Imagining moral enhancement practices

Jona Specker

Imagining moral enhancement practices

Jona Specker

Imagining moral enhancement practices
Specker, J.

This thesis is a result of a research project funded by the NWO program Our Brain as Capital. The Ethical Desirability of Moral Bioenhancement (project number 326-20-001)

ISBN: 978-94-6361-113-8
Layout and printed by: Optima Grafische Communicatie (www.ogc.nl)
Cover artwork: Lily de Bruijn (yosoy_lily@hotmail.com)

Imagining Moral Enhancement Practices

Nadenken over praktijken van morele mensverbetering

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 28 juni 2018 om 15.30 uur

door

Jona Specker
geboren te Amsterdam

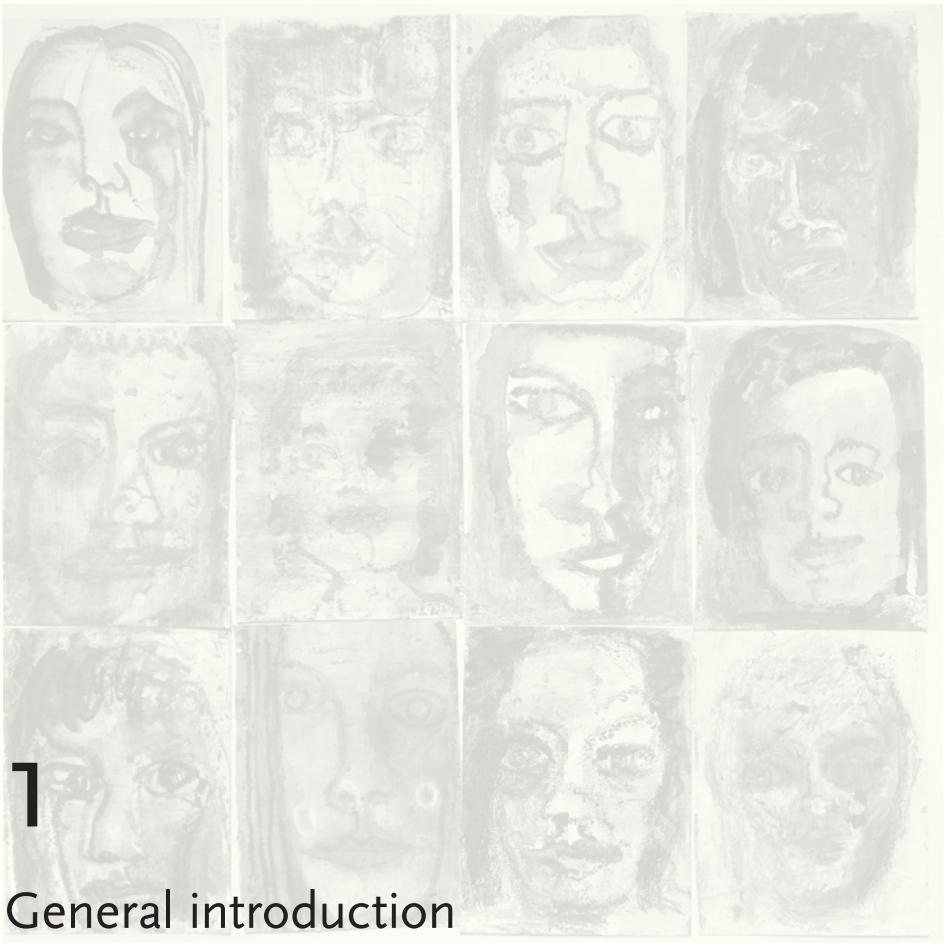
Erasmus University Rotterdam

Promotiecommissie

Promotoren
Prof.dr. M.H.N. Schermer
Prof.dr. I.D. de Beaufort

Overige leden
Prof.dr. S. van de Vathorst
Prof.dr. A.L. Bredenoord
Prof.dr. G. Meynen

Table of contents



1

General introduction

Prelude

The inimitable sci-fi story *Altruizine, or A True Account of How Bonhomius the Hermetic Hermit Tried to Bring About Universal Happiness, and What Came of It* written by the Polish writer Stanislaw Lem narrates the manifold ways in which good intended attempts to create a better society by disseminating a pioneering psychopharmacological substance can backfire (Lem, 2014 [1965]). In this story, the robot Bonhomius leaves the cave he has spent over sixty years meditating in to help his fellow creatures: "it dawned on me that to spend a life in solitude was wrong, for truly, did all my exceedingly profound thoughts and strivings of the spirit ever keep one rivet from falling, and is it not written that thy first duty is to help thy neighbour and not to tend to thine own salvation" (Lem, 2014 [1965], p. 246).

Bonhomius is accompanied by his mentor Klapaucius, a 'constructor' with exceptional engineering skills. They turn to Gnost, a hesitant supercomputer "able literally to contain the Universe itself within its innumerable memory banks" (Lem, 2014 [1965], p. 256) and ask him to help them find a way to eliminate misery and bestow happiness on everyone. Despite his considerable scepticism – numerous earlier attempts have failed miserably – the supercomputer eventually complies and prints them a recipe for 'Altruizine', a drug that triggers profound empathy[1] by replicating the emotional states of others in one's vicinity. Its workings are detailed below:

> ALTRUIZINE. A metapsychotropic transpitting agent effective for all sentient homoproteinates. The drug duplicates in others, within a radius of fifty yards, whatever sensations, emotions and mental states one may experience. Operates by telepathy, guaranteed however to respect one's privacy of thought. Has no effect on either robots or plants. The sender's feelings are amplified, the original signal being relayed back in turn by its receivers and thereby producing resonance, which is as a result directly proportional to the number of individuals situated in the vicinity. According to its discoverer, ALTRUIZINE will insure the untrammelled reign of Brotherhood, Cooperation and Compassion in any society, since the neighbours of a happy man must share his happiness, and the happier he, the happier perforce they, so it is entirely in their own interest that they wish him nothing but the best. Should he suffer any hurt, they will rush to help at once, so

1 A word about concepts here. In the multidisciplinary field of empathy research, the concept of empathy is used in a variety of ways, to refer to a range of distinct but related (emotional, motivational, and cognitive) psychological phenomena (Batson, 2009; Decety & Cowell, 2014). Arguably, the kind of empathy experienced as a result of Altruizine, resembles what in the literature is termed "emotional contagion": experiencing the feelings of those in your immediate vicinity.

> as to spare themselves the pain induced by his. Neither walls, fences, hedges, nor any other obstacle will weaken the altruizing influence. The drug is water-soluble and may be administered through reservoirs, rivers, wells and the like. Tasteless and odourless. One millimicrogram serves for one hundred thousand individuals. We assume no responsibility for results at variance with the discoverer's claims. Supplied by the Gnost. Computerized representative of the Highest Poss. Lev. Devel. (Lem, 2014 [1965], pp. 267-268)

Bonhomius sets off to a planet inhabited by humans and diffuses the Altruizine drug. Chaos ensues. Overwhelmed by the one-on-one transmission of other peoples' emotions and sensations, people start to roam around in mob-like hordes and to violently attack each other.

> By dawn the Altruizine effect was so strong, that if one nostril itched, the entire neighbourhood for a mile on every side would respond with a shattering salvo of sneezes; those suffering from chronic migraines were abandoned by their families, and doctors and nurses fled in panic when they approached – only a few pale masochists would hang around them, breathing heavily. And then there were the many doubters who slapped or kicked their compatriots, merely to ascertain whether there was any truth to this amazing transmission of feelings everyone spoke of, nor were these compatriots slow in returning the favour, and soon the entire city rang with the sounds of slaps and kicks. (Lem, 2014 [1965], pp. 271-272)

Bonhomius' and Klapaucius' honest attempt to bring about a better society by radically enhancing people's empathy, has failed.

'Introduction to the introduction'

One could easily dismiss this pessimistic and rather gloomy story about the use of a pharmacological agent to increase human empathy as rather farfetched, and an interesting thought experiment at best. Nonetheless, one of the most lively and vehemently debated issues in the bioethical literature of the last decade concerns closely related proposals. Since 2008, the so-called 'moral enhancement debate' asks whether we should actively pursue the development of moral enhancement technologies, and whether it would be permissible – or even obligatory – to put them to use, provided that these interventions would be effective and safe?

Whereas 'traditional methods' of moral betterment (such as upbringing, socialization and education) are arguably as old as humanity itself, the debate on moral bioenhancement focuses on the desirability of methods based on novel or emerging biomedical insights, or the use of biomedical methods. The debate follows a significant rise in fundamental research on the (neuro)biological and genetic underpinnings of human (and animal) morality – moral emotions, moral cognition, moral judgment, moral behaviour. Potential interventions that are being discussed in the literature on moral bioenhancement range from various types of psychopharmaceuticals, neurostimulation, and genetic selection and engineering.

Central topics of debate have been how to understand and define moral enhancement; whether (safe and effective) moral enhancement technologies should be mandated; and the question whether means matter morally, i.e. whether there are inherent ethical differences between biomedical and nonbiomedical means of moral enhancement. Just as earlier human enhancement debates (dealing with the ethical desirability of enhancing beauty, sports, mood, cognition and memory), the debate on moral enhancement is highly polarized. Proponents argue that moral enhancement is our only hope in averting disaster, while opponents fear that moral enhancement would *provoke* disaster.

However, until now, the theoretical debate on moral (bio)enhancement has a strong speculative character and mostly precedes and runs ahead of realistic scientific possibilities. Moreover, to date, the debate risks lacking focus and real world impact, as different commentators fail to agree on how to understand and define moral enhancement, and on what kind of examples would constitute cases of moral enhancement.

This under-examination of potential *practices of moral enhancement* is significant because without specifying intended users, contexts of implementation, and the goals and objectives of developing and applying potential biomedical possibilities for optimizing morally relevant capacities, it is not clear who should be concerned about this debate. At the same time, existing or emerging practices that already contain elements of 'moral enhancement' but as yet lack ethical attention and ethical scrutiny, risk staying out of sight.

In order to address this problem, this thesis focuses on (present and emerging) moral enhancement *practices*, in order to identify ethical issues that are not necessarily part of the current debate on moral enhancement. Rather than distinguishing in a blanket fashion desirable from undesirable moral enhancement *technologies*, this thesis intends to formulate conditions and ethical requirements for ethically justifiable moral enhancement practices.

Setting the scene: The new sciences of morality and the debate on moral bioenhancement

The promises of the new sciences of morality

For the last couple of decades, a growing body of research into the so-called "new sciences of morality" studies neurobiological and behavioural genetic models of human morality (Slaby, 2013; Yoder & Decety, 2017). A plethora of scientific disciplines, including social and behavioural neuroscience, biology, evolutionary psychology, experimental psychology, and primatology, study phenomena moral philosophers have examined conceptually for ages.

Primatologists study our evolutionary ancestors, mainly great apes like bonobos and chimpanzees, in order to identify potential continuities between these species and ours. Frans de Waal for example argues that we can find a kind of "proto-morality" in these apes, in the form of empathy, reciprocity, and a sense of fairness that can rightly be called he thinks "the building blocks of morality" (de Waal, 1996, 2009; de Waal & Sherblom, 2018).

In the field of evolutionary and cognitive psychology theorists look upon morality as an evolutionary adaptation and ask how something morality-like might have been advantageous, that is fitness enhancing, for our ancestors (Katz, 2000; Nichols, 2005; Kitcher, 2007). Their research gives way to discussions about whether (and if so, how much of) human morality should be considered 'innate'. Some hypothesize that humans come to this world with an innate capacity for moral judgments already in place (Hauser, 2006). Others go even further and speculate about specific moral judgments being innate, in the form of for example a universal bias against 'unauthorized killing' that forms a necessary precondition for close knit communities to live and work together.

In the early 2000's, in various 'morality labs', the first neuroscientific experiments were set up, aimed at mapping the neural circuits that are involved in moral thinking and moral judging and to find perhaps a specialized moral centre in the human brain (Damasio, 2000; Greene, 2001; Haidt, 2001; Greene & Haidt, 2002). These experiments raised questions about the role played by emotion and cognition in moral judgments, and stirred discussions about which of these two should be considered as a more decisive factor. Later, attention shifted from trying to uncover "domain-specific morality", i.e. identifying brain regions specifically dedicated to morality, to mapping more domain-general processes, such as theory of mind (Young & Dungan, 2012).

