

# How (not) to Argue For Moral Enhancement: Reflections on a Decade of Debate

Norbert Paulo<sup>1</sup> · Jan Christoph Bublitz<sup>2</sup>

Published online: 24 May 2017

© The Author(s) 2017. This article is an open access publication

**Abstract** The controversy over moral bioenhancement has fallen into a stalemate between advocates and critics. We wish to overcome this stalemate by addressing some of the key challenges any moral enhancement project has to meet. In particular, we shall argue that current proposals are unpersuasive as they, first, fail to diagnose the often complex causes of contemporary moral maladies and, second, are premised on methodological individualism. Focusing on brains and minds neglects social and environmental factors. Solving the mega-problems of today very likely requires more than transforming individual brains, it requires structural and higher-level changes. By itself, moral bioenhancement is thus insufficient for solving these problems. Moreover, we outline some of the yet unresolved problems of (democratic) legitimacy a mandatory state-run bioenhancement project faces and show why they cannot be defused through analogies with moral education. Finally, normatively less worrisome means of enhancing morality, such as psychotherapies affording self-exploration, are already available. Moral bioenhancement may thus not even be necessary for solving today's mega problems. The overemphasis on speculative future technological fixes may cloud the view on possibilities to tackle pressing problems instantly.

**Keywords** Moral enhancement · Bioenhancement · Climate change · Methodological individualism · Moral education · Anti-behaviorist fallacy

“Advances in the cognitive sciences have made neurotechnology imperative... The cognitive sciences are now the critical sciences; they will determine the answer to the ultimate moral question of human survival. Given the urgency of this immediate survival problem, the cognitive and social sciences must enable us to control [our] animalistic, barbaric, and primitive propensities ... and subordinate these negatives to the uniquely human moral and ethical characteristics of love, kindness, and empathy. ... We can no longer afford to rely solely on the traditional, prescientific attempts to contain human cruelty and destructiveness. The techniques and appeals of religion, moral philosophy, law, and education ... are, in themselves, no longer appropriate [in] the present survival urgency. ... The work on the effects of direct stimulation of certain areas of the brain, ... the effects of certain drugs on exciting, tranquilizing, or depressing the emotional and motivational levels of the individual; ... suggest that we might be on the threshold of that type of scientific biochemical intervention which could stabilize and make dominant [our] moral and ethical propensities ... and subordinate, if not eliminate, ... negative and primitive behavioral tendencies. ....

It is possible to object to the era of neurotechnology on ‘moral’ grounds and to assert that these suggestions are repugnant because they are manipulative and will take away [the] natural right to make errors—even those errors which perpetrate cruelties and destruction upon other human beings. In the light of the realities of and possible consequences of nuclear weaponry, these allegedly moral arguments seem mockingly, pathetically immoral. It

---

✉ Norbert Paulo  
norbert.paulo@sbg.ac.at

Jan Christoph Bublitz  
christoph.bublitz@uni-hamburg.de

<sup>1</sup> Department of Social Sciences and Economics, Philosophy of Law and Social Philosophy, University of Salzburg, Churfuerstr. 1, 5020 Salzburg, Austria

<sup>2</sup> Faculty of Law, University of Hamburg, Rothenbaumchaussee 33, 20148 Hamburg, Germany

would seem that [we] could afford to indulge in this type of abstract, prescientific moralizing in the past when [the] most destructive weapons were clubs, bow and arrows, or even gunpowder. To continue this type of thinking in an age when nuclear weapons are capable of destroying millions of human beings in a single irrational man-made event would seem to be a form of self-defeating and immoral rigidity.”

Observers of the moral enhancement debate will find this line of reasoning all too familiar. Surprisingly, however, these passages are not part of a contemporary defense of moral bioenhancement but excerpts from the 1971 Presidential Address to the American Psychological Association, delivered by Kenneth Clark (1971, 1054–1056). To give it a more modern tone, we only replaced “biological” and “behavioral” science with “cognitive” science; and “psychotechnology” with “neurotechnology”. Clark was an acclaimed researcher, famous for his work on the psychological effects of segregation, and contributed to the landmark decision *Brown versus Board of Education*, 347 U.S. 483 (1954) (cf. Clark 1974). His Presidential Address went down in history as the call for “peace pills” and, apparently, did a great disservice to his scientific reputation (Baker 1971). Written in the same year as *B. F. Skinner* published his *Beyond Freedom and Dignity* and promised an end to repressive state structures through behavioral modification (Skinner 1971), Clark’s words express the optimistic hope for science-based social and moral progress of the early 1970s. At the end of that decade, public outcry, human rights concerns and political and judicial investigations put an end to, among other things, the behavior modification programs in public institutions in the US (Mills 1998; Rutherford 2006). The Belmont report was published and the field of bioethics emerged to critically reflect upon the technological alterations of humans.

Clark’s suggestions strike us as almost evidently true and patently absurd at the same time. It is this tension that also characterizes the contemporary moral bioenhancement debate, which might well be read as a continuation of past debates on “social engineering” or a “psychocivilized society” (Delgado 1969) in light of more advanced technology. It is a bit surprising, then, that the present debate proceeds largely undisturbed by reference to historical precedents. To be sure, this does not make any argument wrong. But historical context renders criticism of and opposition to all too enthusiastic proposals more understandable. Just as in the turbulent ’70s, the current debate tends to yield opposing camps, pro and contra bioenhancement, without much conciliatory ground. After a decade, the debate seems to have fallen into a stalemate: it is caught up in side-debates, such as whether it is easier to kill or save a large number of people in one act, or in issues metaethics discusses with greater complexity, such as the role of emotions in moral

judgment (rationalism vs. sentimentalism), or in mutual accusations of “misunderstandings” and not taking others’ criticism seriously.<sup>1</sup> We wish to contribute to overcoming this stalemate because we think that a morally better world is, indeed, urgently needed. However, developing a feasible moral enhancement project requires recognizing and accommodating critical objections to biomedical transformations of the *conditio humana*. After some introductory remarks on the current state of the debate and how it is often misleadingly framed (Sects. 1 and 2), we shall rehearse some of the main problems any persuasive moral enhancement program has to overcome. Our central claim is that moral *bioenhancement* is likely not sufficient (Sect. 3) and perhaps not even necessary to tackle pressing global problems. Furthermore, in light of normative considerations, bioenhancements are more problematic than traditional means of moral enhancement, such as moral education (Sect. 4). This urges to reconsider systematic use of the latter rather than deferring action to the advent of speculative technologies. Nonetheless, we will argue that already-available bioenhancements may be conducive to improving moral behavior and sketch possible fields of application (Sect. 5).

## 1 Current State of the Debate

Several key aspects drive the opposition to the moral bioenhancement project as put forward by *Savulescu* and *Persson* and, more moderately, by *Thomas Douglas* (2008). The first are the hardly containable dangers of misuses of technologies altering political opinions and personal traits of citizens. They become salient in a historical perspective that inevitably conjures up images of lobotomy, mind-control, brainwashing and coercive persuasion. What could go wrong if science developed and states employed novel forms of power over citizens’ minds? The potential horror of expanding the struggle for minds, traditionally the domain of ideology, to biological foundations should be obvious.

A second source of concern is the techno-utopian dimension of moral bioenhancement. It may come in a bioconservative version, rejecting chemical interventions into human nature (which we leave aside here as we find it unpersuasive), but may also come in the form of the more general thought that cures for a problem should be

<sup>1</sup> Paradigmatic examples of the back-and-forth about who misinterpreted whom are the exchanges between Ingmar Persson and Julian Savulescu and John Harris, which continue after dozens of articles and two books (Harris 2016; Persson and Savulescu 2016), and between Persson and Savulescu and Michael Hauskeller (Persson and Savulescu 2015; Hauskeller 2016).

different in kind from its causes. After all, technological mastery over nature has brought the world to the brink of the ecological collapse; and the same is true for dangers from weapons of mass destruction. Is more of the same a promising solution? Don't we, rather, need a paradigm shift away from technological fixes and the underlying belief in technological potency? At least resting one's hopes in yet-to-be-invented technologies tends to distract from (better use of) the means already available.

