, inflicting such acts would contribute to the disturbance of normatively desirable human relations which aim to address the question of agency in social space.

Next, we might think of relational autonomy as a term that covers the social *resources* people have to develop and exercise their agential powers. Intuitively it seems that assisting, supporting, and validating others has a role to play in facilitating people’s agency. A friend helping a friend; a state providing education; a parent raising a child – these are all everyday cases of social relationships as resources conducive to the development and exercise of individual agents’ autonomy, provided that such relations do not also perpetuate oppression. The notion of relational autonomy encourages us to think very broadly about the kinds of scaffolding relations between people that *help*. This framing, then, demands that we substantively evaluate and question whether such scaffolding relations might be *improved* or made better – and how they might be made so.

Thus, the *success* of autonomy as a social phenomenon is subject not only to human relationships in general, but also to the *kinds* of relationships people hold with one another in a society, and the ways that such relations are overseen, encouraged, reinforced, and regulated. In my view, *this* is the kind of autonomy we ought to care about as a value worth promoting: a concept that examines the ways that people stand socially in relation to one another, how people cooperate and uphold certain duties and obligations to one another, and how such acting together might shape the kinds of individual allowances and powers people have about what they can *do*. As it so happens, then, intersubjective *moral* agreements about what each agent might do plays a rather crucial role in the standard-setting and regulation of human relationships that aim to produce relational autonomy.

How does this discussion square with the issue about the permissibility of MBE as hinged on autonomy considerations? Hopefully, I have shown in this section that if we conceptualize autonomy in the more relational way that I've described above, then interventions that foster communal and cooperative relationships would be an essential part of *achieving*, rather than undermining, autonomy. And the question of whether agents can act like agents in a *social* space rather than as though isolated from the community, will be answered by something like *moral* motivation and behaviour. If MBE can be utilized as a way to ensure a greater *degree* of relational autonomy, then the former is plausibly a key resource and vehicle for the latter; and it would also be surprisingly congruous with contemporary feminist attempts to reconcile traditions of autonomy with socially sympathetic considerations. For instance, if MBE interventions could make people on balance *less* likely to assault or injure others, to threaten them, oppress them, dominate them, discriminate against them, and so on, and *more* likely to act in pro-social, just, and non-oppressive ways, we could “reconcile the value of autonomy with social obligations to mitigate vulnerability” [53]. The agential restrictions experienced by those who find themselves in precarious social circumstances – whether it be because they are persecuted by others for their identity, because they are unable to escape an unequal relationship, and so forth – could be alleviated with this restoration of relational autonomy.

This opens up novel opportunity for us to talk about the permissibility of MBE, given its potential consonance with a type of autonomy that plays out in accordance with normatively attractive community standards. Claims to the effect that MBE is permissible can be made *more* robust by establishing a positive relationship between MBE and a type of autonomy that reflects this degree of normative value, as compared with the individualistic types that have thus far been more frequently debated in the literature. By making the concept of autonomy in operation more valuable and normatively attractive than before, we would no longer need to make autonomy *compete* with the value of morality.

Furthermore, this relational turn would enable us to clarify and inform some ancilliary concerns about MBE brought up earlier, regarding questions such as establishing *what* ought to be targeted, and what *kinds* of moral behaviour should be promoted. A relational

approach gives us a natural solution to the question of what to target and why, in the context of MBE. MBE would not be about forcing individual agents to “do good” in some abstract sense. It would also not involve some objective list of virtuous traits, attempts to make people better ‘motivated’ to be moral, nor even be about expanding individual rational and deliberative faculties (though these elements may also help). Rather than laying out a separate rationale or scheme of ‘morality’, the aim would be to develop interventions that can be used to practically boost *relational* autonomy, the normative value of which we have already defended.

Thus, programmes to advance MBE contextualized in tandem with the value of relational autonomy substantiates for us the specific moral targets for MBE: as hinted at in a previous paragraph, interventions which lower individual tendencies towards domination, boost adherence to cooperative and altruistic behaviour, improve conformity to standards of justice, and so on, are all justified objectives. These objectives, which naturally arise out of our relational desideratum to promote a qualified and communal version of free agency, are very much in line with Persson and Savulescu’s suggestion that “altruistic concern and a concern for justice” [54] would enhance our adherence to morality. The advantage of my view is that we would not even need to go so far as to defend the moral urgency of issues like climate change and the existence of weapons of mass destruction to make MBE sound appealing against the value of autonomy. MBE would, instead, be developed as a way to *make good* on everyone’s relational autonomy by aiming to promote normatively acceptable social relations that will preserve that autonomy.