In the meantime, a rich landscape of studies on the 'emerging neurosciences of' has appeared, ranging from the emerging neuroscience of justice motivation and justice sensitivity (Decety & Yoder, 2017), empathy (Marsh, 2018), intergroup relations (Cikara & Van Bavel, 2014), social decision-making (Yoder & Decety, 2017), violence (Miczek et al., 2007), to pro-social behaviour (Yamasue et al., 2012), etc. etc. The same goes for (behavioural) genetic studies on for example violence and aggression (Baum, 2013; Asherson & Cormand, 2016), and antisocial personality disorder and psychopathy (Viding et al., 2005; Tracy et al., 2010).

In criminological research a previously improbable interest in biological and social underpinnings of antisocial behavior is visible (Raine, 2013, pp. xii-xiii; Glenn & Raine, 2014). Bio-social criminology stresses the interplay between social and biological factors: "What is important about this new line of research is that it is not a reification of the outdated nature versus nurture debate, but rather an entirely new perspective that highlights the dual influences of genetic and environmental factors in the etiology of crime and delinquency" (Beaver et al., 2015, p. 109).

Moreover, these biological, 'brain-based' explanations of (im)morality[2] have given rise to related debates on the relevance of this growing body of knowledge for a variety of domains (Blank, 2007; Levitt & Manson, 2007; Mendez, 2009; Zik & Roberts, 2015). In the field of 'Neurolaw' for example, the meaning and implications of neuroscience for the law and legal practices is investigated (Meynen, 2014, p. 819; Morse, 2017). Neuroscientific information and techniques are already being used in court (de Kogel & Westgeest, 2015). This research has also stirred debate about the question whether in the future brain imaging could be used to (aid in) accurately predicting future violence and other forms of antisocial behaviour (Poldrack, 2017; Poldrack et al., 2018), and help in tailoring treatment to individual needs.

> Although single genes only contribute to a small proportion of the overall variance in antisocial behaviour, identifying genes that confer risk may aid in the development of treatment methods that could potentially be tailored to specific risk factors of the individual. In addition, it may improve our understanding of the biological pathways that lead to antisocial behaviour. (Glenn et al., 2013, p. 4)

The promises of these new sciences of morality thus go beyond improved understanding of biological underpinnings of morality. As moral (and immoral) behaviour are

2 In *The Anatomy of Violence*, Adrian Raine argues that criminals have "broken brains, brains that are physically different from those of the rest of us" (Raine, 2013, p. 180).

explained in terms of (disordered) brain functions, commentators have also speculated about how insights from neuroscience could benefit conflict resolution and foster peace (Bruneau, 2015) or even ways in which potential neurobiological interventions could be used to correct potential "social or moral dysfunction":

> Understanding the functional anatomy of moral judgment and the anatomical and functional differences between social and antisocial brains is a prerequisite for developing socially and morally acceptable neurosurgical interventions that treat social or moral dysfunction. (de Ridder et al., 2009, p. 156)

In short, the new sciences of morality not only add to and challenge previous predominantly non-biological scientific understandings of human morality (Abend, 2013), but go a step further as well, and suggest novel methods to deal with behavioural problems, such as pharmacological means, neurostimulation, or even neurosurgical interventions as mentioned in the quote above. As Adrian Raine writes in his book *The Anatomy of Violence*: "It's a bitter pill for many criminologists and psychologists to swallow, but medications do work in controlling and regulating aggression in children and adolescents" (Raine, 2013, p. 292). Current techniques still fall short however, and a range of methodological and ethical challenges remain to be met.

There are of course historical predecessors to these biological approaches to moral and immoral behaviour. Felix Schirmann narrates in what ways the moral brain has been an object of scientific study before, and the 19th century ways of "thinking of immoral persons in terms of disordered brains altered what it means to be immoral. Though contested, the immoral persons discussed were conceived of as viable objects for brain science, psychiatry and medicine: their misdoings were observed on a behavioral level, their actions were evaluated with ethical standards and they were punished according to the law. Observation, evaluation and reprimand were part of the social purview, yet the respective experts sought the causes for their misconduct in their neurobiology." (Schirmann, 2013a, p. 43; 2013b) Notorious is the identification of 'human types', among which 'the born criminal', by the Italian physician and anthropologist Cesare Lombroso, whose ideas were influential in the first half of the twentieth century (for example in the eugenics movement). During the second half of the twentieth century, biological explanations for criminal or immoral behaviour were strongly resisted. For instance, in The Netherlands, professor in criminology Wouter Buikhuisen was deterred from doing research on biological features of criminals, next to social factors

(Berkvens, 2009; de Haan, 2009; Oosterhuis, 2014).[3] Now, however, renewed attention for genetic and biological risk factors for criminality has emerged.

Interest in the neuroscience of morality has not been confined to the academic world, but has attracted considerable interest from the general public and from policy makers as well. Popular science books abound, both about the *moral* (Hauser, 2006; Churchland, 2011) and the *immoral* brain (Fallon, 2013; Raine, 2013), and about empathy or lack thereof (Baron-Cohen, 2011). Fascination with psychopathic or criminal genes and brains is reflected in considerable attention in the media (Kahn, 2012; Dennett, 2013; Frentz et al., 2015; Levy, 2015a; Hagerty, 2017), although resistance to the advance of so-called 'brain culture' is visible as well (Dehue, 2011; Bloemink, 2013; Dehue, 2014; de Vrieze, 2017; Koelewijn, 2017).

Also here, attention goes beyond fascination with biological explanations of moral and immoral behaviour, and the question is asked whether we can put this new knowledge about the underpinnings of human moral and immoral behaviour to good use. For example, can we teach people (e.g. our children, or medical doctors) to have more empathy (Boodman, 2015; Krznaric, 2015; Eerkens, 2017), for example by using virtual reality (Polak, 2015; Berdik, 2017)? Or, one step further, by means of a simple pill? Plentiful op-eds about the desirability of so-called 'morality pills' have been published (Kahane, 2011; Singer & Sagan, 2012; Crockett, 2014b; Levy, 2015b; Dubljević, 2017; Mandelbaum, 2017; Rampton, 2017; Tremonti, 2017) and on various online forums people have been discussing the possibility and desirability of developing, propagating, and using such pills (Pellissier, 2012; Aeon, 2015).[4]

3 "by the 1970s it became sociological common sense that fatalism, determinism, reductionism, sexism – a naturalization and legitimation of existing relations of power – would follow inescapably from any engagement with the reality of human biology – as either an ontological question – what were humans really like? – or as an epistemological one – what can biology tell us about the forms of life that humans have made for themselves? Human biology was relevant only in that it provided the preconditions for language, meaning and culture, whose form and content must be accounted for in non-biological terms." (Rose, 2013, p. 10)

4 Fascination with potential ways to improve people morally, including taking a pill, is evident in the television series *The Good Place* as well (Goddard et al., 2016). In one of the first episodes, protagonist Eleanor Shellstrop needs a quick and dirty way to improve herself morally in order to extend her place in The Good Place: "It is time to make me good, partner! How do we do it? Is there a pill I could take, or something I could vape?" But instead of giving her a pill her partner Chidi, a former professor of moral philosophy, decides to teach her philosophy as a means to become a better person.

The bioethical debate on moral enhancement

In the current academic debate on moral bioenhancement it has been suggested that biomedical interventions will aid in solving a number of pressing societal issues, such as crime and violence, or will aid in solving climate change - either by enhancing pro-social tendencies and emotions or by inhibiting anti-social tendencies and emotions. The debate started in 2008, with the publication of two articles advocating moral enhancement in the *Journal of Applied Ethics* (Douglas, 2008; Ingmar Persson & Savulescu, 2008).[5]

In his first paper on moral enhancement, Thomas Douglas main intention is to debunk what he calls the Bioconservative Thesis (Douglas, 2008). It posits that "Even if it were technically possible and legally permissible for people to engage in biomedical enhancement, it would not be morally permissible for them to do so" (Douglas, 2008, pp. 228-229), because, the bioconservative argument goes, although the enhancement might be good for the individual, it might be bad for others. Douglas then presents moral enhancement as a type of enhancement that is not necessarily disadvantageous for others, and can even benefit others (and not necessarily yourself). A good target for moral enhancement according to Douglas, are the "counter-moral emotions":

> I have in mind those emotions which may interfere with all of the putative good motives (moral emotions, reasoning processes, and combinations thereof) and/or which are themselves uncontroversially *bad* motives. Attenuating such emotions would plausibly leave a person with better future motives, taken in sum. (Douglas, 2008, p. 231)

The central question Douglas poses is whether it would be ethically permissible for individuals to morally enhance themselves.

In a book and in a series of articles, Ingmar Persson and Julian Savulescu have argued that there is an urgent need to explore the possibility of using the emerging science of morality to develop means to improve - *enhance* - moral dispositions (Ingmar Persson & Savulescu, 2008, 2012; Savulescu & Persson, 2012). Later they synthesized these ideas in a book: *Unfit for the Future: The need for Moral Enhancement* (Ingmar Persson & Savulescu, 2012). They argue that people's moral capacities are essentially 'flawed', and that there

5 During the course of this research project, the amount of journal articles, books, academic conferences, and media attention for moral enhancement has risen significantly. A systematic literature review we performed at the beginning of this project (up until April 2014) yielded 85 academic articles discussing moral (bio)enhancement (see chapter 2). The amount of publications on moral enhancement now (spring 2018) is nearing 300.

exists a fundamental mismatch between our moral psychology and today's conditions of human life (Ingmar Persson & Savulescu, 2013). In addition there is, they posit, a considerable and growing potential for "a significant improvement of various aspects of human cognition by biomedical means", the risk that a small but potent minority will acquire the power to do great harm, grows (Ingmar Persson & Savulescu, 2008). To counter this risk, they argue, "cognitive enhancement would have to be accompanied by a moral enhancement which extends to *all* of us, since such moral enhancement could reduce malevolence" (Ingmar Persson & Savulescu, 2008, p. 166). An extensive moral enhancement of humankind, in conclusion, is our only hope in averting "ultimate harm."[6]

They ask whether we could, through our knowledge of "the new sciences of behavioural control", strategically influence people's moral dispositions and behaviour (Savulescu & Persson, 2012), and discuss a range of lines of research that they feel could provide new and effective means of influencing human choices, such as priming (affecting choice by unconscious stimuli) and nudging (influencing individual choice by changing the choice architecture of the environment). In addition, they discuss how a series of currently used pharmacological substances may be used to affect moral behaviour as a side effect.

Persson and Savulescu have defended mandatory programs if people fail to morall enhance themselves voluntarily:

> If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them. (Ingmar Persson & Savulescu, 2008, p. 174; 2011see also: ; Ingmar Persson & Julian Savulescu, 2014b; Ingmar Persson & Savulescu, 2016)

More recently, they have argued that people even have a moral duty to morally enhance themselves (Ingmar Persson & Savulescu, 2017).

The God Machine

In a thought experiment not much different from Lem's futuristic story about *Altruizine* this chapter started with, Julian Savulescu and Ingmar Persson entertain the idea of

6 The concept of Ultimate Harm reminds of research on so-called existential risks (or x-risks); risks that could lead to human extinction or civilizational collapse (Bostrom, 2013), studied for example in the University of Cambridge based Centre for the Study of Existential Risk (https://www.cser.ac.uk/).

an all-powerful, self-learning, self-developing bioquantum supercomputer, named The God Machine that would intervene in human action to prevent "great harm, injustice or other deeply immoral behaviour from occurring" (Savulescu & Persson, 2012, p. 413):

> The God Machine would monitor the thoughts, beliefs, desires and intentions of every human being. It was capable of modifying these within nanoseconds, without the conscious recognition by any human subjects. (Savulescu & Persson, 2012, pp. 412-413)

Without the person involved being consciously aware, the God Machine would 'change his mind', for example by altering his intention to murder an innocent person. The machine is designed to ensure maximum individual freedom, Savulescu and Persson maintain: people's choices to do the morally right thing remain fully autonomous. It is only when a person is about to act in a clearly immoral way that the machine intervenes. Yet, the person in question would never be aware if the intervention, and as such she would still *experience* a life of complete freedom.

Would this God Machine be a bad idea, Savulescu and Persson ask, and, why would it be? The God Machine now makes impossible what in 'the old days' was prohibited by law, on pain of punishment. People are not free to murder now. Surely this small infringement of a person's autonomous choice is preferable to physical incarceration, or worse?