Thirdly, the reductionist framing of global problems as deficits located in the minds and brains of individual actors warrants skepticism. Many writers in favor of moral bioenhancement do not spend much ink on a comprehensive analysis of the causes of contemporary moral maladies but suggest that they could, possibly, be overcome by bioenhancement. Their view on social and environmental factors appears clouded by the individualist and reductionist temper of the times. But without a firmer understanding of the range of causes for current moral problems, solutions are, at best, premature. This is particularly worrisome if they involve quite high costs, in terms of significant interferences with individual liberties and obstacles to liberal democratic orders.

These concerns come into play to different degrees, depending on the level of argument. The further we move from abstract thought experiments of ideal pills in cases devoid of social context towards real-life applications, the more these concerns come to bear. The debate could improve in precision if authors made the level of their argument more transparent, because the lamented "misunderstandings" are partly due to different argumentative levels rather than substantive disagreements. This contributes to the elusive nature of the current debate. Proponents of moral bioenhancement roughly invoke this minimal set of claims: We ought to improve moral behavior; biomedical means can contribute to this end; therefore, *pro tanto* and unless other means are preferable, we ought to develop and use such means. This is a valid argument. But concerning the soundness of the premises, proponents of moral bioenhancement often resort to mere stipulations or far-fetched possibilities. And while it does indeed seem hard to reject them out of hand, there might be good arguments against them. Furthermore, many critics do not question the premises in-principle, but are rather cautious concerns, built on aspects such as those touched on above. Unfortunately, because different levels, in-principle objections and cautious concerns are not kept apart, productive dialogue often does not set in.

The case for the most comprehensive moral bioenhancement project, forwarded by *Persson* and *Savulescu*, is based on the following premises:

1. To avert ultimate harm to humankind, we urgently have to address, among other things, the problems of climate change, global injustice, and dangers from weapons of mass destructions.
2. These problems are the result of morally defective behavior.
3. Morally defective behavior is caused by psychological deficits in individual minds, e.g., temporal and spatial parochialism.
4. Solving these problems requires remedying these individual-psychological defects.
5. Traditional means of enhancing moral behavior have proven incapable of redressing these problems, possibly because our genetic or biological make-up has emerged in conditions very unlike those of today.
6. In the not too distant future, biomedical means might be developed that afford the altering of those biological and psychological deficits.
7. There is no morally relevant difference between biomedical and traditional means of improving moral behavior (parity principle).
8. Therefore, we urgently need to develop and use biomedical means to improve moral behavior.

In what follows, we shall critically investigate these premises. Our objections are not in-principle; i.e., they do not necessarily rule out any form of moral bioenhancement. But any reasonable moral enhancement project has to take these objections into account. At the end of the article, we will sketch a number of enhancement options that are far less objectionable than the ones currently discussed.

## 2 Framing of the Debate

Although we concur with the first premise, we wish to note some objections to the framing of the moral bioenhancement debate along above lines. It prejudices the strength of premise 1, as well as the costs one is willing to incur in pursuit of such a project. *Persson* and *Savulescu* open their book, *Unfit for the Future*, with these lines: "The extraordinary progress of scientific technology has increased our powers of action so that we can cause ultimate harm, i.e. render worthwhile life forever impossible, by weapons of mass destruction or by deleterious climate change and environmental destruction" (*Persson* and *Savulescu* 2014, 1). Likewise, other proponents of moral bioenhancement allude to some of humankind's biggest problems. *Molly Crockett*, for instance, begins her talks with the Israeli-Palestinian conflict. *David DeGrazia* reminds readers of genocides in Rwanda and Bosnia, slavery, forced prostitution, mass starvation and that there are over a billion people without access to water (*DeGrazia* 2014, 362). *Douglas*

invokes problems such as world poverty, climate change and war (Douglas 2008, 230) or alludes to the “Nazi’s Final Solution, Mao’s Cultural Revolution, and recent genocide in Rwanda and the Balkans” (Douglas 2013, 160). Well, how could any reasonable person object to biomedical means in order to prevent atrocities of that scale? But is this really the issue? Grand promises to redress problems of this magnitude seem suggestive; they significantly frame the debate and may cloud critical judgment or be outright misleading. If anyone had any means of ending or preventing conflicts of this scale, we would not be writing these words. But no one has, purports to have, or is remotely likely to have, in the foreseeable future (if ever), any neurobiological means of preventing humans from engaging in atrocities such as genocide or slavery.

A range of studies have shown that moral decisions and behavior, as well as related properties such as trust, can be manipulated at a biological level, through chemical and electric interventions (Zak et al. 2005; Zak 2012; Crockett et al. 2008). As proof-of-principle, this is noteworthy and may provide some challenges for strong rationalist accounts of moral decision-making. Nonetheless, effect sizes are low and mental domains or properties cannot be targeted very specifically; cross-effects on other domains are unclear at present. Researchers thus rightly caution that their findings “are a bit more subtle and nuanced than implied in [DeGrazia’s] article, as is often the case in neuroscientific studies of complex human behaviour” (Crockett 2014, 370). In recent writings, *Persson* and *Savulescu* have distanced themselves from too “confident proposal[s] about moral bioenhancement,” which claim “that more or less by itself, [bioenhancement] will enable us to cope with the great moral problems facing humanity”, and instead advocate a “cautious proposal,” i.e., one that “concedes the need to fit in ‘a programme of more extensive moral bioenhancement. .. with more thoroughgoing traditional moral education and with various possible reforms of laws and other social institutions’” (Persson and Savulescu 2015, 50). In light of the empirical work on which they are grounded, the ‘confident’ claims appear, to put it mildly, overblown. The debate is, thus, best understood at the level of thought experiments. But then, proponents of bioenhancement should abstain from conflating ‘cautious’ suggestions with ‘confident’ allusions. Conversely, critiques fight straw men by constantly reminding that no effective bioenhancers exist (Wiseman 2016, 46 f.). This might be true but it does not affect the most visionary and interesting aspects of the moral bioenhancement proposals.

Thought experiments surely have their place in philosophy. Likewise, mirroring the present through science fiction can be illuminating, especially when pondering the society-wide impact of technologies, which is speculative by nature. We do not wish to join the chorus of those

dismissive of moral enhancement, given the current lack of effective means. However, and this is our methodological caveat: thought experiments have to keep their distance from concrete realities or historical precedents. If understood as offers of concrete solutions, they may become misleading. For instance, we suspect that every high school teacher who teaches kids empathy and self-control has stronger pro-social influences than neuro-tools. Just imagine the absurdity of a teacher opening a school board meeting on new pedagogical techniques by referring to the Rwandan genocide. But it is more than absurdity. Appeals to the Rwandan genocide or the Holocaust evoke strong emotional reactions. As paradigmatic instances of dramatic moral failures, they spur unconditional acceptance of whichever means there might be of preventing them. Proponents invoke these examples not coincidentally but because of this reaction. They thereby set the wrong tone from the start: discursive honesty demands that we do not cast critics of bioenhancement in the light of being opposed to means of averting genocide. These examples are misused, especially if arguments, in the end, only pertain to more realistic applications such as rehabilitating offenders or preventing tax evasion.