One might have reservations, however, about whether the notion of relational autonomy as laid out above really is instructive enough to enable a clear scheme of moral targeting in MBE development. While adopting MBE to upkeep relational autonomy – say, by keeping incidences of domestic violence low via behavioural targeting – does not seem particularly objectionable, we might have different intuitions about targeting that is *extremely* encroaching on free choice and authenticity. For instance, it would appear inappropriate to attempt to dull traits like *passion* or *risk-taking* in general just in case intervening with such traits in this way, on balance, made everybody safer or better off. This could be an example of a loss

that people are *not* generally normatively prepared to sacrifice for the sake of greater communal freedom. The *forms* of MBE that should be endorsed under the relational approach to autonomy, then, remains an open question to some extent. But here I would reiterate that threshold-setting for MBE should not be such a great obstacle from a relational perspective. It does not *preclude* values like free choice and authenticity, but merely recasts such concepts with a view to consider agents' social environments; so, determining the kinds of interventions appropriate to maintain relational autonomy would involve ongoing appraisals on the various ways we might balance these considerations for coexisting agents.

Conclusion

In this paper, I surveyed the ways that discussion about moral bioenhancement and autonomy have generally been conducted in the philosophical literature. I claimed that these debates were right to identify that the permissibility of MBE may plausibly hinge on the issue of autonomy. Yet the permissibility, or impermissibility of MBE, would be overstated if we relied only on individualistic conceptions of autonomy that extol, primarily, *free choice* and *authenticity*. The individualistic view, I said, does not provide a normatively attractive basis for a view of autonomy that ought to *matter* in context. I argued, instead, that we might see a more positive relationship between MBE and autonomy if we formulated the latter in a more *relational* manner.

My way of framing autonomy moved us away from well-worn discussions regarding MBE and individualistic conceptions of autonomy. An emphasis on relational autonomy instead helped generate a novel exploration of the consonance between MBE and autonomy. This framing then allowed us to draw out the relatively underexplored, *positive* position that MBE may be importantly *conducive* to autonomy.

References

1. Douglas, Thomas. 2008. Moral Enhancement. *Journal of Applied Philosophy* 25 (3): 228–245.
2. Shook, John R., and James J. Giordano. 2017. Moral Bioenhancement for Social Welfare: Are Civic Institutions Ready? *Frontiers in Sociology* 2 (21): 1–5.
3. Shook, J.R. 2016. My Brain Made Me Moral: Moral Performance Enhancement for Realists. *Neuroethics* 9: 199–211.
4. Persson, Ingmar, and Savulescu, Julian. 2012. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press.
5. Specker, J., F. Focquaert, K. Raus, et al. 2014. The ethical desirability of moral bioenhancement: a review of reasons. *BMC Medical Ethics* 15: 67.
6. Specker, J., M.H.N. Schermer, and P.B. Reiner. 2017. Public Attitudes Towards Moral Enhancement: Evidence that Means Matter Morally. *Neuroethics* 10: 405–417.
7. Ahlskog, R. 2017. Moral Enhancement Should Target Self-Interest and Cognitive Capacity. *Neuroethics* 10: 363–373.
8. Huang, P., and P. Tsu. 2018. Biomedical Moral Enhancement in the Face of Moral Particularism. *Royal Institute of Philosophy Supplement* 83: 189–208.
9. Schaefer, G.O., and J. Savulescu. 2019. Procedural Moral Enhancement. *Neuroethics* 12: 73–84.
10. Gibson, R.B. 2020. The epidemiology of moral bioenhancement. *Medicine, Health Care, and Philosophy* 1–10.
11. Zarpentine, C. 2013. The Thorny and Arduous Path of Moral Progress: Moral Psychology and Moral Enhancement. *Neuroethics* 6: 141–153.
12. Focquaert, F., and M. Schermer. 2015. Moral Enhancement: Do Means Matter Morally? *Neuroethics* 8: 139–151.
13. Harris, John. 2011. Moral Enhancement and Freedom. *Bioethics* 25 (2): 102–111.
14. Persson, I., and J. Savulescu. 2016a. Enharrismen: A Reply to John Harris about Moral Enhancement. *Neuroethics* 9: 275–277.
15. Persson, I., and J. Savulescu. 2016b. Moral Bioenhancement, Freedom and Reason. *Neuroethics* 9: 263–268.
16. Lavazza, A., and M. Reichlin. 2019. Introduction: Moral Enhancement. *Topoi* 38: 1–5.
17. DeGrazia, David. 2014. Moral enhancement, freedom, and what we (should) value in moral behaviour. *Journal of Medical Ethics* 40: 361–368.
18. Lewis, Jonathan. 2020. Autonomy and the limits of cognitive enhancement. *Bioethics* 1–8.
19. Dworkin, Gerald. 2015. The nature of autonomy *Nordic Journal of Studies in Educational Policy* 2.
20. Frankfurt, Harry. 1971. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68: 5–20.
21. Watson, Gary. 2004. "Free Agency" In *Agency and Answerability: Selected Essays*. Oxford: Oxford University Press.
22. Hauskeller, Michael. 2017. Is It Desirable to Be Able to Do the Undesirable? Moral Bioenhancement and the Little Alex Problem. *Cambridge Quarterly of Healthcare Ethics* 26 (3): 365–376.
23. Sparrow, R. 2014. Better living through chemistry? A reply to Savulescu and Persson on 'moral enhancement.' *Journal of Applied Philosophy* 31: 23–32.
24. Reichlin, Massimo. 2019. The moral agency argument against moral bioenhancement. *Topoi* 38: 53–62.
25. Persson, Ingmar, and Julian Savulescu. 2015. The Art of Misunderstanding Moral Bioenhancement Two Cases. *Cambridge Quarterly of Healthcare Ethics* 24 (1): 48–57.