The primary purpose of this thought experiment has been to counter an objection by one of moral enhancement most outspoken critics, philosopher and bioethicist John Harris. He reasons that moral bioenhancement forms a fundamental threat to human freedom and autonomy, in particular to what he has called "the freedom to fall", the ability to choose based on reason and argument, and the ability to make an *ethical choice* (Harris, 2011, 2013a, 2014):

> Ethics is for those occasions on which altruism fails; or for those people who fail to think and feel, and/or who are not disposed to do as they should! Moral reasoning is needed to identify the appropriate objects for sympathy, empathy and the sort of generalized love that is the conclusion of a moral argument and which is often expressed as 'love thy neighbour as thyself'. We will, I believe, always need to use moral reasoning to act as a guide to our emotions and as a way of checking that we are having appropriate feelings in appropriate circumstances and for appropriate objects. (Harris, 2013a, p. 171)

Michael Hauskeller's critique on the God Machine focuses on the aspect of external manipulation and control the God Machine impersonates.

> We treasure human freedom, not the metaphysical one that very likely is a mere illusion, but the social, relational one. It defines what we are, our very humanity. From that perspective even the God machine poses a threat because the problem would still be that we are being controlled by a machine that has been designed with the explicit goal of controlling us. Whether we are actually aware of the manipulation or not is not relevant, or is not the main problem. The main problem is the fact that we are being controlled by, per definition, somebody else's will, and the fact that the purpose of our being controlled is that we conform to certain moral standards does not make it any better, but rather the contrary, because instead of being ends in ourselves, we thereby become means to the end of morality, which ultimately makes morality itself pointless because we no longer have a good reason to treat each other as ends if we are *in fact* nothing but means. (Hauskeller, 2017, p. 375)

The difficulty of defining moral enhancement

From the beginning, the debate on moral enhancement has been characterized by profound disagreements about the way the concept of moral enhancement should be defined. An exchange between neuroscientist Molly Crockett and ethicists John Harris and Sarah Chan can serve as an example (Crockett et al., 2010a, 2010b; Harris & Chan, 2010; Chan & Harris, 2011). According to Crockett et al., enhancing the neurotransmitter serotonin in the brain influences moral judgment and moral/pro-social behaviour, as their study shows that it diminishes the willingness to harm another person; it increases *harm aversion:* "our subjects were less likely to endorse personal moral harms and less likely to punish unfairness in the UG [Ultimatum Game]" (Crockett et al., 2010b, p. 17436). John Harris and Sarah Chan present an opposing view, and argue that serotonin seems to preclude and cloud moral judgment, rather than improving it.[7]

7 Harris Wiseman makes a similar point, mentioning both hesitations about the safety and efficacy of using SSRI's as a moral enhancement intervention, and concerns about whether SSRI's are "worthy of the name" of moral enhancement: "It is hard to imagine a serious moral outlook that can equate chemically induced pacification of behavior with a genuine moral improvement of the individual in question. Add to this that SSRIs cannot be made to work in a contextual fashion, as they are "all-or-nothing," brute agents that require periods of building up and tapering off, and that the only justifiable "populations" they might be effective in treating, assuming no abreaction occurs, are those with more serious mental health issues, in which case, it is more likely that stronger antipsychotics or sedatives will be used rather than SSRIs." (Wiseman, 2014b, p. 27).

> It is not clear that 'enhancing aversion to personally harming others' is something that would promote either moral behaviour in the public at large or indeed, harm reduction more generally, unless it was capable of much more nuanced effects than seems evident from the reported research. Thus, if serotonin affects moral behaviour, it does so adversely by impairing moral judgment, subjugating it to emotional instinct. We should be wary of assertions claiming that serotonin has a role in moral judgment; the opposite seems to be the case. It may enhance aversion to violence, but it does not enhance moral behaviour; it can increase, rather than diminish, harm to others and bypass the use of moral reasoning. (Harris & Chan, 2010, p. E183)

Underlying these disputes often lie profound philosophical battles about what constitutes morality. Harris believes that true moral growth can be achieved through cognitive means, by employing reason; he is a rationalist (Baertschi, 2014). Directly manipulating moral emotions, as advocated by for example Douglas (Douglas, 2008, 2013), would according to Harris lead to *moral decline*, not moral growth.

> Although both emotion and reasoning thus affect moral decisions, of the two, it must be reasoning that pulls in the direction of morality. A moral agent is not just someone who performs actions with moral consequences, she is a person who cares about doing the right thing. Such a person must have a way of deciding whether what her emotions prompt her to do, what strikes her as the right thing to do, really is the right thing to do. She will need to think things through, identify the relevant principles she accepts, the values she holds and the moral objectives she believes are right, and apply them to the present circumstances; and to do this she must use moral reasoning. (Chan & Harris, 2011, p. 130)

Taxonomy of definitions of moral enhancement, and their normative implications

In the following quotes, John Shook calls attention to the wide variety of definitions given:

> Too many discussions are proceeding as if both the meaning and the possibility of moral enhancement were already widely understood and agreed upon. (…) Asking such questions, and offering answers, depend on assigning some sense or another to "moral enhancement." However, clear and precise definitions of "moral enhancement" are not to be found; what has been called "moral" enhancement ranges from feeling empathic concern to increasing personal responsibility all the way to heightening respect for global fairness. (Shook, 2012, p. 3)

And:

> anyone using the term 'moral enhancement' as if everyone knows what is meant must either be simplifying matters to the point of negligence, or trying to speak only to those already in local moral consensus. (Shook, 2012, p. 4)

We identified different definitions of moral (bio)enhancement that have been given in the literature, and developed a taxonomy of different definitions and uses of the concept of moral (bio)enhancement as well as their normative implications (Raus et al., 2014).[8]

Some commentators provide a specific definition of moral enhancement. Thomas Douglas for example defines moral enhancement as:

> interventions that will expectably leave an individual with more moral (*viz.* morally better) motives or behaviour than she would otherwise have had. (Douglas, 2008, p. 229)

Other commentators do not offer such an explicit definition, and for the purpose of this taxonomy we therefore had to reconstruct the way the concept was used by the different authors.

Focus on the intervention or on the individual

The first distinction concerns different interpretations of the question *what is moral enhancement?* Whereas some definitions focus on the intended effect of a particular *intervention* or technology (i.e., "interventions that are intended to improve our moral capacities such as our capacities for sympathy and fairness" (DeGrazia, 2014, p. 361)), other definitions focus on questions about when an individual can be considered to have been morally enhanced, or what it means to say that an *individual* is morally enhanced. James Hughes for example argues that moral enhancement goes beyond "the jacking up of virtue with neurochemicals", but that it should be conceived more broadly, in terms of "taking conscious control of our lives to build the kind of character we want to have" (Hughes, 2011).

---

8 This paragraph draws on a paper by Kasper Raus et al., of which I am a co-author: Raus, K., Focquaert, F., Schermer, M. H. N., Specker, J., & Sterckx, S. (2014). On defining moral enhancement: A clarificatory taxonomy. *Neuroethics, 7*(3), 263-273. When discussing the different definitions and uses of moral enhancement in this paragraph, 'we' refers to the authors of this paper.

Broad versus more specific interventional means

Another relevant distinction we found concerns broad versus more specific definitions. Some authors prefer a more restrictive definition of moral enhancement that focuses on biomedical and genetic means only. These authors also tend to prefer the concept of moral *bio*-enhancement or *genetic virtue* (Walker, 2009, 2010). Others include any practice that brings about – or intends to bring about – positive changes in a person's moral capacities, including practices such as (moral) education and (talk) therapy. Because genetic and biomedical interventions (for example pharmaceutically or by means of (deep) brain stimulation) are generally considered more invasive and more precarious than non-biomedical ones, it might be expected that using "the concept of moral enhancement for *all* interventions or only for those interventions that are (most) invasive, can serve to bias the normative debate towards or away from a conclusion of permissibility or desirability" (Raus et al., 2014, p. 265).

Enhancing individuals versus enhancing humanity

A further way in which definitions of moral enhancement differ, is based on the question who is considered the target of moral enhancement: individuals, or groups of persons? Thomas Douglas, especially in his earlier work on moral enhancement, explicitly focuses on individual persons, as illustrated by the previously mentioned definition. Ingmar Persson and Julian Savulescu on the other hand, have advocated enhancing "the moral character of humanity" (Ingmar Persson & Savulescu, 2008), and have defended mandatory moral enhancement *programs*, aimed specifically at those most in need of moral improvement. As we argued in the taxonomy, these definitions have profound normative implications. For justifying moral enhancement applied solely to individuals, justifications related to autonomy and personal benefit might suffice. However, in justifying moral enhancement on a societal level one would need to turn to other justifications such as issues of justice or of achieving a common good. Moreover, concerns about altering human nature would then come into view as well.

Moral treatment versus moral enhancement

A further distinction concerns the distinction between moral treatment and moral enhancement. Whereas some definitions understand moral enhancement to be any form of moral improvement, regardless of whether a person's moral functioning before enhancement could be considered below average, average, or above average; other definitions clearly damarcate bringing people to an average or 'normal' level of moral functioning from enhancing beyond the average level. Improving towards average would be considered moral treatment, improving beyond average would be

considered enhancement.[9] Yet, defining moral normalcy raises questions. As there is no objective way of determining what falls within the range of 'normal' moral behaviour or functioning of moral capacities, every classification is, necessarily, a normative one.

John Shook and James Giordano have proposed an alternative classification between "moral (re)habilitation" (instilling or restoring some degree of moral capacity and responsibility in someone); "moral normalization" (an improvement of already-existing moral capacity towards society's standard of good moral conduct); and "surpassing enhancement" (improvement above regular requirements of common morality) (Shook & Giordano, 2016b, 2016a, 2017).

Intended versus effective interventions

A further difference concerns definitions that focus on either the intention to enhance versus interventions that actually have a positive effect. John Harris for example understands moral enhancement interventions to be those interventions that have an actual effect: "I do not define enhancements in terms of the intention or the motivation of those who produce them but rather in terms of their effect" (Harris, 2014, p. 372). On the other hand, the definition by Douglas given above, and the definition by David include *intended or expected but failed* interventions (Douglas, 2013; DeGrazia, 2014):

> Interventions that are intended to improve our moral capacities such as our capacities for sympathy and fairness. (DeGrazia, 2014, p. 361)

For them, what is important in determining whether an intervention is to be considered a moral enhancement is the *intention* with which the intervention is used, not its actual effect. This distinction between intended and actual effect is relevant for morally justifying moral enhancement. Because if an intervention can only be labelled a moral enhancement if it has a positive effect, then one important reason to oppose moral enhancement is eliminated. All problematic cases where no enhancement is reached or a person is left *less* moral than before the intervention, would not be considered cases of moral enhancement at all. A consequence of this way of justifying moral enhancement is that it would avoid debate concerning potential risks and side-effects of interventions.

9 As Dorothee Horstkotter and colleagues argue: "if there is a health problem, medical treatment is the reasonable reaction, while enhancement, either moral or otherwise, does not arise." (Horstkötter, Berghmans, & de Wert, 2012)

Capacities-oriented versus behaviour-oriented interventions

Some commentators in the moral enhancement debate label a certain intervention a moral enhancement depending on its (real or intended) effect on a person's *behaviour*. Others see moral enhancements as interventions that target or are intended to target a person's *capacities of moral reflection*. In the taxonomy, we have referred to this difference as the difference between a behaviour-oriented and a capacities-oriented intervention.