### 3 The Insufficiency of Moral Bioenhancement

#### 3.1 From Cures to Causes, and Back Again

Things are different when humankind’s great problems are not only used as an introduction for arguments to much smaller ends but when they are the very subject of inquiry, as in *Persson* and *Savulescu*’s work concerning existential threats. There, the problem is rather a lack of thorough analysis of the causes of the problem that leads to an overemphasis of specific means. This pertains to premises 5 (traditional means are insufficient) and 8 (biomedical means should be developed) in the argument above. Although currently available means have evidently failed to cure all major moral defects, it is unclear whether they are inept in-principle, and whether neurobiological means are necessary or sufficient for that end. The quest for solutions that work touches upon a general problem of the moral enhancement debate. It is helpful to distinguish between causal explanations of the genesis of a state of affairs, on the one hand, and possible means of changing it on the other. It is possible to find the latter without understanding the former. For instance, we may find out that aspirin alleviates a headache without understanding its causes. Conversely, preconceptions about causes may blind one to possible cures. Despite the Marxist adage that in interpreting the world in various ways, philosophers have missed the chance to change it, we suspect that finding proper cures

for a social-political problem on the scale of global warming very likely fails without a comprehensive diagnosis of its causes (Marxists, of course, have a profound explanation that predetermines the necessary remedies). Strikingly, such a diagnosis is largely missing in the current debate, especially with respect to the relation between individual moral defects (at a psychological or neurobiological level) and higher-structure levels such as nation states, global institutions, culture or the economy.

Current proposals are premised on the idea that global problems can be cured at an individual level (premises 3 and 4). *Persson* and *Savulescu*, for instance, presume so because they locate the ultimate causes of the problems in the psychology of individuals. However, we wish to remind that this is, at present, a debatable stipulation. The high-level *explanandum*—the collective inability to tackle the future effects of the interrelated actions of several billion people—is explained at one of the lowest levels, namely the level of individual psychology and neurobiology. There are certainly many levels in between. Assuming that higher-level processes are fully explainable at lower levels expresses a specific position on perennial controversies in the philosophy of science, the reducibility of sociology to psychology and of psychology to neurobiology.<sup>2</sup> We can, surely, only scratch at its surface. To illustrate the reductionist main problem, consider this analogy: Imagine physicists declared that because everything, including brains, CO<sub>2</sub> emissions and lethal weapons, is made from atoms, we could solve the mega-problems of today by re-arranging the atoms on the planet in a different order (*atomic moral enhancement*). Even though we cannot find in-principle objections (apart from feasibility), the absurdity of the proposal illustrates the necessity of observing adequate levels of explanation.

Typically cited candidates at an individual psychological level (premise 3) are cognitive biases, lack of self-control or emotional capacities such as empathy, or defects in more general traits, such as a sense of justice. Indeed, biases such as temporal and spatial parochialism seem to correlate with the structure of global problems such as inaction on averting climate change. Moreover, the world would likely be a better place if people were less self-interested and had more cognitive empathy. However, a persuasive moral enhancement proposal would have to show more precisely that (and which) mental processes or characteristics could be transformed to effect substantive changes on a global scale. However, we wish to voice a more fundamental skepticism about the individualistic premises (3 and 4). Is global inaction

on climate change really due to individual moral deficits? Are people not taking the threat seriously because they are biased?

Let us look at some data: A recent global survey (Stokes et al. 2015) in 40 nations found that 54% of the population consider climate change a *very* serious problem (85% say it is at least a somewhat serious problem); and 78% support their countries in signing new international treaties to cut down on emissions. Interestingly, a majority, especially in so-called developing countries, consider the negative effects of climate change to already be setting in, so it is not just a future problem. Also, people in the two top polluting countries, the USA and China, are markedly less concerned than people elsewhere or than the global median (45% and 18% respectively). What does this show? First, temporal and spatial parochialism do indeed seem to play a role: those who feel more immediately affected are also more concerned and more supportive of policies targeting climate change. However, the parochialism biases are insufficient to explain the interesting findings of the survey. For one, there is not a strong correlation between being concerned and supporting political action; even those who neither feel affected nor think it is a *very* serious problem support political action on climate change. People seem to take global problems seriously even if they are not directly affected. Further, there is an 18 point divide between support for climate action in the US (69%) and support in Western Europe (87%). How to explain that? Cognitive and emotional traits of US-Americans and Europeans are, we suppose, almost identical, and both regions are top polluters. If we want to explain the 18 point difference, we will very likely not succeed by comparing the biochemical properties of their brains. Changes in public opinion seem to have a whole array of reasons, not only ones rooted in individual psychology (see Page and Shapiro 1992, Chap. 8). But most importantly, the survey clearly shows political support from a large majority of the population to cut down on emissions through signing new international treaties. This contrasts, strikingly, with the repeated failures in treaty negotiations. How can we understand this? Evidently, the causes of failure do not primarily lie at the level of individual beliefs on climate change. The survey did not inquire into the willingness of people to endure setbacks to their lifestyle (though 67% believe that reducing the effects of climate change requires major lifestyle changes). Citizens' support for climate actions might just be cheap talk. However, it seems much more likely that something else along the way from citizens' minds to the conference tables in Copenhagen or Paris has gone astray. It might be that political representation is not working; it might have to do with powerful stakeholders, national interests, the protection of specific industries and global power structures. Then, psychological deficits seem to be neither the right

<sup>2</sup> For an introduction, see Rosenberg (2012, Chap. 9); for a recent in-depth discussion, see the papers in Zahle and Collin (2014).



level at which to understand the causes, nor the level at which to find proper solutions.

### 3.2 Individualism and the Anti-Behaviorist Fallacy

This survey illustrates our larger methodological point, skepticism about the individualistic approach. Premises 3 and 4 are based on methodological individualism, the view that all higher level social processes can ultimately be exhaustively explained (and, hence, remedied) at the level of the individual. It presupposes the reducibility of societal, cultural, historical or economic matters to the behavior of individual persons (and their psychology or even neurobiology). *Persson* and *Savulescu* explicitly endorse this position. In a recent response to one of the authors' criticisms, that they fail to address higher-level processes which lead to the exploitation of natural resources (Bublitz 2016b), they write:

we don't believe in the existence of 'higher level' social processes or structures that are independent of the minds of a *large majority* of the citizens, such that they would persist even if these minds are altered ... We agree [that economies drive exploitation of resources], but we ask what drives the latter. Ultimately, we believe that this must be human psychology, their motivation and reasons. Suppose that most human beings, throughout their history, had had a psychology and ideals like, say, Diogenes the Cynic (who is said to have lived in a barrel); then humanity today surely wouldn't be at the brink of depleting the natural resources of the planet.<sup>3</sup>

We agree that neither climate change, nor weapons of mass destruction, nor global injustice would pose problems if all lived like Diogenes. However, for one, an ascetic way of life, withdrawn from sociality, or a "back to the woods" approach as espoused by radical ecologists, do not appeal to us as promising solutions to global problems of the twenty-first century. We need solutions for over seven billion people that live densely together, form innumerable groups and structures and have to cooperate and coordinate their actions to generate the resources from which they live. But even more importantly, *Persson* and *Savulescu's* point suggests more than it can prove: while it is true that without individual human beings, no social facts, societies or economies would exist, this does *not* imply that we can explain—or remedy—dysfunctional higher-level processes at the individual level. The relation between the individual, societal and other levels, whether reductions are possible and what they entail more precisely is a central topic in the

philosophy of science. Full reducibility is at least a highly controversial claim, which seems to be forgotten too often in the face of the reductionist *zeitgeist*.

We do not wish to commit ourselves to strong positions such as *Durkheimian* group minds but would like to suggest that some supra-individual elements may well have to be recognized in political affairs, even by individualistic approaches (List and Spiekermann 2013), namely interactions between individuals and their environment. Further, understanding the actions of collectives requires taking relations between individuals into account. These relations differ from the internal properties of the individual. Consider this illustration: The physical properties of the chair on which you are sitting are not exclusively explainable by the properties of its constitutive elements, atoms. The chair is a particular assembly of atoms and its properties diverge from other assemblies (say, of a window). Both window and chair are, we suppose, made out of atoms only but they diverge in the compositional structure of the atoms, which determines their higher-level properties. If we want to know why the window breaks if you throw your chair at it, it does not suffice to look at the levels of atoms alone. The structure of the composite also constrains the actions and properties of the atoms it is made of. Even in paradigmatically reductionist physics, we have to observe compositional or structural features.