26. Diéguez, Antonio, and Véliz, Carissa. 2019. "Would Moral Enhancement Limit Freedom?" *Topoi* 38(1): 29–36.
27. Hirschmann, Nancy J. 1996. Toward a Feminist Theory of Freedom. *Political Theory* 24 (1): 46–67.
28. Taylor, Charles. 1991. *The Ethics of Authenticity*. Cambridge: Harvard University Press.
29. Bublitz, J.C., and Merkel, Reinhard. 2009. Autonomy and Authenticity of Enhanced Personality Traits *Bioethics* 23(6): 360–374.
30. Kraemer, Felicitas. 2011. Authenticity Anyone? The Enhancement of Emotions via Neuro-Psychopharmacology. *Neuroethics* 4: 51–64.
31. Christman, John. 2009. *The Politics of Persons: Individual Autonomy and Socio-historical Selves*. Cambridge: Cambridge University Press.
32. Glannon, Walter. 2008. Psychopharmacological Enhancement. *Neuroethics* 1: 45–54.
33. Erler, Alexandre. 2012. One man's authenticity is another man's betrayal: A reply to Levy. *Journal of Applied Philosophy* 29 (3): 257–265.
34. Glannon, Walter. 2015. *Neuroethics: Cognitive Enhancement*. Oxford University Press: Oxford Handbooks Online.
35. Bolt, L.L. 2007. True to oneself? Broad and narrow ideas on authenticity in the enhancement debate. *Theoretical Medicine and Bioethics* 28 (4): 285–300.
36. Elliott, Carl. 2011. Enhancement technologies and the modern self. *Journal of Medicine and Philosophy* 36: 364–374.
37. Schaefer, G.O., Kahane, Guy, and Savulescu, Julian. 2014. Autonomy and Enhancement. *Neuroethics* 7: 123–136.
38. Parens, Erik. 2005. Authenticity and Ambivalence: Toward Understanding the Enhancement Debate. *The Hastings Center Report* 35(3): 34–41.
39. Lavazza, Andrea. 2019. Moral Bioenhancement Through Memory-editing: A Risk for Identity and Authenticity? *Topoi* 38: 15–27.
40. Walker, Mary Jean, and Mackenzie, Catriona. 2020. Neurotechnologies, Relational Autonomy, and Authenticity. *IJFAB: International Journal of Feminist Approaches to Bioethics* 13(1): 98–119.
41. Ashford, Elizabeth. 2000. Utilitarianism, Integrity, and Partiality. *The Journal of Philosophy* 97 (8): 421–439.
42. Huang, Pei-Hua. 2015. Authenticity, Autonomy, and Enhancement. *Dilemata* 7 (19): 39–52.
43. Westlund, Andrea. 2009. Rethinking Relational Autonomy. *Hypatia* 24 (4): 26–49.
44. Nedelsky, Jennifer. 1989. Reconceiving Autonomy: Sources, Thoughts, and Possibilities. *Yale Journal of Law & Feminism* 1: 7–36.
45. Code, Lorraine. 2000. Perversions of Autonomy and The Subjectation of Women. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. By Catriona Mackenzie and Natalie Stoljar. Oxford: Oxford University Press.
46. Friedman, Marilyn. 2014. Relational Autonomy and Independence. In *Autonomy, Oppression, and Gender*, ed. Andrea Veltman and Mark Piper. Oxford: Oxford University Press.
47. Mackenzie, Catriona, and Stoljar, Natalie. 2000. Introduction: Autonomy refigured. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. By Catriona Mackenzie and Natalie Stoljar. Oxford: Oxford University Press.
48. Wenner, Danielle M. 2020. Nondomination and the Limits of Relational Autonomy. *IJFAB: International Journal of Feminist Approaches to Bioethics* 13(2): 28–48.
49. Friedman, Marilyn. 2000. Autonomy, social disruption, and women. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. Catriona Mackenzie and Natalie Stoljar. Oxford: Oxford University Press.
50. Veltman, Andrea, and Piper, Mark. 2014. Introduction In *Autonomy, Oppression, and Gender*, ed. Andrea Veltman and Mark Piper. Oxford: Oxford University Press.
51. Barclay, Linda. 2000. Autonomy and the social self. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, ed. Catriona Mackenzie and Natalie Stoljar. Oxford: Oxford University Press.
52. Oshana, Marina A.L. 2015. Is Social-Relational Autonomy a Plausible Ideal? In *Personal Autonomy and Social Oppression: Philosophical Perspectives*, ed. Marina A.L. Oshana. New York: Routledge.
53. Mackenzie, Catriona. 2019. Feminist innovation in philosophy: Relational autonomy and social justice. *Women's Studies International Forum* 72: 144–151.
54. Person, Ingmar, and Julian Savulescu. 2019. The duty to be morally enhanced. *Topoi* 38: 7–14.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.