If moral enhancement interventions can target a person's behaviour or their capacities of moral reflection, any intervention can have one of four results. It can result in a person (1) reflecting the same, and acting the same; (2) reflecting the same, but acting differently; (3) reflecting differently, but acting the same; and (4) reflecting differently, and acting differently. Definitions of moral enhancement differ as to the question which of these should be considered cases of moral enhancement. In identifying which interventions can be labelled moral enhancements, some commentators rely heavily or solely on the intervention's achieved (or intended) effect on a person's behaviour [e.g. (Douglas, 2013)]. For those commentators, moral enhancement occurs only when a person, because of an intervention, *acts* differently. According to them, if a person reflects differently, but acts the same because of an intervention, this should not be considered a case of moral enhancement. A range of other authors do not think that the requirement for successful behaviour change suffices to label an intervention a moral enhancement, and instead categorise them as a form of behaviour control (e.g. Harris) or moral therapy (e.g. Agar):

> I take moral enhancement to involve enhancing our ability to think ethically (...), not manipulating the probability of some reacting in ways that *others* deem ethical. (Harris, 2014, p. 373)

> While the manipulation of moral emotions might change the behaviour of an individual, it does not provide any content, for example, norms or values to guide one's behavioural response. (Jotterand, 2011, p. 6)

> (i) forcing agents to act rightly, (ii) preventing agents from acting wrongly, and (iii) making it harder for moral agents to act wrongly fail to constitute genuine moral enhancement. (Simkulet, 2012, p. 17)

For these commentators, as well as for those who use similar lines of argumentation (e.g. (Lev, 2012; Baertschi, 2014)), the criterion for whether an intervention constitutes a moral enhancement is not (only) behaviour, but rather whether the intervention (also) affects or improves an individual's capacities for moral reflection. In the quote

below, Harris appears to embrace the most ambitious and restrictive definition of moral enhancement; acting differently and reflecting differently:

> It seems to me that moral enhancement, properly so called, must not only make the doing of good or right actions more probable and the doing of bad ones less likely, but must also include the understanding of what constitutes right and wrong action. (Harris, 2014, p. 172)

Active involvement versus passive receiving

A final distinction is that between a process of moral enhancement in which the individual is actively involved, versus a process in which the person is a passive recipient. This distinction does not concern the *target* of a moral enhancement, but rather *the way in which the enhancement is achieved*. A moral enhancement by way of active involvement would then be an enhancement requiring conscious mental processes in the subject as a means to achieve its result. Moral education would be a classical example. In contrast, moral enhancement involving passive receiving would either be enhancement in an immediate way where no active involvement was possible (e.g. a pill with immediate effect) or enhancement by way of a process that required no deliberate involvement of the recipient (e.g. classical conditioning).

This distinction is relevant since, for some authors, the goal of moral enhancement is for individuals to become more virtuous, and often these authors follow Aristotle in claiming that becoming virtuous is always a conscious and deliberate process, where the way to becoming virtuous is just as essential as the result (being virtuous). Chris Zarpantine talks about "the thorny and arduous path of moral progress" (Zarpentine, 2013, p. 141), while Jotterand states:

> Virtue is a behavioral habit under the supervision of reason that can be taught and learned. The control and manipulation of moral emotions by technological means reduce the human mind to neurochemical processes and threaten the very essence of moral agency, that is, autonomy. (Jotterand, 2011, p. 7)

The distinction is used to argue for the idea that the means by which moral enhancement is achieved matter morally (Focquaert & Schermer, 2015). Interventions where the individual is passive (more likely in the case of interventions that work directly on the

brain), Focquaert and Schermer argue, are more likely to compromise autonomy and disrupt identity than interventions in which the individual is an active participant.[10]

Conclusion: Definitions are not normatively neutral

As discussed above, definitions of moral enhancement vary widely – Harris Wiseman speaks of the "tremendous plurality the term conceals" (Wiseman, 2017, p. 398) – and have important normative implications. The chosen definition conveys which practices an author considers more or less ethically desirably. Moreover, the concept of moral enhancement sometimes appears to assume that improving a person's morality is unproblematically good – why else would it be called *moral* enhancement? Finally, as the sciences of morality are evolving, it is not yet clear what will be possible in the future, what kind of moral enhancement technologies will be developed, and how people will react to and interact with them.

Therefore, to a certain extent, this thesis takes the moral enhancement debate itself as the main unit of investigation. It asks in what ways the concept of moral enhancement is used, what kind of reasons participants in this debate have given for or against its desirability, what kind of interventions they have proposed, and what kind of problems they think it will solve. To be able to analyse these aspects, it is unhelpful to have a pre-defined notion of what 'moral enhancement' is or should be.

In this thesis the concept of moral enhancement is used to refer to attempts to influence, change, or optimize moral capacities, and limit immoral capacities, with the aim of changing a person's moral emotions, moral reasoning, or moral behavior. This includes instances where someone (e.g. a practitioner or policy maker) discusses or implements interventions that have an effect on a person's moral capacities, but that person does not necessarily discuss those interventions in explicitly *moral* terms. Finally, this working definition is not limited to biomedical interventions but also includes nonbiomedical, traditional means.

This thesis: Commitment to practices

With the ambition of counterbalancing the highly polarized debate and offering a more nuanced vision of what moral enhancement entails, this thesis focuses on (present and emerging) moral enhancement *practices*. The ultimate goal of doing so is to

10 Whether means matter morally has been a topic of intense debate (see chapter 2 and 6). See also Reichlin (Reichlin, 2017), and Bublitz (Bublitz & Merkel, 2014; Bublitz, 2015, 2016).

identify ethical issues that are not necessarily part of the current debate on moral enhancement.

Similarly, Harris Wiseman has advocated a 'practical-realities first' approach to potential moral bioenhancement interventions, implying that speculation about moral bioenhancement should account for "the specific practical realities to be found on the ground level, which are not at all incidental but the very realities around which the abstractions of the debate must be made to shape themselves (not the other way around)" (Wiseman, 2016, p. 13). A focus on practices can function as a correction to overly wild[11] scenarios, and at the same time call attention to practices where elements of stimulating moral betterment are visible. Much of what is being discussed is not possible yet, if ever (bringing some authors to question whether moral enhancement should be considered science fiction, rather than science fact (Dubljević, 2017; Dubljević & Racine, 2017)).[12]

Arguably, in previous enhancement debates, one of the central ethical questions concerned the desirability of expanding upon current, therapeutic interventions or practices, and to use them "beyond therapy" (Kass, 2003) or to go "beyond what is necessary to sustain or restore good health" (Juengst, 1998, p. 29). The quite extensive literature about the desirability of the non-medical use of prescription drugs meant to treat ADHD (Attention Deficit Hyperactivity Disorder) can serve as an example (Greely et al., 2008; Schermer, 2014). In the case of moral enhancement, it is often far less clear what that current practice is, and in what way the proposals made would impact

11 There is no straightforward answer to the question on how to decide whether a certain scenario should indeed be considered overly wild, or, alternatively, represents an important thought experiment. Arguably, imagination is key for scientific discovery (de Mey, 2006) and ethical analysis alike. At the same time, an important normative function of thought experiments might be to stimulate imagining of "possible usages and desirabilities" (van de Werff et al., 2016, p. 98) in order to expose and elucidate relevant ethical issues. An overly wild or speculative scenario arguably "violates conditions of intelligibility, squanders the scarce and valuable resource of ethical concern, and misleads by casting remote possibilities or philosophical thought-experiments as foresight about likely technical developments. In effect, it deflects consideration from the transformative technologies of the present" (Nordmann, 2007, p. 31, and see chapter 3). Björn Hofmann proposes to introduce a so-called *hype test,* which reads as follows: "*What is special with this technology (e.g., gene editing), compared to existing methods, that makes it succeed in improving human social characteristics to make the world a better place for all?*" (Hofmann, 2018).

12 As Stephan Schleim argues: "notwithstanding the influences of neuro-collaborations and related funding schemes within academia, we scholars indeed collectively overestimate the practical and translational social impact of this research so far. The communicated promises as well as the scholarly and public attention given to these possibilities are, in my view, in no way justified by the scientific possibilities" (Schleim, 2014, p. 3), also see (Crockett, 2014b).

on that practice. Oftentimes, what an author considers relevant practices remains implicit, or is discussed in very general terms only.

For example, one domain that is often implied, but not made explicit, is the field of forensic psychiatry. Participants in the moral enhancement debate have discussed the use of neuro-interventions for offenders or forensic patients who are suffering from various cognitive, motivational and emotional impairments as examples of moral enhancement. As such impairments may involve risk factors for various kinds of immoral behaviour (e.g., sexual crimes, violence, and racism), proponents argue that moral bioenhancement could provide new ways to achieve successful recidivism reduction and rehabilitation (Douglas, 2008; S. Carter, 2016).

However, the diversity of definitions of moral enhancement that are available suggests that not everyone will agree that psychiatric treatments that address neurobiological risk factors for deviant behaviour should be understood as proper instances of moral enhancement (Raus et al., 2014; Reichlin, 2017). Nonetheless, outspoken proponents of moral bioenhancement Persson and Savulescu have argued that a number of psychiatric disorders can be characterized as "moral defects", and therefore, that treating these disorders should indeed be understood as moral enhancement (Savulescu & Persson, 2012).

> The opposite of promoting another's interests is damaging another's interests. Traits which increase harm to others cause immoral behaviour. The paradigm is psychopathic personality disorder, but other personality disorders such as antisocial personality disorders, borderline personality disorder and narcissistic personality disorder can cause great harm to those who come into contact with these individuals. The reduction in these tendencies are thus moral enhancements. (Savulescu & Persson, 2012, p. 410)

Likewise, DeGrazia has characterized the treatment (or prevention) of antisocial personality disorder as a uncontroversial example of moral enhancement (DeGrazia, 2014), and Douglas has discussed institutions of criminal justice as institutions that are arguably "already engaged in a kind of moral enhancement" (Douglas, 2014c, p. 1245).

To conclude, intended users, contexts of implementation, and the goals and objectives of developing and applying potential biomedical possibilities for moral improvement are often not specified. In order to expose and elucidate relevant ethical issues concerning moral enhancement, as well as to formulate ethical requirements, this thesis focuses on potential *practices* of moral enhancement.

The following questions have been guiding my research:

1. What is the debate on moral bioenhancement about?
   a. What do participants in this debate understand moral bioenhancement to be? How do they define and demarcate moral bioenhancement?
   b. What is the rationale proponents give for the need for moral bioenhancement and what are the main ethical arguments given for and against moral bioenhancement?
   c. What ethical issues are most prominent in the debate?
2. What is a sensible way to approach the question of the ethical desirability of moral bioenhancement?
   a. From where should normative analysis start?
   b. For which professional practices and for whom is the debate on moral bioenhancement relevant? Who should care about/ participate in the debate?
   c. Can a focus on potential *practices* of moral enhancement help identify ethical issues that are relevant, but left out or under examined in the debate so far?
3. What are conditions and ethical requirements for ethically justifiable moral enhancement practices?

Methodology

In order to answer these research questions, in this thesis, a variety of research methods is employed. In order to get a better grasp of potential practices of moral enhancement, in this thesis I have employed a mixed methods approach, making use of empirical research methods such as qualitative interviews and vignette studies, as well as of more traditional philosophical methods such as conceptual and argumentative analysis.

In his book *Experiments in ethics* Kwame Anthony Appiah warns against considering anything experimental as alien to philosophy (Appiah, 2008). Philosophy nowadays is mostly defined in terms of what it is not – psychology, physics, anthropology. Appiah argues against this crude and a-historical view on what philosophizing consists of, and consisted of in the past: the cliché of a separate group of 'professional thinkers' that look down on anything not purely theoretical and "confined themselves to the realm of unsullied abstraction" (Appiah, 2008, p. 9). Doing so misrepresents the interests and fascinations of many of the canonical philosophers: "what's novel isn't the experimental turn; what's novel was the turn away from it" (Appiah, 2008, p. 6).

Also (bio)ethics has witnessed what has been called "an empirical turn" (Hope, 1999; Molewijk et al., 2004; Sugarman, 2004; Borry et al., 2005), indicating that increasingly normative analysis is being supplemented with empirical input, as well as the working together of more empirical and more philosophical oriented researchers. A recent systematic review of methodologies used in bioethical research demonstrates a striking heterogeneity of methodologies, both quantitative and qualitative (Davies et al., 2015).

There is an extensive literature discussing the benefits and downsides of integrating empirical research with normative analysis and appropriate ways of doing so. Strong et al. list a range of familiar advantages of incorporating empirical research in bioethics, next to philosophical reasoning (Strong et al., 2010). The strength of empirical research in bioethics, they write, is that it allows for:

> The description of the experience of individuals or populations with respect to morally relevant issues; the description of the attitudes, beliefs, moral opinions, reasoning patterns and decision-making of those involved in a certain practice; the generation of data which can challenge authority, dogma, convention, norms and experience by showing how practice varies; and the identification of moral issues that have escaped the attention of ethicists, but are relevant in a specific context, including those that are not obvious because they are embedded in practice. (Strong et al., 2010, p. 318)

Central challenges concern both the empirical methodology, and the integration of empirical and normative research. With regards to methodology, it can be questioned what kind of empirical information is most useful – the opinions, arguments or expressed moral views of various stakeholders, or perhaps rather their actual moral actions and decisions? And which stakeholders should be included? Moreover, what are the best ways to collect data on stakeholders' moral views: open interviews, questionnaires, focusgroups, or vignettes? As regards the integration of empirical data in moral reasoning a number of models have been proposed, such as "integrated empirical ethics" (Molewijk et al., 2004), "broad reflective equilibrium" (van Thiel & van Delden, 2010), "reflexive balancing" (Ives, 2014), and "symbiotic empirical ethics" (Frith, 2012), just to name a few.