We suspect the same is true with respect to individual persons and collectives. But whereas the intrinsic physical properties of atoms are not changed by the structure they form, it is very likely that the minds and brains of individuals are influenced by higher-level and environmental processes such as society, culture, economy, etc. In contrast to visions of the '70s, environmental factors have fallen out of favor in current debates, partially due to the move from behaviorism to cognitive sciences. The central claim of classic behaviorism is that behavior, moral and immoral, is the product of environmental influences, natural and social, on the organism. The environment selects behavior, with operant and classic conditioning as the transmitting mechanisms. To behaviorists, the subject is not an interesting level of explanation, as behavior is not much more than the output of a complex reflex machine. Consequently, changes in human behavior can be produced by changing environmental conditions only. This line of thought gave rise to the behaviorist social engineering projects (Skinner 1979). In their view, mere moral bioenhancement cannot produce lasting behavior changes as long as persons are exposed to the same environmental stimuli. After some feedback loops, old behavior will set in again.

Although behaviorism has been surpassed by cognitive science largely because of its neglect of the mind, the idea of strong external influences on minds and brains is not discredited, let alone disproven. Interpreting the findings

<sup>3</sup> Neuroethics, forthcoming.

of cognitive and neuroscience as showing that minds and brains develop independently of environmental stimuli would be to commit what we call the *anti-behaviorist fallacy*. One day, the brain sciences will likely be able to explain all physical processes in the brain but they cannot, by themselves, explain the larger causes of these processes. For behaviorist the cause is to survive under given natural and social conditions (Skinner 1981). Whereas the behaviorist explains the person from the outside in, some understandings of neuroscience seem to explain the person from the inside out. These different approaches may well end in a chicken and egg problem: does the organism or the environment come first? Again, without committing ourselves to a position, we consider it highly probable that human behavior is a product of complex interactions between the person and her environment, at many interwoven levels. Social psychology and social neuroscience attest to this. Further, we should not forget that criminology has sought to study reasons for and causes of criminal behavior for over a century. Even the most biologically oriented criminologists acknowledge the importance of environmental and social factors for deviant behavior and endorse a bio-social model (Raine 2013).

Then, however, good reasons speak against full reductionism; intrinsic properties of brains are not the sole bedrock level of explanation. Structural and environmental features have to be taken into account if we want to understand (and redress) the processes that lead to global problems. While they are shaped by individual people and human psychology, the reverse is equally true: individual behavior and minds are themselves shaped by external and environmental forces detectable at higher levels. This has ramifications: a successful moral enhancement project very likely either needs to alter environmental factors too, or has to create persons *immune* to such external influences (which is why *Persson* and *Savulescu's* example of the eremite *Diogenes* works so well). As immunity to external forces is a highly unlikely trait, it appears that moral bioenhancement is likely *insufficient* to attain the desired results without concomitant social changes.

Any closer look at factors contributing to climate change presumably reiterates this basal claim. Ending the devastation of the rainforest likely fails if only brain chemistry or mental processes of locals are targeted. It requires economic, cultural and other structural changes (which, by themselves, may suffice). Or take the often invoked example of the Holocaust: any serious historian cannot but reject outright a one-dimensional explanation. The idea of an anti-Nazi pill that would have averted the catastrophe seems equally mistaken. To be sure, if a large majority of Germans had been Stoics, Nazi ideology might not have prevailed. But there were reasons and causes for them not being Stoics, just as there were for *Diogenes*, and they

do not lie in brain chemistry. Genocide examples further seem off the mark because they are paradigmatic instances of the powerful influence of mass psychology: peer and group pressure and ideology, which have to be tackled at the supra-individual level. This seems to be the key finding of social psychological research (*Zimbardo* 2007). More abstractly: Although there is a bottom-up causal influence from individual minds to global problems, it is likely more indirect and muddled than the individualistic premises of bioenhancement proponents suggest. In addition, it is likely a two-directional relation, with top-down constraining influence as well. Then, premises 3 and 4 have to be restated more inclusively and a successful moral enhancement project has to change both individual and higher levels, preferably at the same time. Therefore, by itself, moral bioenhancement is likely *insufficient* to alter human behavior and to solve global problems.

#### 4 Legitimatory Problems of Moral Bioenhancement

Whether moral bioenhancement is *necessary* depends on the strength of premise 5, the purported impotence of other means, to which we will turn in the final section. Let us, for now, assume that a safe morality pill that alters humankind's mental structure for the better is found. Should we use it? Well, assuming that it is powerful enough to avert genocide or ultimate harm: yes. But what about less serious concerns? The use of moral bioenhancers yields several legitimatory problems that any persuasive proposal has to overcome, especially if it involves state-run or mandatory enhancements of citizens.

##### 4.1 Moral Relativity and Overlapping Consensus

Any argument in favor of moral enhancement presupposes the possibility of identifying the moral good, which evidently touches upon key controversies in metaethics and normative ethics. Is it always good to be less aggressive or to feel more empathy with others? Probably not—the moral goodness or badness of aggression or empathy is likely to depend, first, on the particular context and, second, on the moral theory one endorses. Proponents attempt to meet the challenge from moral relativity in several ways.

Rather than directly 'implanting' controversial beliefs, such as the utilitarian calculus or the categorical imperative, *Persson* and *Savulescu* seek to enhance moral motivation or sensitivity to moral reasons, so that people are more likely to feel motivated by them. However, enhancing the fittingness between moral reasons and moral motivation nonetheless presupposes knowledge of what counts as a moral reason. They try to solve this problem by invoking

rather uncontroversial elements, such as enhanced altruism, or a ‘sense of justice’, and sometimes a common-sense morality, one that is, as they say, “reflected in the moral norms of leading religions, like the Golden Rule of Christianity, which has equivalents in other world religions. We rely on this moral core when we give our children moral education, and we suggest it could be relied on for the purposes of moral bioenhancement” (Persson and Savulescu 2015, 53). *DeGrazia* employs a similar idea. The problem of determining the right values is not particular to bioenhancements; parents or teachers face them every day. One should, he writes, “[s]tick to improvements that represent *points of overlapping consensus among competing, reasonable moral perspectives*” (DeGrazia 2014, 364, his italics).

But there are problems with such views of a common-sense morality or an overlapping consensus among moral perspectives. To illustrate their dubiousness, consider some of *DeGrazia’s* suggestions for dispositions which, according to him, all reasonable perspectives would conceive as moral defects and fitting candidates for elimination: anti-social personality disorder, moral cynicism, narcissistic personality disorder, self-absorption, prejudice against the interests of those outside one’s group of identification, inability to focus on unpleasant realities, weak will, inability to find creative solutions to difficult problems involving competing interests and values, and an inability to grasp subtle details that are of undeniable moral relevance (DeGrazia 2014, 364). These and other dispositions may have detrimental effects on moral behavior. Yet, it is anything but clear that all or even many instantiations of such dispositions fall under an overlapping consensus among competing moral views. Notably, not even enhancing altruism is unanimously believed to be morally good. Ethical egoism, for instance, has a respectable tradition within moral philosophy (cf. Shaver 2015) and remains, for better or worse, one of the major views on common-sense morality, most clearly seen in economic matters. Or consider *DeGrazia’s* suggestion to regard prejudice against the interests of those outside one’s group of identification as a clear example of an undeniable moral defect. To be sure, the statement of this ‘defect’ would need to be fleshed out more clearly to determine its moral quality. But we expect a hard time convincing dyed-in-the-wool communitarians that favoring one’s own group is clearly morally wrong and rightly described as a ‘prejudice.’ So, we fear that the overlapping consensus might turn out to be a rather narrow one.