In my opinion, the most challenging issue concerns the question what type of normative claim or conclusion one can generate through conducting or studying empirical research. Davies et al. distinguish two "poles of methodological orientation", one *dialogical*, the other *consultative* (Davies et al., 2015). In dialogical approaches "the ethical analysis and reaching of a normative conclusion is part of the research encounter

itself" (Davies et al., 2015, p. 8), whereas in consultative approaches ethical analysis is "undertaken post-engagement, by the researcher or the research team. (…) stakeholders feed into the ethical analysis, but are not involved in it directly" (Davies et al., 2015, p. 9). The approach to empirical data generated in the context of this thesis is mostly consultative in nature.

A review from 2012 maps empirical research on human enhancement specifically (Schuijff & Munnichs, 2012). Empirical studies on enhancement employ a diversity of methodologies, ranging from more quantitatively to qualitatively oriented, and mostly focus on studying public attitudes and preferences. The authors argue that empirical research, and research on public opinion specifically, can enrich the largely expert-only debate on human enhancement (Dijkstra & Schuijff, 2015). Because, they argue, consulting the public may aid in better understanding what shapes public preferences, and this may help in policy making. Ethical analysis of potential practices of moral enhancement indeed needs to include and account for views of relevant stakeholders such as potential users, potential medical professionals or 'operators', policy makers, and the general public. Indeed; "the question of moral bioenhancement is a paradigmatic instance in which understanding public attitudes is critical" (Reiner, 2013, 2017).

Outline of this thesis

Chapter 2 contains a systematic clustering of ethical arguments that emerge in the moral enhancement debate (based on a systematic search of the literature). Each argument is discussed separately, and the debate as a whole is assessed. It is concluded, first, that there is little discussion on what distinguishes moral bioenhancement from treatment of pathological deficiencies. Furthermore, remarkably little attention has been paid so far to the safety, risks and side-effects of moral enhancement, including the risk of identity changes. Finally, many authors overestimate the scientific as well as the practical feasibility of the interventions they discuss, rendering the debate too speculative.

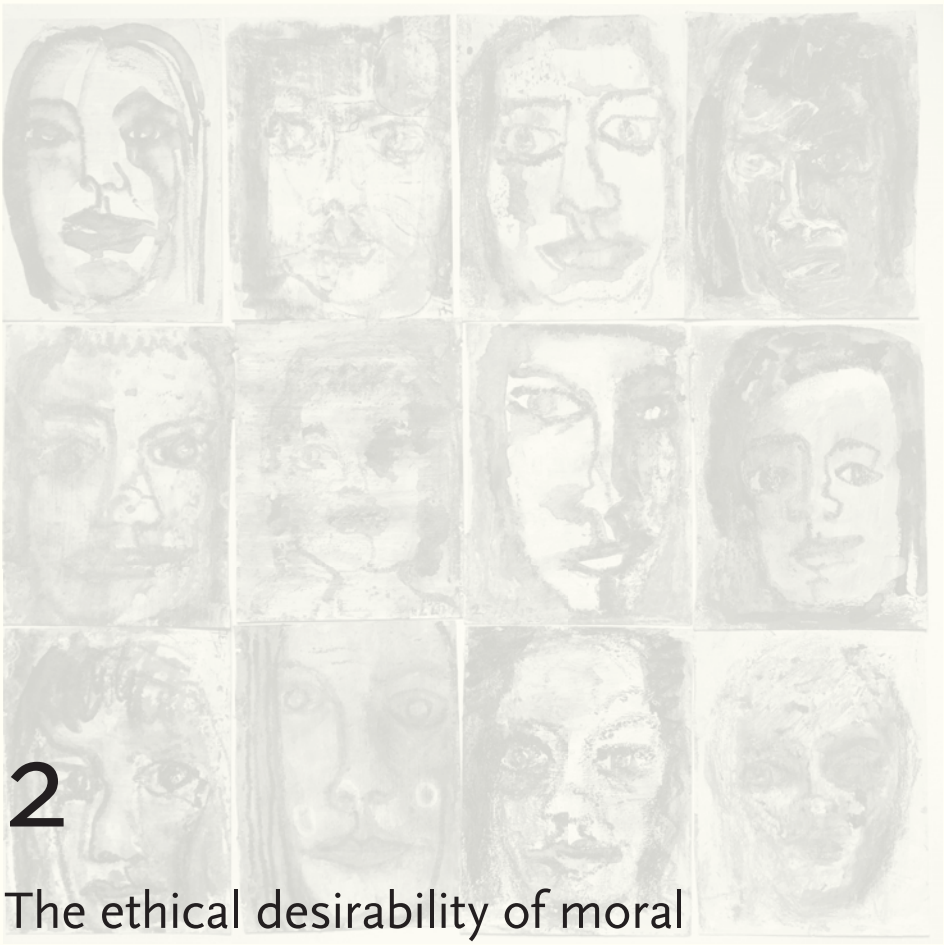
Chapter 3 discusses a number of contexts or domains in which (future) moral bioenhancement interventions possibly or most likely will be implemented. By looking closely at similar or related existing practices and their relevant ethical frameworks, we identify ethical considerations that are relevant for evaluating potential moral bioenhancement interventions. Domains that are examined are, first, debates on the proper scope of moral education; second, proposals for identifying early risk factors

for antisocial behaviour, and third, the difficult balancing of the best interests of the individual and third party concerns in (forensic) psychiatry.

Chapter 4 and **Chapter 5** report on a series of interviews with forensic practitioners on the prospects, threats, and limitations of integrating neurobiological and behavioural genetic interventions in forensic psychiatric practices. Forensic practitioners were interviewed about their expectations as well as moral views regarding potential applications of current neurobiological and behavioural genetic research aiming to understand (and possibly help prevent, contain, or treat) violent and antisocial behaviour. Potential applications that were included were potential biomedical possibilities to lower aggression, the possible usage of neuroimaging in assessing legal responsibility, and the potential use of biomarkers in assessing risk for future violent and antisocial behaviour. Furthermore, these interviews explored to what extent forensic practitioners consider moral development and moral growth to be a part of their current professional practices and to what extent they think that stimulating moral development is a legitimate objective in the context of forensic psychiatric treatment. In addition, we asked how forensic practitioners balance public safety and risk management concerns with the interests and wellbeing of the individual patient.

Chapter 6 reports on an empirical study aimed at gaining insight into the reasons that the public may have for endorsing or eschewing pharmacological moral enhancement for themselves or for others. The study is driven by three hypotheses: first, the hypothesis that the degree to which members of the public support an empathy-enhancing moral enhancement program depends on whether or not the means employed are pharmacological or non-pharmacological. Second, that people are less supportive of pharmacological moral enhancement of their own children than they are of other people's children. Third, that the degree to which respondents support these programs depends on whether they imagine themselves or someone outside their immediate circle of concern to participate. In other words, we expect that the distinction between self and other is relevant.

Chapter 7 reflects on the main findings of the preceding chapters, as well as their limitations. It revisits the main research question, formulates normative conclusions, and suggestions for future research.



2

The ethical desirability of moral bioenhancement: A review of reasons

Specker, J., Focquaert, F., Raus, K., Sterckx, S., & Schermer, M. H. N. (2014). *BMC Medical Ethics, 15*(1), 67.

Abstract

Background

The debate on the ethical aspects of moral bioenhancement focuses on the desirability of using biomedical as opposed to traditional means to achieve moral betterment. The aim of this paper is to systematically review the ethical reasons presented in the literature for and against moral bioenhancement.

Discussion

A review was performed and resulted in the inclusion of 85 articles. We classified the arguments used in those articles in the following six clusters: (1) why we (don't) need moral bioenhancement, (2) why it will (not) be possible to reach consensus on what moral bioenhancement should involve, (3) the feasibility of moral bioenhancement and the status of current scientific research, (4) means and processes of arriving at moral improvement matter ethically, (5) arguments related to the freedom, identity and autonomy of the individual, and (6) arguments related to social/ group effects and dynamics. We discuss each argument separately, and assess the debate as a whole. First, there is little discussion on what distinguishes moral bioenhancement from treatment of pathological deficiencies in morality. Furthermore, remarkably little attention has been paid so far to the safety, risks and side-effects of moral enhancement, including the risk of identity changes. Finally, many authors overestimate the scientific as well as the practical feasibility of the interventions they discuss, rendering the debate too speculative.

Summary

Based on our discussion of the arguments used in the debate on moral enhancement, and our assessment of this debate, we advocate a shift in focus. Instead of speculating about non-realistic hypothetical scenarios such as the genetic engineering of morality, or morally enhancing 'the whole of humanity', we call for a more focused debate on realistic options of biomedical treatment of moral pathologies and the concrete moral questions these treatments raise.

Background

Should we develop and implement interventions that aim to improve people's morality? Ever since the publication of two papers in a special issue of the *Journal of Applied Philosophy* in 2008 (Douglas, 2008; Ingmar Persson & Savulescu, 2008), the ethical desirability of moral bioenhancement has been the subject of intense debate. Whereas 'traditional methods' of moral betterment (such as upbringing, socialization and education) are arguably as old as humanity itself, the debate on moral bioenhancement focuses specifically on the desirability of the use of biomedical methods. Interventions that are being investigated in the literature range from various types of psychopharmaceuticals, deep brain stimulation (DBS), and genetic selection and engineering.

In a previous paper (Raus et al., 2014) we examine the different ways in which the concept of moral bioenhancement is used in the literature: what different authors understand its main goals to be, what would count as a success, and what kind of interventions would and would not fall within their proposed definitions. In this paper, we ask what reasons and arguments have so far been given in the debate on the ethical desirability of moral bioenhancement. We do this by mapping out the different arguments that have been presented in the debate up till now (see subsection 'Arguments used in the debate'). We aim to provide a complete overview including both the main arguments in the debate as well as the less commonly voiced arguments. In the final critical appraisal section (see subsection 'Critical appraisal of the current debate'), we analyze the kinds of arguments given, thereby critically assessing the issues and concerns that have been discussed, as well as identifying those issues and concerns that up to now have been neglected. This section represents our own interpretation and views concerning the debate and the arguments that are given. We argue for a shift in focus of the debate towards a discussion of more realistic interventions for specific target groups.

Discussion

Methodology

In order to give a comprehensive overview of the debate so far, we conducted a literature search to collect all publications that discuss 'moral (bio)enhancement' since the start of the debate in 2008. With assistance from a reference librarian, we selected suitable databases in bioethics. In September 2013, we searched these databases to find all publications that mention moral (bio)enhancement. For the specific search terms and strings per database, please consult Table 1. The results of these searches were downloaded to Endnote, and duplicates were removed.