#### 4.2 State Enforcement and Neutrality

But even when we suppose, for the sake of argument, a *minima moralia* exists—would this permit states to administer bioenhancements? By confining the scope of values it enforces to those within the overlapping consensus,

states might not run afoul of their duty to neutrality. Even more, the possibility of altering moral views may make one inclined to recast the issue of neutrality. For the idea of neutrality, since the day of *Locke’s* Letter Concerning Toleration (1689/2010), has resulted from the seemingly unavoidable fact of moral pluralism. Political liberals such as *Rawls* and *Larmore* emphasize that disagreements about value, unlike scientific disagreements, remain contested—there simply is no factual matter that could be solved so that the disagreement about values would vanish (cf. Larmore 2005; Rawls 2005). Neutrality is a *response* to the challenge of forming a unified state in the face of insurmountable disagreement over substantial matters. Insofar as these disagreements can be resolved by moral interventions, the question arises whether states ought to be neutral? Why not favor a particular (e.g. majority) position and offer anyone else moral bioenhancement? Put differently, the question is whether pluralism or neutrality are first order principles, good for their own sake, or second order principles, instrumentally good to install a peaceful order among disagreeing citizens (cf. Paulo 2016).

Moreover, the real and profound problem that *DeGrazia* and others fail to address is where the legitimacy to “reduce or eliminate” those beliefs outside of the overlapping consensus arises from, particularly in citizens who do not wish to change their views. If they hold moral views outside of the realm of reasonable doctrines—to take a provocative one: racism—the state may, perhaps, discriminate against those views through anti-racist campaigns, or implement education policies that impede such opinions being formed in the first place (and, uncontroversially, ban racist actions). However, beyond that, the crucial question *DeGrazia* oversteps is this: Why is it legitimate to target the emotional dispositions of an outspoken racist? The right to hold opinions, including racist ones, without interference is among the most fundamental rights (Art. 19 Universal Declaration of Human Rights). It is usually not understood to be conditional upon the beliefs of others or upon whether it stands the test of reasonableness (Bublitz 2016a). What proponents need to provide is an argument to override this guarantee. The figure of the overlapping consensus does not seem to be the right place to look for it. In fact, to invoke it as a justification for interferences with basic rights would reverse the argumentative role it is supposed to play in political philosophy. The overlapping consensus *limits* state authority, i.e., state powers are legitimate *only* if justifiable to views within the consensus. But it does not, in turn, impose duties on citizens to only hold views within the consensus. It is one thing to reject justifying immigration policies vis-a-vis racists but a completely different matter to oblige a racist to reject racism. Policing views outside of the scope of the overlapping consensus is a politically highly dangerous proposal.



### 4.3 Moral Bioenhancement and Democratic Legitimacy

This leads us to a general worry about the compatibility of the moral bioenhancement project with democracy. Freedom of opinion, we suggest, is a political right necessary for democratic legitimacy. If the state alters moral opinions, it undermines the very legitimacy on which it is grounded (Bublitz 2016b; Paulo and Bublitz 2016). *Persson* and *Savulescu* show that a democratic political order might increase the problem of morally deficient individual minds exponentially: if psychological deficiencies befall a majority of the electorate, majoritarian rule will likely let bad individual decisions accumulate and prevail over more reasonable views. That motivates *Persson* and *Savulescu*'s intriguing solution: governments should alter the psychological traits of citizens through moral bioenhancement. A different mindset of the electorate may afford the implementation of urgently needed drastic, yet unpopular, policies. Enhanced voters may, for instance, be more supportive of measures to tackle global warming or to elect a 'green' government. Of course, as a *large majority* is needed, this entails changing the minds of citizens against their will. At the same time, however, they disavow installing an authoritarian regime: "In our view, the solution to the mega-problems of today, if there is one, lies not in a shift to an authoritarian type government, but in moral enhancement of the citizens in democracies" (*Persson* and *Savulescu* 2014, 8). Accordingly, they portray state-driven moral bioenhancement of the electorate as a democratic solution.

However, such a strategy would hardly observe the limits of democracy. Recall the famous opening passage of *Jean-Jacques Rousseau*'s *The Social Contract*, where he outlines the key task of political philosophy: to inquire into the legitimacy of government "taking men as they are and laws as they can be made to be" (*Rousseau* 1762/2002, 155). Taking people as they are is a cornerstone of democratic thought—*Persson* and *Savulescu*'s moral bioenhancement proposal, in contrast, reverses this idea. If governments, through moral bioenhancement, severely manipulate the will of individuals—and thereby, of the people as a whole—a circular relation ensues. Government is not representing the will of the people but vice versa. Thus, a government that manipulated the will of the people could not be grounded in democratic legitimacy—even if it did so for noble reasons. Proponents of large-scale, state-driven and partially mandatory moral bioenhancement should, thus, be clear that they advocate a non-democratic order (*Paulo* and *Bublitz* 2016).

### 4.4 Moral Education Versus Bioenhancement

Proponents of moral bioenhancement commonly try to defuse worries about legitimacy by drawing analogies with

traditional moral education. They argue that if traditional moral education is desirable, so is moral bioenhancement (premise 7). Biomedical means are allegedly not relevantly different from traditional forms of moral education, which are universally accepted, be it through parents or religious or public institutions such as kindergartens or schools. However, we think there are a number of important disanalogies: for one, the education of children differs normatively from educating adults. Parents—and, subsidiarily, the state—have a *duty* to support the development of children, including education in moral matters. Not educating children amounts to a failure of discharging this duty. There is, however, no relevantly similar duty to educate adults. On the contrary, the relation between adults is one of equals. Adults can reject involuntary education as paternalistic. The analogy thus presupposes what it seeks to establish: that a duty to educate exists with respect to adults.

Moreover, moral education is a very complex matter. Urging one's infant time and time again not to pull the dog's tail is very different from trying to convince a teenager to follow the Golden Rule. The infant is less likely to respond to carefully presented arguments. In the pre-conventional stages of moral development, to use Kohlberg's terminology, the child is more likely to learn through sanctioning, obedience and, later, self-interested cooperation. *Persson* and *Savulescu* seem to have those stages in mind when comparing traditional and biomedical moral enhancement: "For those of us who are content with the empirical knowledge of common sense and science, it is surely evident that when small children are taught language, religion, basic moral rules, or whatever, this education is just as effective, irresistible, and irrevocable as biomedical intervention is likely to be" (*Persson* and *Savulescu* 2015, 52). We doubt that this stipulation is true and that teaching basic moral rules is "effective, irresistible, and irrevocable." Children are individually different, respond differently to their environment and, even though parents are enormously influential and often manipulative, every act of teaching is an attempt that can fail. As *Michael Hauskeller* correctly observes, education is not simply "injecting knowledge, beliefs and behavioural dispositions into an empty container and then, once it is filled, sealing it for good so that it cannot get out again" (2016). Children make up their own minds more and more, in correspondence with their increasing developmental stages. To be sure, children can be drilled and morals imprinted. But these forms of education, at least in the later developmental stages, are themselves worrisome and pedagogical scholarship has been discussing other, more dialogical forms for decades (on the problem of children's autonomy and state neutrality, see *Callan* 2004; *Fowler* 2015).

The central difference between biomedical and traditional means is that the latter are set in the "space of

reasons”, whereas biomedical means bypass it. Children are taught to play by the rules of reason and, as soon as they are cognitively able to do so, moral education is primarily an exchange of reasons and reflection, of arguments and counterarguments. To an increasing extent during development, children are taught—and should be encouraged—to make up their own minds. The more capacities a child acquires, the more parental education turns into an attempt of persuasion rather than mindless imprinting (Sparrow 2014). At some point, usually in the post-conventional stages of moral development, children might even successfully convince their parents or teachers of a moral rule they did not endorse. So biomedical means bypass individual reason and self-control, while traditional moral education usually respects them. Furthermore, an argument by analogy from manipulative forms of education seems to miss its target. Does the fact that there are manipulative and indoctrinative forms of education, inevitably due to children’s development, provide any argument in favor of expanding and increasing these forms beyond the inevitable? We do not think so. The education analogy thus does not hold.

#### 4.5 Neutral Moral Insight?

Changing others’ opinions and beliefs through reasons and argument rather than biomedical means is also preferable for epistemic reasons. Any metaethical position has to concede that moral views can be subjected to rational critique and should be revisable in light of better arguments, which implies a certain “open-ended normativity” of moral discourse (Buchanan 2012). Effective biomedical forms that confer moral “insights” or form certain virtues, however, suggest a fixed and contained view of the moral realm. Instead of an open-ended normativity, the search for moral truth is concluded and revisions in light of new arguments or changed circumstances are likely impeded. This relates to a problem recently pointed out by Parker Crutchfield (2016). Suppose a specific moral belief in an individual is changed through biomedical means and that the person is aware of her new beliefs. Proponents of moral bioenhancement may say she has acquired a new moral insight. The problem is, however, that the person does not have new reasons for having formed that belief, as it has not been generated by evidence or argument. Furthermore, no emotional experience—such as the unexpected loss of a loved one—has led to the change. In other words, the enhanced person lacks any form of justification for her newly acquired belief. She may even realize that she has acquired a new belief that is not warranted. This poses the epistemological question whether this new belief really is a belief *strictu sensu* (or merely a different mental state, such as a vague seeming). As rational beings, we are interested in *justified* beliefs. Moreover, it might well be the case that moral

beliefs acquired without justifying reasons are less likely to motivate behavior than beliefs acquired in a traditional way.

In this section, we have reviewed some of the central problems of legitimacy that moral bioenhancement projects, which seek to impose moral values or beliefs upon citizens in non-communicative or non-rational ways, have to face: the challenge of relativity; missing legitimacy grounds; problems of democracy if states alter the will of the people. The often invoked analogy with moral education cannot defuse these worries. These problems cast premise 7—the parity principle—into serious doubt.

### 5 Prospects of Moral Enhancement

From the foregoing, we can derive the following desiderata: *First*, any persuasive moral enhancement project has to embark on the painstaking journey to find the causes of the problems it seeks to cure. Most importantly, it has to integrate higher-level processes, such as the structure of the growth-economy that drives exploitation of resources, but also political structures that may misrepresent the people primarily affected by global injustice and climate change (Paulo and Bublitz 2016). Confining one’s view to psychological traits of individuals tends to overlook multilevel relations, gets lost in translation between them or is oversimplified. *Secondly*, a moral enhancement project very likely has to be embedded in broader changes at levels other than individual psychology. The positive side is that we can start with those changes right away. There is no need to wait for speculative future technological fixes. The emphasis on future means also tends to cloud the view of already available strategies and may postpone urgently needed action. *Thirdly*, a meaningful moral enhancement project has to take concerns about legitimacy seriously. Improving morality is not *eo ipso* legitimate. Altering citizens’ opinions may interfere with strong rights to freedom of opinion, even of those outside the purview of the overlapping consensus, and raises profound problems in light of democratic ideals. Therefore, in scenarios short of ultimate harm, only *voluntary* moral bioenhancement projects are normatively feasible. *Fourthly*, the significance of the qualitatively novel way of altering morally relevant traits at the biological level through biological means, rather than in the space of reasons through evidence and argument, has to be fully appreciated. Strong reasons speak against discarding a central enlightenment idea: that we solve collective matters by the force of arguments, not by non-rational powers such as those of pharmacology.

We surely share the aims of moral enhancement: making the world a better place. And while we hold it necessary to induce changes at institutional, economic and other structural levels, the individual level is, of course, relevant

as well. Only an integrated multilevel approach has the chance to succeed. The interesting question is, therefore, how could morality at the individual level be improved?

Much current writing takes a somewhat fatalistic approach: it is said that, since the heydays of Greek philosophy, human morality has not improved much. Traditional means of moral enhancement appear impotent to induce wide-ranging, meaningful change. As this is due, so it is speculated, to genetic and biological causes, moral bioenhancement is necessary (premise 5). This is, of course, an empirical claim, not substantiated by much evidence. Authors such as Steven Pinker paint a more optimistic picture of humankind's moral trajectory (Pinker 2011). Be this as it may, we are skeptical about the premise that the available means to improve human morality have been used exhaustively to no avail.

It is illuminating to put this question in a broader perspective. The development of individual persons seems to be a long and perhaps never-ending process of balancing self-interested desires and demands of sociality. This process plays out in the external world, through institutions prohibiting and correcting flawed and incentivizing good behavior; and it plays out within the individual, through the internalization of norms. The *Freudian* model of the mind, with the conflict between the id and ego, and the super-ego as the mediating instance, nicely illustrates this (Freud 1989). The process of balancing conflicting drives and demands is the process of civilization. So, it is likely shortsighted to claim that our biological heritage makes us unfit for the demands of the future. We are and always have been in this struggle. In this perspective, it seems unfounded to suppose that current problems are of a kind that the system evolved to adapt to changing environments cannot handle. The empirical evidence is also less damning than sometimes suggested. Take spatial and temporal parochialism. People often act with a bias towards the near and discount long-term effects of actions. However, psychology does not show that people are *unable* to act differently. One of the key claims of the relevant literature is that there are two systems, one of which is prone to parochialism, activated under specific conditions. In conditions more favorable to long-term rational thought, people can overcome their intuitions and make good judgments (Kahneman 2012; Heath 2014; Bloom 2016). The lesson then, it seems, is to create favorable conditions for better choice, e.g., by institutional design enhancing rationality.

Moreover, let us assume that some psychological processes need fine-tuning. Finding out which ones have to be weakened or strengthened, where the problem of morality really resides, requires a much closer look at developmental and moral psychology. *Prima facie*, for instance, it seems that guilt and shame function by transforming external expectations into internal demands. Amplifying

both, therefore, appears to be a promising bioenhancement strategy. However, it is anything but clear whether a collective guilt trip makes this world better off, because it likely has secondary effects, such as increased self-deception and avoidance. So, sophisticated intrapsychic understanding and empirical research are needed to make reasonable proposals. Moral psychology over the last decades has produced a wealth of findings but this knowledge is neither systematically applied, nor brought to practical use. Or, to invoke a practical example, take the rehabilitation of criminal offenders. Some methods seem to work under some conditions. Thus, while traditional techniques may have to be updated, we have not even begun to conduct any large and systematic research project into enhancing moral behavior in these ways. There is no political or cultural project worth its name that exists to enhance morality in light of modern research. "Nothing works" is a premature position.

Pessimism may suggest that such endeavors are futile. However, a broader perspective suggests that there are effective ways to alter people's desires and choices. The problem is, rather, that respective research is, in large parts, mainly used for business purposes, especially marketing. For decades, an entire subfield of psychology, with its own journals and research centers, has been devoted to creating demands and making people buy things. So, we wonder, if only parts of the amount of money that goes into commercial marketing were spent on pro-morality campaigns, would we not likely see significant improvements? One well-known example of how knowledge from behavioral economic and psychology can be used for good purposes is nudging, i.e., the design of 'choice architectures' in a way that leads people more or less "automatically" (without deliberation, that is) to make decisions that are deemed good (Thaler and Sunstein 2008). A classic example would be a cafeteria where fruits are presented at eye level, thus making it more likely that consumers reach for the healthier option, because unhealthy foods such as chocolate bars are slightly harder to see and reach (but still available). Along those lines, significant moral improvements may be reached through means that have proven effective in non-moral settings. Furthermore, what if we harnessed the mass media for such purposes? How often is talk in the mass media about climate change as compared to trivialities? To us, there is a striking gap between the opportunities to influence people and the degree to which they are used for moral campaigning. In fact, it seems that, to an overwhelming extent, mass media is used for purposes which contribute to, for example, climate change (consumerism; constantly creating desires for new products). In short, there are a number of promising means readily available; what is missing is the individual and political will to put them to better uses.