Table 1: Search terms and strings

Database	Search string
Embase	(morality/de AND ('genetic enhancement'/de OR 'medical technology'/de OR 'neurosurgery'/de)) OR (((moral* OR virtue* OR virtuous OR biomedical* OR bio-medical) NEAR/6 (enhanc* OR bioenhanc* OR manipulat*))):ab,ti
Medline OvidSP	((morals/ OR Moral Development/ OR Virtues/) AND ("Genetic Engineering"/ OR exp "Biomedical Enhancement"/ OR "neurosurgery"/)) OR (((moral* OR virtue* OR virtuous OR biomedical* OR bio-medical) ADJ6 (enhanc* OR bioenhanc* OR manipulat*))).ab,ti.
Web-of-science	TS=(((moral* OR virtue* OR virtuous OR biomedical* OR bio-medical) NEAR/6 (enhanc* OR bioenhanc* OR manipulat*)))
PsycINFO OvidSP	((morality/ OR Moral Development/ OR Virtue/) AND ("Genetic Engineering"/ OR "neurosurgery"/)) OR (((moral* OR virtue* OR virtuous OR biomedical* OR bio-medical) ADJ6 (enhanc* OR bioenhanc* OR manipulat*))).ab,ti.
PubMed publisher	(Moral enhanc*[tiab] OR Moral bioenhanc*[tiab]) NOT medline[sb]
Scopus	TITLE-ABS-KEY(((moral* OR biomedical* OR bio-medical) W/3 (enhanc* OR bioenhanc*)) AND (ethic* OR bioethic*))
Google Scholar	"moral (enhancement|bioenhancement|enhancing)"|"moral bio enhancement"
Scirus – preferred web/ ProQuest	"Moral enhancement" OR "Moral bioenhancement"

Table 2: Selection of publications

Database	Initial results	Results after deduplication
Embase	1027	1008
Medline OvidSP	820	178
Web-of-science	1191	788
Scopus	449	261
PsycINFO OvidSP	427	253
PubMed publisher	17	10
Google Scholar	192	142
ProQuest	75	58
Scirus	5	3
Total	4203	2701

Based on title and abstract, we excluded all articles that are not directly related to moral bioenhancement. As the main aim of this paper is to provide an overview of the ethical reasons for and against moral bioenhancement in the debate so far, we included only those publications in which authors explicitly mention moral bioenhancement. We excluded from our analysis publications on the moral status of post-persons, unless there was an explicit reference to the debate on the desirability of moral bioenhancement. We also excluded those publications that discuss moral bioenhancement but were not written in English (N=14). We discussed those publications that we were less sure about (N=99) until consensus was reached. In April 2014 we repeated the exact same search, in order to retrieve all publications that were published in the interven-

ing period (N=22). All in all, 85 publications were included. For a schematic overview of the selection process, see Figure 1 (Appendix A, p. 65).

We read the full-text of all articles and conducted a thorough thematic document analysis, in which we identified and coded each argument for or against moral bioenhancement mentioned in each publication. Based on this analysis, we formulated six broad clusters of arguments: arguments on the need for moral bioenhancement, on the possibility of attaining sufficient agreement on what moral bioenhancement should involve, on the status of current scientific research, on whether means and processes matter with respect to the desirability of moral bioenhancement, on the effects on the identity and the autonomy of the individual, and finally on the social effects of moral bioenhancement. This clustering was further refined and complemented on the basis of the analysis of all included publications, resulting in the final categories and subcategories that can be found in this article (see Table 3 for an overview of the arguments and sub-arguments we identified).

We have conscientiously attempted to provide a neutral and comprehensive review of the existing arguments, by clearly separating the description of the arguments (see subsection 'Arguments used in the debate') from our critical appraisal of the arguments (see subsection 'Critical appraisal of the current debate').

Arguments used in the debate

Table 3 provides an overview of the clusters of arguments and sub arguments we identified, as well as an overview of the authors addressing the specific argument. The arguments are formulated in a neutral way, and are almost always used by some to argue in favor of moral bioenhancement and by others to argue against it.

Below we will present the arguments we identified in the literature, organized in the following six clusters: (1) why we (don't) need moral bioenhancement, (2) it will (not) be possible to reach consensus on what moral bioenhancement should purport, (3) the feasibility of moral bioenhancement and the status of current scientific research, (4) means and processes of arriving at moral improvement matter ethically, (5) arguments related to the freedom, identity and autonomy of the individual, and (6) arguments related to social/ group effects and dynamics.

In the next section, we will describe these arguments in greater detail by summarizing these six clusters and their components. Given the richness of the publications we

Table 3: Arguments for and against moral bioenhancement, presented in the reviewed literature

Cluster	Argument	Description/ background
1. Why we (don't) need moral bioenhancement		
	There is scope for improvement	Almost by definition, each person can be/ act/ behave better. We therefore all have a moral duty/ imperative/ reasons to enhance ourselves. We have good reasons for wanting to better ourselves. Also: a duty to do the right thing.
	Human biological nature is defective	Humans are innately evil. Evil cannot be eradicated by socialization and education alone. Or: humans are innately good.
	Traditional means are (not) effective enough	Such as education, upbringing, socialization. These will only bring us so far. Or: they do suffice, are attractive and effective.
	Our only hope in averting major disaster	Avoidance of ultimate harm. Some of the world's most important problems can be attributed to moral deficits of individuals. Or: those problems have other causes besides the moral deficits of individuals. Moral enhancement should accompany, or even precede/ prioritize over cognitive enhancement and scientific progress.
	Moral bioenhancement might reduce criminality	Promise of solving immoral and criminal acts. Or: warning that these are not necessarily the same.
2. It will (not) be possible to reach consensus on what moral bioenhancement should purport		
	No consensus on the mechanisms that comprise our moral psychology	The way we should interpret neurobiological findings.
	Behavioral changes alone are (not) enough	Emotions versus moral reasoning. Dependent on view on what is considered worthy of moral appraisal. Behavioral control, or: certain attitudes towards behavior are also necessary (they have cognitive content).
	Ethical systems and theories differ and often disagree	Subjectivity of/ disagreement between main (substantive) moral theories. Individuals and cultures differ, there is moral pluralism. Possibility of being neutral between different conceptions of the good.
	(Im)possibility of considerable consensus	The question whether we can find a common ground, despite moral pluralism. Also: discussions on relativism/ nihilism, objectivism.

	Key articles
	(Douglas, 2008; Faust, 2008; Ingmar Persson & Savulescu, 2008; Agar, 2010; Fröding, 2011; Douglas, 2013; Jones, 2013; Douglas, 2014b; Jebari, 2014; Macer, 2014; Ram-Tiktin, 2014; Beck, 2015; Kahane & Savulescu, 2015)
	(Ingmar Persson & Savulescu, 2008; Walker, 2009; I. Persson & Savulescu, 2010; Sprinkle, 2010; Walker, 2010; Harris, 2011; Hughes, 2013; Zarpentine, 2013; Ingmar Persson & Julian Savulescu, 2014b; Rakić, 2014b; Beauchamp, 2015; Casal, 2015; Hauskeller, 2015; Kahane & Savulescu, 2015; Ingmar Persson & Savulescu, 2015b)
	(Ingmar Persson & Savulescu, 2008; Walker, 2009; Bronstein, 2010; Harris & Chan, 2010; I. Persson & Savulescu, 2010; Harris, 2011; Christen & Narvaez, 2012; Ehni & Aurenque, 2012; Harris, 2013b; Ingmar Persson & Savulescu, 2013; Tonkens, 2013; Zarpentine, 2013; DeGrazia, 2014; Rakić, 2014b; Agar, 2015b; Ingmar Persson & Savulescu, 2015b)
	(Douglas, 2008; Ingmar Persson & Savulescu, 2008; Walker, 2009; Fenton, 2010; I. Persson & Savulescu, 2010; Harris, 2011; Pacholczyk, 2011; Ingmar Persson & Savulescu, 2011; Rakić, 2012; Tennison, 2012; Gunson & McLachlan, 2013; Hughes, 2013; Ingmar Persson & Savulescu, 2013; Tonkens, 2013; Zarpentine, 2013; DeGrazia, 2014; Rakić, 2014b, 2014a; Selgelid, 2014; Wasserman, 2014b; Agar, 2015b; Beck, 2015; J. A. Carter & Gordon, 2015; Hauskeller, 2015; Ingmar Persson & Savulescu, 2015b)
	(Walker, 2009; Brooks, 2012; Beck, 2015; Casal, 2015)
	(Douglas, 2008; Ingmar Persson & Savulescu, 2008; Walker, 2009; Bruni, 2011; Chan & Harris, 2011; Fröding, 2011; Jotterand, 2011; Schaefer, 2011; Shook, 2012; Douglas, 2013; Triviño, 2013; Baertschi, 2014; Lechner, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Beck, 2015)
	(Ingmar Persson & Savulescu, 2008; Spence, 2008; Walker, 2009; Chan & Harris, 2011; Harris, 2011; Jotterand, 2011; Schaefer, 2011; Christen & Narvaez, 2012; William Paul Kabasenche, 2012; Rakić, 2012; Shook, 2012; Simkulet, 2012; Tennison, 2012; Douglas, 2013; Harris, 2013a, 2013b; Harris, 2013b; Jones, 2013; Baertschi, 2014; DeGrazia, 2014; Douglas, 2014b; Harris, 2014; Jebari, 2014; Jotterand, 2014; Morioka, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Savulescu et al., 2014; Sparrow, 2014a, 2014b; Agar, 2015b; Beck, 2015; Casal, 2015; Kahane & Savulescu, 2015)
	(Douglas, 2008; Ingmar Persson & Savulescu, 2008; Walker, 2009; Agar, 2010; Sprinkle, 2010; Bruni, 2011; Fröding, 2011; Jotterand, 2011; Schaefer, 2011; Lev, 2012; Ingmar Persson & Savulescu, 2012; Shook, 2012; Harris, 2013b; DeGrazia, 2014; Lechner, 2014; Macer, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Savulescu et al., 2014; Sparrow, 2014a, 2014b; Beck, 2015; Casal, 2015; Hauskeller, 2015; Kahane & Savulescu, 2015)
	(Douglas, 2008; Walker, 2009; Arnhart, 2010; Walker, 2010; Pacholczyk, 2011; Schaefer, 2011; Lev, 2012; Shook, 2012; Simkulet, 2012; Sio et al., 2012; DeGrazia, 2014; Jebari, 2014; Macer, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Ram-Tiktin, 2014; Savulescu et al., 2014; Sparrow, 2014b; Agar, 2015b; Hauskeller, 2015)

Table 3: Arguments for and against moral bioenhancement, presented in the reviewed literature (*continued*)

Cluster	Argument	Description/ background
	Situation- and role-dependency	Situation dependency of what counts as an improvement (morally). Different roles, assessments of situations. Weighing relevant reasons to act. One virtue can turn into a vice dependent on the situation.
	Human enhancement versus treating mental disorders	Enhancing humanity or treating mental disorders. Moral element in mental disorders.
3. The feasibility of moral bioenhancement and the status of current scientific research		
	Status of current scientific research	Further research is needed or, technological possibilities are already there.
	Complexity of our moral psychology/ biology	Makes it doubtful that we will gain sufficient understanding. Is morality genetically/ biologically determined? For example: are virtues and vices heritable? Is the core of our moral dispositions malleable by biomedical and genetic means? Danger of reductionism: we should not overlook the impact of the socio-cultural environment.
	Unintended or undesirable side effects	Interventions have effects beyond the intended effects (also: bluntness of the instruments). A 'baby and bathwater' problem. Moral bioenhancement might even lead to the opposite: not moral progress but moral decline.
	Scientific rigor, standards	Research ethical questions about standards of good/ sound science. Is scientific experimentation permissible, given that 'lack of moral virtue' is not a disease?
4. Means and process of arriving at moral improvement matter ethically		
	Other (non-biomedical) methods are preferable	Such as moral training, socialization or (self-) education. Taking a pill might seem 'all too easy' or too disconnected from ordinary human understanding. Are biomedical means intrinsically bad? Also: man is not supposed to play God.
	There is no principled difference between traditional and biomedical means	Results matter, the means less so. Perhaps the difference lies in the irrevocability/ irreversibility of biomedical means.
5. Arguments related to the freedom, identity, and autonomy of the individual		
	Moral bioenhancement might threaten the freedom of the individual	Moral bioenhancement might impair our freedom and diminish our freedom to act on bad motives. It might subvert moral agency.
	Moral bioenhancement might endanger our identity and autonomy	Questions about personal identity, and 'true' versus 'brute' self. Enhancer decides on outcome of moral bioenhancement (paternalism). Might compromise autonomous, informed choice.

	Key articles
	(Douglas, 2008; Faust, 2008; Sprinkle, 2010; Chan & Harris, 2011; Harris, 2011; Jotterand, 2011; Hallgren, 2012; Douglas, 2013; Agar, 2014; Jebari, 2014; Morioka, 2014; I. Persson & J. Savulescu, 2014; Sparrow, 2014a, 2014b; Wasserman, 2014b; Casal, 2015; Kahane & Savulescu, 2015)
	(Agar, 2010; Horstkötter, Berghmans, & de Wert, 2012; Agar, 2014; DeGrazia, 2014; Jotterand, 2014; Wiseman, 2014a; Agar, 2015b; Beck, 2015; Casal, 2015; Hauskeller, 2015)
	(Ingmar Persson & Savulescu, 2008; Walker, 2009; Arnhart, 2010; Walker, 2010; Pacholczyk, 2011; Shook, 2012; Harris, 2013b; Jones, 2013; Zarpentine, 2013; Agar, 2014; Crockett, 2014a; Jotterand, 2014; Savulescu et al., 2014; Casal, 2015; Ingmar Persson & Savulescu, 2015b)
	(Douglas, 2008; Ingmar Persson & Savulescu, 2008; Walker, 2009; Blackford, 2010; Harris & Chan, 2010; Sprinkle, 2010; Walker, 2010; Fröding, 2011; Jotterand, 2011; Pacholczyk, 2011; Ehni & Aurenque, 2012; Hallgren, 2012; William Paul Kabasenche, 2012; Shook, 2012; Harris, 2013b; Jones, 2013; Ingmar Persson & Savulescu, 2013; Zarpentine, 2013; Agar, 2014; Crockett, 2014a; Douglas, 2014b; Jotterand, 2014; Lechner, 2014; I. Persson & J. Savulescu, 2014; Robichaud, 2014; Sparrow, 2014a, 2014b; Wiseman, 2014a; Agar, 2015b; Beauchamp, 2015; Beck, 2015; J. A. Carter & Gordon, 2015)
	(Bronstein, 2010; Harris & Chan, 2010; Sprinkle, 2010; Walker, 2010; Chan & Harris, 2011; Harris, 2011; Pacholczyk, 2011; Horstkötter, Berghmans, & de Wert, 2012; William Paul Kabasenche, 2012; Lev, 2012; Tennison, 2012; Douglas, 2013; Jones, 2013; Zarpentine, 2013; Baertschi, 2014; Crockett, 2014a; DeGrazia, 2014; Harris, 2014; Jebari, 2014; Sparrow, 2014b; Agar, 2015b; Beauchamp, 2015; Beck, 2015; Casal, 2015; Kahane & Savulescu, 2015)
	(Bronstein, 2010; Shook, 2012; Tonkens, 2013)
	(Douglas, 2008; Faust, 2008; Walker, 2009; Harris & Chan, 2010; Sprinkle, 2010; Fröding, 2011; Harris, 2011; Christen & Narvaez, 2012; Ehni & Aurenque, 2012; William Paul Kabasenche, 2012; Douglas, 2013; Harris, 2013a; Hughes, 2013; Jones, 2013; William P. Kabasenche, 2013; Douglas, 2014b; Jebari, 2014; Casal, 2015)
	(Walker, 2009, 2010; Christen & Narvaez, 2012; Ingmar Persson & Savulescu, 2012; DeGrazia, 2014; Jebari, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Sparrow, 2014a)
	(Douglas, 2008; Faust, 2008; Walker, 2009; Chan & Harris, 2011; Harris, 2011; Jotterand, 2011; Pacholczyk, 2011; Schaefer, 2011; Curtis, 2012; Horstkötter, Berghmans, & de Wert, 2012; Ingmar Persson & Savulescu, 2012; Rakić, 2012; Simkulet, 2012; Douglas, 2013; Harris, 2013a; Harris, 2013a, 2013b; Ingmar Persson & Savulescu, 2013; DeGrazia, 2014; Harris, 2014; Jebari, 2014; Ingmar Persson & Julian Savulescu, 2014b; Rakić, 2014b; Savulescu et al., 2014; Sparrow, 2014a; Beck, 2015)
	(Douglas, 2008; Faust, 2008; Spence, 2008; Hubbeling, 2009; Agar, 2010; Blackford, 2010; Bronstein, 2010; Chan & Harris, 2011; Curtis, 2012; Ehni & Aurenque, 2012; Horstkötter, Berghmans, & de Wert, 2012; Ingmar Persson & Savulescu, 2012; Harris, 2013a; Savulescu et al., 2014; D. Shaw, 2014; Sparrow, 2014a; Beck, 2015)

Table 3: Arguments for and against moral bioenhancement, presented in the reviewed literature (*continued*)

Cluster	Argument	Description/ background
	Despite concerns about individual liberty and autonomy, a trade-off is justified	The advantages outweigh the disadvantages.
6. Arguments related to social/ group effects and dynamics		
	Moral bioenhancement benefits others	Unlike other types of enhancements (cognitive, cosmetic, sports). Or: who benefits? The individual or society as a whole?
	Moral bioenhancement might foster abuse	Moral bioenhancement might induce free-riding (e.g. prisoner's dilemma). The virtuous exposed to exploitation by the vicious. It may lead to moral decline.
	Moral bioenhancement might undermine moral diversity and moral debate	It might diminish opportunities for ethical thinking/ debate. Reasonable pluralism. Moral bioenhancement might generate social inequalities, elitism.
	Risks of utopian derailing	Progressive, well-intended, yet… Utopian. Interventions will be used recklessly or overenthusiastically. Moral perfectionism.
	Mandatory implementation or free/ parental choice	State neutrality versus free choice. Danger of tyranny/ discrimination.

studied, it will not be possible to take account of all the arguments in great detail. However, in the following paragraphs we hope to sketch the outlines of the discussions held so far and to provide an overview of the main arguments identified under clusters one through six. Relevant subthemes will be discussed under each cluster.

1. Why We (Don't) Need Moral Bioenhancement

The arguments gathered under this first cluster address the question as to why (or whether) we in fact need moral bioenhancement. What kinds of problems we hope it would eradicate, what its advantages are compared to other methods, and how it relates to traditional methods of moral betterment. It is clear that most proponents of moral bioenhancement feel the need to offer some story on why there is in fact an urgent need for it. Opponents or sceptics may doubt whether we need moral bioenhancement at all.

	Key articles
	(Douglas, 2008; Pacholczyk, 2011; Ingmar Persson & Savulescu, 2012; Douglas, 2013; DeGrazia, 2014; Harris, 2014; Ingmar Persson & Julian Savulescu, 2014b; I. Persson & J. Savulescu, 2014; Rakić, 2014b; Savulescu et al., 2014; Selgelid, 2014)
	(Douglas, 2008; Bronstein, 2010; Baertschi, 2014; Jebari, 2014; Marshall, 2014; Sparrow, 2014b)
	(Douglas, 2008; Faust, 2008; Walker, 2009; Pacholczyk, 2011; Ehni & Aurenque, 2012; Shook, 2012; Douglas, 2013; Harris, 2013b; Triviño, 2013; Jotterand, 2014; Morioka, 2014; Ram-Tiktin, 2014; Robichaud, 2014; Sparrow, 2014a, 2014b; Wilson, 2014; Wiseman, 2014a; Beauchamp, 2015; Beck, 2015)
	(Agar, 2010; Brooks, 2012; Shook, 2012; Sio et al., 2012; Harris, 2013b; Jotterand, 2014; Lechner, 2014; Marshall, 2014; Ingmar Persson & Julian Savulescu, 2014a; I. Persson & J. Savulescu, 2014; Ram-Tiktin, 2014; Robichaud, 2014; Sparrow, 2014b; Wilson, 2014; Wiseman, 2014a; Casal, 2015)
	(Faust, 2008; Bronstein, 2010; Sprinkle, 2010; Walker, 2010; Ehni & Aurenque, 2012; Douglas, 2013; Harris, 2013b; Jones, 2013; Joyce, 2013; Sparrow, 2014a, 2014b)
	(Faust, 2008; Ingmar Persson & Savulescu, 2008; Walker, 2009; Arnhart, 2010; Bronstein, 2010; Curtis, 2012; Lev, 2012; Ingmar Persson & Savulescu, 2012; Rakić, 2012; Shook, 2012; Sio et al., 2012; Joyce, 2013; Triviño, 2013; Lechner, 2014; Macer, 2014; Morioka, 2014; Ingmar Persson & Julian Savulescu, 2014b, 2014a; I. Persson & J. Savulescu, 2014; Rakić, 2014b, 2014a; Selgelid, 2014; Sparrow, 2014b; Wasserman, 2014b; Wiseman, 2014a)

There is scope for improvement

Almost by definition, most if not all people would benefit from an improvement in their moral character. Different authors vary, however, with respect to the kind of changes they would like to see implemented: changes in moral behavior, will-power, or moral agency and insight.

Because the moral character of most people is suboptimal (or even defective by nature), every person has good reasons to morally better herself. The general argument holds that we have a moral duty to enhance ourselves, and that if we need moral bioenhancement to reach this goal, we should consider it: "it is not that taking medicine is intrinsically moral or immoral, it is that a human subject can use medication as a means to assist them towards a moral end: reducing future harm. Such a person exhibits altruism" (Spence, 2008, p. 180). The only right attitudes towards one's own bad motives and impediments are "*non-acceptance* and a *desire for self-change*", Thomas Douglas (Douglas, 2008, p. 235) maintains.

Where Douglas (Douglas, 2008) presents the recognition that there is room for improvement as an argument for moral bioenhancement, others are of the opinion that although we can agree that the world we live in now is far from optimal, it is not clear why this would be a reason in favor of moral bioenhancement. For example, according to Nicholas Agar: "We don't need superior moral vision to understand that poverty, climate change, and terrorism are bad things. (…) We do need enhanced effort and perhaps enhanced nonmoral powers to fix poverty, climate change, and terrorism but we don't need enhanced moral vision to recognize that they need fixing" (Agar, 2010, p. 75).

Human biological nature is defective

In defense of the need for moral bioenhancement to morally better ourselves, Ingmar Persson and Julian Savulescu argue that there is a fundamental mismatch between our moral psychology and today's conditions of human life (Ingmar Persson & Savulescu, 2013, p. 124). Because human moral psychology evolved in conditions that are radically different from those in today's world, we should alter human moral psychology by biomedical and genetic means, they argue:

> People encode the race of each individual they encounter, and do so via computational processes that appear to be both automatic and mandatory. (…) If genetic and biomedical means of enhancement could counter such natural tendencies, they could have a crucial role to play in improving our moral character, that could complement traditional social and educational means of moral enhancement. (Ingmar Persson & Savulescu, 2008, p. 168)

Others add that besides being ill equipped for today's conditions of human life, human beings are innately evil to a greater or lesser extent. Wickedness is an indispensable part of human nature. If we want to eradicate evil, we have to alter these immoral innate (biologically determined) tendencies of human beings. Socialization, upbringing and education will bring us only so far. According to Mark Walker, precisely because humans are evil by nature, we need biomedical interventions in order to effectively alter human nature for the better: "For sure, it may be possible to minimize some contemporary evil through better socialization, but it will never be possible to eliminate it so long as human nature remains unaltered" (Walker, 2009, p. 29).

On the other hand, Robert Sprinkle argues that the observation that evil may not be an eliminable feature of the human condition should temper our hopes regarding the possibility of effectively addressing all forms of evil, not raise them: "I, for one, never held such a hope" (Sprinkle, 2010, p. 89). John Harris turns the 'humans are evil by

(biological) nature'-argument around and argues that there is an inborn human goodness: "We have certainly evolved to have a vigorous sense of justice and right, that is, with a virtuous sense of morality" (Harris, 2011, p. 104).

Traditional means are (not) effective enough

In addition to stating the need for moral bioenhancement due to our defective moral nature, some authors argue that traditional means are ill-equipped or less effective as compared to biological and/or genetic means. As mentioned by David DeGrazia, surely, we already have at our disposal many different means of enhancing our moral capacities: methods such as "explicit moral instruction, mentoring, socialization, carefully designed public policies, consciousness-raising groups, literature and other media that encourage moral reflection, and individual efforts at improvement" (DeGrazia, 2014, p. 361). However, Persson and Savulescu argue that these means are not nearly effective enough to help us counter the great evils of our time: "Biomedical and genetic means may be much more effective in terms of both how thoroughly and quickly they could improve everyone in need of improvement" (Ingmar Persson & Savulescu, 2008, p. 168).

Others, for example John Harris and Jamie Bronstein, feel that this line of argument wrongly minimizes the moral progress that has been made through those tried and tested traditional methods (Bronstein, 2010, p. 86; Harris, 2011, p. 104), and argue that these methods still offer many possibilities for moral improvement.