Moreover, when we look for psychological causes of the problem, we find some evident candidates largely absent from current moral bioenhancement proposals. First of all: self-criticism. Everyday experience tells us that people do not have problems observing moral rules; in fact, they are quick to point out the wrongdoings of others. But what is much harder to attain is a self-critical stance towards oneself. Given the inherent perspectivity of all human perception, this does not come as a surprise. Self-criticism requires taking a detached view of oneself and is hurtful towards our fragile egos. For these reasons, there seems to be a moral blind spot—oneself (for a powerful recent exposition of this view, see Batson 2015). We indulge in moral rationalization rather than self-critical work on ourselves. But surely, we can learn constructive self-criticism, especially if our behavior is mirrored by others. A range of psychotherapeutic treatments is available to explore one's depths when distress is too high. These means could also be used to improve the self-understanding of non-disordered people. But there are no systematic ways in which people are incentivized and supported to learn to understand and deal with themselves. Not even at school do we teach kids the basics about their psyche or techniques to better understand themselves. In other words: there is a striking lack of opportunities for people to undergo self-criticism and self-growth. Psycho- or other forms of therapy are only available in very few places in the world, are often rather expensive and rarely covered by insurance or health plans. Great powers of self-transformation lie dormant. A global program of learning to become a more self-reflective person, grounded in evidence-based assessments of various talk therapies, could be rolled out within a couple of years.

One may object that people are not interested in undertaking hard work on themselves. But this might be an underestimation of the power of morality. A glimmer of hope that has not been given sufficient attention arises from the idea of a “moral self” (Prinz, forthcoming). Being a morally good person seems to be one of the central characteristics upon which persons build their (narrative) identity. If we look at moral failures, we encounter, indeed, a striking fact: It is not that people consider themselves immoral; on the contrary, the overwhelming majority thinks that they act morally and say that being moral is highly important to them. Even the Nazis managed to convince themselves that they did the morally right thing (Batson 2015). Of course, this involves a lot of self-deception and hypocrisy. But nonetheless: because being morally good is an intricate part of one's identity, people will be very willing to declare their allegiance to morally right conduct. And even if this is more for external display rather than an expression of true internalized morality, this is a promising hook

with which a systematic voluntary moral enhancement project may persuade people to participate and make intrinsic commitments.

### 5.1 Biomedical means

In such programs, biomedical means could play a valuable role. Instead of imposing attitudes or dispositions or permanently altering neurobiology, they could conduce specific processes in (experimental) therapeutic sessions. They could facilitate processes of self-understanding by overcoming ego-defenses; or they could afford perspective taking. Again, we don't have to look into the future; everyday experience shows us possible applications. Take, for instance, the most widely used (and widely illegal) substance, cannabis. Apparently, it creates feelings of destressing, being less aggressive, etc. Cannabis advocates like to call it “the drug against wars”. Why not give cannabis, integrated into therapy, to overly aggressive people, to let them experience a less aggressive view of the world? What about the “insights”, the ego-dissolving effects that generate feelings of being intimately connected with the natural world, that psychedelic substances are said to possess? At this point, we shall only remind readers about the movements of the 1960s and 1970s that explicitly combined biomedical means with the quest for individual and collective moral progress. Much would have to be said about their failures but it seems they were not caused by impotent substances; their means were not merely *moral hallucinogens*. They failed because of other (cultural and societal) factors. Those, again, could be remedied, and much more easily than other biomedical means be invented. The counter-movements of the 1970s, with all their failures and successes, seem to us as potential anchor points for developing a moral enhancement project that embeds biomedical means with cultural and societal transformations. Such a kind of self-development, through substance augmented self-exploration, largely avoids the pitfalls mentioned above.

### 5.2 Epistemic enhancement

In times of post-truth, truthiness and alternative facts, we also wish to emphasize the importance of capacities to make good moral judgments, which mainly fall into the domain of cognitive enhancement and are therefore less controversial. The idea is sometimes couched as ‘procedural’ moral enhancement (Schaefer and Savulescu 2016b, a). Obviously, there are, among others, psychological causes for the worrying trend. We expect the cognitive science of post-truth to become a hot topic soon. A moral enhancement project that would draw on such research and counteract *failures of rationality* by improving moral



deliberation would be less worrisome than other forms, at least in light of the idea of a democratic society of rational persons. One might consider it as creating and safeguarding necessary cognitive competencies for the democratic *Citizen*. Furthermore, it seems that most of the people susceptible to post-truth thinking, and that is part of the paradox, do declare their allegiance to norms of rationality. Thus, just as most people want to be moral, they want to be rational. And such self-conceptions—and self-deceptions—may lay the ground for their voluntary participation.

### 5.3 Aesthetic Perceptual Enhancement

Finally, in the imaginative style of the debate, we wish to mention a form of enhancement that escapes charges of moral relativity and epistemology concerns—aesthetic enhancements. Consider the unrealistic but theoretically interesting science-fiction short story *Liking what you see* by Ted Chiang (2015), in which he envisions a kind of moral bioenhancement that has to do with calliagnosia. As a fictional neurologist explains: “The condition is what we call an associative agnosia, rather than an apperceptive one. That means it doesn’t interfere with one’s visual perception, only with the ability to recognize what one sees. A calliagnosic perceives faces perfectly well; he or she can tell the difference between a pointed chin and a receding one, a straight nose and a crooked one, clear skin and blemished skin. He or she simply doesn’t experience any aesthetic reaction to those differences” (Chiang 2015, 285). In the story, calliagnosia can be switched on and off and has no significant side-effects. It is used to prevent “lookism”, i.e., unconscious prejudice against unattractive people, which arguably would count as moral progress. It might have modest individual, local and short-term effects or—when made compulsory, as beautifully discussed in the story—deep individual, wide societal and long-term effects.

## 6 Conclusion

However, there is no need to wait for tomorrow’s wonder-technologies. Systematic and promising moral enhancement projects could start today. A lot could be done without raising fundamental normative concerns. The work of Douglas, Persson, Savulescu and others has placed this important issue on the agenda. We concur with the aims but have objections over the proper cures. The individual psychological level can be a legitimate target, if undertaken voluntarily. Moral progress might well require, in the end, changing individuals. But that does not imply that all change has to start from there. Only a multilevel approach seems to be promising. This leads back to the opening

words by *Kenneth Clark*. He, after all, suggested morally enhancing political leaders only:

There is no way of predicting the personal and emotional stability of leaders with the life and death power over mankind. .... The masses of human beings are now required to live and continue to work on faith, hope, denial, and the acceptance of the chances that their powerful leaders will have the strength to use their power wisely and morally. Given these contemporary facts, it would seem logical that a requirement imposed on all power-controlling leaders—and those who aspire to such leadership—would be that they accept and use the earliest perfected form of psychotechnological,[neurotechnological] biochemical intervention which would assure their positive use of power and reduce or block the possibility of their using power destructively. This form of psychotechnological neurotechnological medication would be a type of internally imposed disarmament. It would assure that there would be no absurd or barbaric use of power. It would provide the masses of human beings with the security that their leaders would not or could not sacrifice them on the altars of their personal ego pathos, vulnerability—and instability (Clark 1971, 1056).

While it is intriguing to speculate how history might have unfolded if *Brezhnev*, *Castro*, *Mao Tse-tung* and *Nixon* had followed *Clark’s* proposal and negotiated over global affairs under the influence of peace pills, enhancing leaders seems almost a modest suggestion. It would evade many thorny problems, as one could make a compelling case that holding office requires, or should be made conditional upon, heightened moral competence. This could be attained by psychological or biomedical interventions.

In these days, *Clark’s* words appear as a timely commentary on recent events. To many, it seems that one political leader, perhaps the most powerful person on this planet, suffers from what *Clark* described as power-pathos. The people cannot rest assured of the emotional stability of their leaders. In the *New York Times*, a group of psychiatrists recently openly stated that the President of the United States seems to suffer from emotional instabilities that make him incapable of serving as president (Dodes and Schachter 2017). The response by *Allen Frances*, one of the drafters of the DSM, is instructive for our inquiry. Apart from questions of diagnostic criteria, he writes: “The antidote to a dystopic Trumpean dark age is political, not psychological” (Frances 2017). The controversy over the right levels of analysis and remedy has evidently become a matter of great relevance. Just as we argued that it is misleading to overemphasize the individual level, so too it would

be to fully disregard it. The only way out of our moral maladies is political *and* psychological.

**Acknowledgements** Open access funding provided by University of Graz.

### Compliance with Ethical Standards

**Conflict of interest** The authors have no conflict of interests to declare.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Baker R (1971). "Oberver." *New York Times*, September 16.
- Batson CD (2015) What's wrong with morality? a social-psychological perspective. Oxford University Press, Oxford
- Bloom P (2016) Against empathy: the case for rational compassion. Harpercollins, New York
- Bublitz JC (2016a) Moral enhancement and mental freedom. *J Appl Philos* 33(1):88–106. doi:10.1111/japp.12108
- Bublitz JC (2016b) Saving the world through sacrificing liberties? a critique of some normative arguments in unfit for the future. *Neuroethics* 1–12. doi:10.1007/s12152-016-9265-8
- Buchanan, AE (2012) The open-ended normativity of the ethical. *Analyse Kritik* 34(1):81–94
- Callan E (2004) Creating citizens: political education and liberal democracy. Oxford University Press, Oxford
- Chiang T (2015) Liking what you see: a documentary. In: *Stories of your life and others*, Picador pp 281–327
- Clark KB (1971) The pathos of power: a psychological perspective. *Am Psychol* 26(12):1047–1057. doi:10.1037/h0032217
- Clark KB (1974) *Pathos of power*. Harper & Row, New York
- Crockett MJ (2014) Moral bioenhancement: a neuroscientific perspective. *J Med Ethics* 40(6):370–371. doi:10.1136/medethics-2012-101096
- Crockett MJ, Clark L, Tabibnia G, Lieberman MD, Robbins TW (2008) Serotonin modulates behavioral reactions to unfairness. *Science* 320(5884):1739. doi:10.1126/science.1155577
- Crutchfield P (2016) The epistemology of moral bioenhancement. *Bioethics* 30(6):389–396. doi:10.1111/bioe.12239
- DeGrazia D (2014) Moral enhancement, freedom, and what we (should) value in moral behaviour. *J Med Ethics* 40(6):361–368
- Delgado JMR (1969) *Physical control of the mind: toward a psychocivilized society*. Harper & Row, New York
- Dodes L, Schachter J, (2017) A mental health warning on Trump. *New York Times*, February 14
- Douglas T (2008) Moral enhancement. *J Appl Philos* 25(3):228–245
- Douglas T (2013) Moral enhancement via direct emotion modulation: a reply to John Harris. *Bioethics* 27(3):160–168
- Fowler T (2015) In defence of state directed enhancement. *J Appl Philos* 32(1):67–81. doi:10.1111/japp.12068
- Frances A (2017) An eminent psychiatrist demurs on Trump. *New York Times*, February 15
- Freud S (1989) *Das Ich und das Es* (1923). In: *Psychologie des Unbewußten*. Studienausgabe, 7th edn. vol 3. Fischer, Frankfurt am Main
- Harris J (2016) *How to be good: the possibility of moral enhancement*. Oxford: Oxford University Press.
- Hauskeller M (2016) The art of misunderstanding critics: the case of Ingmar Persson and Julian Savulescu's defense of moral bioenhancement. *Camb Q Healthc Ethics* 25(1):153–161. doi:10.1017/S0963180115000407
- Heath J (2014) *Enlightenment 2.0. Restoring sanity to our politics, our economy, and our lives*. Harper, New York
- Kahneman D (2012) *Thinking, fast and slow*. Penguin, London
- Larmore C (2005) Respect for persons. *Hedgehog Rev* 7(2):66–76
- List C, Spiekermann K (2013) Methodological individualism and holism in political science: a reconciliation. *Am Polit Sci Rev* 107(4):629–643. doi:10.1017/S0003055413000373
- Locke J (1689/2010) A letter concerning toleration and other writings. In: Goldie M (ed). *Liberty Fund*, Indianapolis
- Mills JA (1998) *Control: a history of behavioral psychology*. NYU Press, New York
- Page BI, Shapiro RY (1992) *The rational public: fifty years of trends in Americans' policy preferences*. University of Chicago Press, Chicago
- Paulo N (2016) Liberal perspectives on moral enhancement. *Ethics Politics* 18:397–421
- Paulo N, Bublitz C (2016) Pow(d)er to the people? Voter manipulation, legitimacy, and the relevance of moral psychology for democratic theory. *Neuroethics* 1–17. doi:10.1007/s12152-016-9266-7
- Persson I, Savulescu J (2014) *Unfit for the future: the need for moral enhancement*. Oxford University Press, Oxford
- Persson I, Savulescu J (2015) The art of misunderstanding moral bioenhancement: two cases. *Camb Q Healthc Eth* 24(1):48–57. doi:10.1017/S0963180114000292
- Persson I, Savulescu J (2016) Enharrismment: a reply to John Harris about moral enhancement. *Neuroethics* 9(3):275–277. doi:10.1007/s12152-016-9274-7
- Pinker S (2011) *The better angels of our nature: the decline of violence in history and its causes*. Allen Lane, London
- Prinz J. (forthcoming) *The moral self*. Oxford University Press, Oxford; New York
- Raine A (2013) *The anatomy of violence: the biological roots of crime*. Allen Lane, London
- Rawls J (2005) *Political liberalism*. Columbia University Press, New York
- Rosenberg A (2012) *Philosophy of Social Science*, 4th edn. Westview Press, Boulder, CO
- Rousseau J-J (1962/2002) *The social contract: and, the first and second discourses*. Dunn S (ed) Yale University Press, New Haven
- Rutherford A (2006) The social control of behavior control: behavior modification, individual rights, and research ethics in America, 1971–1979. *J Hist Behav Sci* 42(3):203–220. doi:10.1002/jhbs.20169
- Schaefer O, Savulescu J, (2016a) Better minds, better morals: a procedural guide to better judgment
- Schaefer O, Savulescu J, (2016b) Procedural moral enhancement. *Neuroethics* doi:10.1007/s12152-016-9258-7.
- Shaver R. (2015). Egoism. In: Zalta EN (ed) *The stanford encyclopedia of philosophy*, Spring 2015, Metaphysics Research Lab, Stanford University, Stanford <https://plato.stanford.edu/archives/spr2015/entries/egoism/>.
- Skinner BF (1971) *Beyond freedom & dignity*. Knopf, New York
- Skinner BF (1979) The non-punitive society. Commemorative Lecture given by B.F. Skinner after receiving an Honorary Doctorate at Keio University, Japan." [http://www.bfskinner.org/wp-content/uploads/2014/02/Skinner\\_1979\\_The\\_Non-Punitive\\_Society.pdf](http://www.bfskinner.org/wp-content/uploads/2014/02/Skinner_1979_The_Non-Punitive_Society.pdf)

- Skinner BF (1981) Selection by consequences. *Science* 213(4507):501–504
- Sparrow R (2014) Egalitarianism and moral bioenhancement. *Am J Bioeth* 14(4):20–28. doi:10.1080/15265161.2014.889241
- Stokes B, Wike R, Carle J, (2015) Global concern about climate change, broad support for limiting emissions. Pew Research Center's Global Attitudes Project. <http://www.pewglobal.org/2015/11/05/global-concern-about-climate-change-broad-support-for-limiting-emissions/>
- Thaler RH, Sunstein CR (2008) *Nudge: improving decisions about health, wealth and happiness*. Yale University Press, New Haven
- Wiseman H (2016) *The myth of the moral brain: the limits of moral enhancement*. MIT Press, Cambridge, MA
- Zahle J, Collin F (2014) *Rethinking the individualism-holism debate*. Springer, Heidelberg; New York
- Zak PJ (2012) *The moral molecule: the source of love and prosperity*. Dutton, New York
- Zak PJ, Borja K, Matzner WT, Kurzban R (2005) The neuroeconomics of distrust: sex differences in behavior and physiology. *Am Econ Rev* 95(2):360–363. doi:10.1257/000282805774669709
- Zimbardo P (2007) *The lucifer effect: understanding how good people turn evil*. Random House, New York