Our only hope in averting major disaster

On top of the need for effective interventions to morally better ourselves, *urgency* is another critical factor that is addressed in the debate. If we succeed in (biomedically) enhancing people's cognitive abilities, some argue, it is of paramount importance to also – or even first – enhance their moral abilities due to the risks that cognitively enhanced human beings may pose to others. In today's technologically advanced world, Persson and Savulescu argue that a "morally corrupt minority" (Ingmar Persson & Savulescu, 2008, p. 163) is increasingly able to inflict major disaster on the majority. Moral bioenhancement might be our only hope of engaging with other major challenges as well. According to DeGrazia:

> The status quo is deeply problematic because there is such an abundance of immoral behavior, with devastating consequences, and serious risk of worse to come. (…) In addition to these harms and injustices, there is the threat of truly massive harm. (…) It is increasingly possible for a small number of individuals to acquire the technical capability of inflicting terrible harm. (DeGrazia, 2014, p. 362)

Harris, however, argues that we should instead embrace cognitive enhancements, as they are our best prospect of self-defense against disaster (Harris, 2011, p. 110). Adam Carter and Emma Gordon argue that because cognitive and moral enhancements are principally interconnected, we should consider potential enhancements "*outwith* any essential reference to a moral/cognitive conceptual dichotomy" (J. A. Carter & Gordon, 2015, p. 160).

David Wasserman questions whether all of these evils can be attributed to individual moral defects, and warns us not to underestimate the role of defective institutions (Wasserman, 2014b, p. 375).

Moral bioenhancement might reduce criminality

Last but not least, some authors such as Walker (Walker, 2009) suggest that moral bioenhancement could achieve a significant reduction in 'evil', referring to criminal behavior such as rape, murder, torture and so on. Others, for example Thom Brooks, warn us that immorality and illegal behavior do not necessarily coincide:

> Morality and law are imperfectly linked at best. First, not all immorality is illegal. Lying is widely regarded as immoral, but not all lying is criminal. (…) Second, not all illegality is immoral. Drug offenses are widely incorporated in most legal systems, but it is unclear at best whether cannabis use is intrinsically immoral. (Brooks, 2012, p. 29)

Because the project of moral bioenhancement and the project of reducing criminality are not necessarily the same, it is argued that we should be careful in suggesting that moral bioenhancement might indeed reduce criminality and using this as an argument for moral bioenhancement.

2. It Will (Not) Be Possible to Reach Consensus On What Moral Bioenhancement Should Purport

In this second cluster we discuss arguments that address the many different kinds of disagreement that influence the different proposals on the way moral bioenhancement should take shape. Opponents of moral bioenhancement think that because of these allegedly fundamental disagreements, moral bioenhancement is a problematic endeavor. Proponents, however, argue that sufficient consensus is possible, and that these differences need not necessarily jeopardize the project of moral bioenhancement.

No consensus on the mechanisms that comprise our moral psychology

The first issue that fuels disagreement on the way moral bioenhancement should take place relates to our limited knowledge concerning our moral psychology. Moreover, according to Persson and Savulescu, how we should interpret and understand findings of moral psychology is not straightforward, and influenced by our preferred view on what constitutes morality: "what morality is, or of what it is to be moral" (Ingmar Persson & Savulescu, 2008, p. 168). Agar warns us that:

> The absence of a consensus upon the mechanisms of morality could prevent any agreement that a proposed moral enhancer could really be enhancing morality, whatever else it may be doing. This skepticism is not the fault of the behavioral and brains sciences, but our own, for failing to agree about which cognitive processes are genuinely relevant to what we want to call morality and moral agency. (Shook, 2012, p. 5)

Behavioral changes alone are (not) enough

Even if a consensus would exist on the mechanisms of morality and how to achieve more moral behavior, several authors question whether it is enough for any moral bioenhancement intervention to have effect on behavior, but not necessarily on other aspects of morality, such as moral reasoning, moral insight, or moral will. Do behavior control interventions constitute moral enhancement or does moral enhancement require an accompanying change in moral agency?

According to Douglas (Douglas, 2008, 2013, 2014b), reducing an individual's tendencies towards violent aggression *directly*, without using cognitive means such as persuasion or deliberation (and assuming this would effectively lead to less immoral behavior), would count as moral enhancement. Harris (Harris, 2011, 2013a, 2013b; Harris, 2013b; Harris, 2014), however, insists that without concurrent changes in a person's moral reasoning, these changes would not amount to moral enhancement at all:

> We will, I believe, always need to use moral reasoning to act as a guide to our emotions and as a way of checking that we are having appropriate feelings in appropriate circumstances and for appropriate objects. If the good involves feeling the right way, how do we know that we are feeling the right way? (Harris, 2013a, p. 172)

Fabrice Jotterand also argues that "the emphasis on the control of moral emotions appears reductive and one-sided in the sense that it conflates moral reasoning (as practical reasoning) with moral psychology (how moral reasoning acts on one's motivational/emotional states)" (Jotterand, 2011, p. 5).

Bernard Baertschi argues that much of the disagreement between Douglas and Harris can be attributed to their different preferred meta-ethical positions (Baertschi, 2014, pp. 66-67). Harris adheres to a rationalist conception of ethics, according to which emotions should only be acted upon through cognitive means, and, as described by Baertschi, "reason furnishes the only genuine moral motives" (Baertschi, 2014, p. 66). Douglas however appears to be a sentimentalist, Baertschi argues, i.e. espousing a view on ethics according to which having the right feeling matters. Their different conceptions of what morality *is* influence their different assessments of whether moral bioenhancement might be effective. Whereas Douglas thinks that direct modulation of emotions is effective (and permissible), Harris denies that direct modulation of emotions even amounts to moral enhancement at all.

Ethical systems and theories differ and often disagree

The wide variety of substantive moral principles that characterizes debates between ethical theorists, also hampers agreement on what would constitute a moral enhancement. John Shook provides the following example:

> Suppose a brain modification transforms a person into someone who now takes the moral deed to be the one maximizing the welfare of all. (…) most utilitarians would soon find fault with this new utilitarian's concrete moral judgments, just as they find fault with each other's. And many deontologists would simply deny that this fresh utilitarian has received moral enhancement at all. (Shook, 2012, p. 4)

Nevertheless, as DeGrazia argues, in one way or another, every program for moral bioenhancement needs to make explicit what it considers to be a moral improvement, based on what kind of principles or theory. Differences between various moral theories are not necessarily purely theoretical, and may lead to different normative judgments concerning a variety of moral dilemmas (e.g. abortion, the death penalty, and euthanasia) (DeGrazia, 2014, p. 363).

Next to disagreement between the major ethical theories, DeGrazia reminds us that individual (groups of) people hold diverse and often conflicting moral outlooks. They differ greatly with respect to the values they adhere to: politically conservative or progressive values for example (DeGrazia, 2014, p. 363).

(Im)possibility of considerable consensus

In addition to the fact that a consensus is currently lacking, some authors view ethical standards as "arbitrary products of cultural history", as Larry Arnhart wonders with regard to Walker's writings (Arnhart, 2010, p. 81). Or, given a pluralistic reality, is it

possible to be neutral with respect to wide-ranging outlooks and ethical systems? Persson and Savulescu, for instance, expect that their proposal for the core of our moral disposition – consisting of altruism, a sense of justice or fairness, and empathy – will be shared by many (Ingmar Persson & Savulescu, 2008, pp. 168-169). They think that despite the deep disagreements between different accounts of right action, some sort of actions ("the willingness to sacrifice one's own interests for the benefit of others", for example) will be viewed as "a moral enhancement, on any account of morality" (Ingmar Persson & Savulescu, 2012, pp. 5-6).

Walker also cautions against overemphasizing the differences, and points to significant overlap between different lists of virtues (Walker, 2009, p. 35). DeGrazia proposes that we "stick to improvements that represent points of overlapping consensus among competing, reasonable moral perspectives" (DeGrazia, 2014, p. 364). Moreover, Filippo Sio and colleagues claim that:

> As in the cases of cognitive enhancement and social progress, reference to some objective standards is also necessary to make the concept of moral enhancement coherent. (Sio et al., 2012, p. 15)

Situation- and role-dependency

In addition to different views on morality and moral behavior, situation- and role-dependency further add to the confusion concerning what should count as a moral enhancement. What should count as an improvement is highly dependent on the specific context and the roles performed in that situation (e.g. detached surgeons to remove brain tumors, impartial judges to administer justice). Wasserman asserts that even slight moral improvements will vary according to the role and context in which these are brought about (Wasserman, 2014b).

Sprinkle approvingly cites Aristotle's assertion that "traits virtuous in moderation might be vices in absence or excess" (Sprinkle, 2010, p. 89). Moreover, Markus Christen and Darcia Narvaez argue that "moral character cannot emerge from a short-term intervention, but, as Aristotle advised, must be shaped with mentoring through multiple situations over time" (Christen & Narvaez, 2012, p. 26). Moreover, Sarah Chan and John Harris refer to a situation in which "serotonin-induced aversion to inflicting direct harm" might have stopped passengers from forcefully stopping a would-be hijacker (Chan & Harris, 2011, p. 131). In situations like these, aggression can be a good thing, although it is clearly bad in others. From this it follows that even enhancing traits that everyone would agree to be 'good', may still not result in an overall, all-purpose, moral enhancement. However, such criticisms on moral bio-enhancement fail to consider, ac-

cording to Kahane and Savulescu, that this is "not at all an argument against enhancement but, rather, an argument for more precisely *fine-tuned* enhancement" (Kahane & Savulescu, 2015, p. 141). For example, by making biomedical interventions sensitive to certain contexts, but not to others.

Human enhancement versus treating mental disorders

Some authors, such as Agar and Jotterand, highlight the difference between enhancing moral capacities of individuals "beyond human norms" (Agar, 2014, p. 369) and treating mental disorders that may or may not contain "an inherent moral element" (Jotterand, 2014, p. 1). Dorothee Horstkötter and colleagues for example argue that "if there is a health problem, medical treatment is the reasonable reaction, while enhancement, either moral or otherwise, does not arise" (Horstkötter, Berghmans, & de Wert, 2012, p. 27).

3. The Feasibility of Moral Bioenhancement and the Status of Current Scientific Research

Under the third cluster, we discuss arguments based on the alleged feasibility of proposals for moral bioenhancement. Whereas proponents of moral bioenhancement are optimistic about the status of current scientific research, opponents warn us that the complexity of our moral psychology and biology make it doubtful that we will be able to develop effective interventions.

Status of current scientific research

Current scientific developments give rise to both high hopes and substantial skepticism. DeGrazia, for example, rather optimistically lists research that may further the science of moral bioenhancement, ranging from the use of selective serotonin reuptake inhibitors as a means to being less inclined to assault people, to deep brain stimulation as a way to reduce aggression (DeGrazia, 2014, pp. 361-362). Walker sees no technological reasons why the pre-implantation sorting of embryos that is presently used to screen for genetic diseases could not be used for selecting virtues (Walker, 2009, p. 31).

Molly Crockett warns us against overstating the conclusions of single studies, and asserts that science in this field is "in its infancy" (Crockett, 2014a, p. 370). With respect to genetic engineering Arnhart observes that clear examples of "specific genetic linkages to virtue that could be altered" are presently unknown (Arnhart, 2010, p. 79).

Persson and Savulescu admit that their proposals for moral bioenhancement are mostly based on hypothetical scenarios, and treatment at this moment is only possible to a very small extent: "A lot more scientific research is needed before we can be made

more altruistic or just by suitable drugs or surgery, or genetic manipulation" (Ingmar Persson & Savulescu, 2008, p. 172).

Although William Kabasenche thinks that human moral psychology is highly complex and therefore that it is difficult to 'engineer' virtues and vices, he feels that these studies "give us important insights into what embodied virtues might look like, and they suggest that we should take our embodied nature seriously" (William Paul Kabasenche, 2012, p. 20).

Complexity of our moral psychology/ biology

Regardless of an optimistic or pessimistic view on the status of current scientific research, it is also argued that character traits, such as those involved in human morality, are highly complex, and therefore that moral bioenhancement is probably not feasible. These reservations are not only expressed with respect to the manipulation of neurotransmitters, but also with respect to the possibilities of genetically engineering virtues.

Arnhart doubts that we are anywhere near having found the "generic correlates of virtue that are clear, strong, and manipulable" (Arnhart, 2010, p. 80). Walker however thinks that, given the fact that much progress has been made in the behavioral genetics of schizophrenia, this field can show the way forward to investigating the possibilities of a comparable behavioral genetics of virtue: "After all, genes for schizophrenia are polygenetic and show intensity of expression and gene-environment interactions" (Walker, 2010, pp. 91-92). Douglas also feels that there are some elements of our moral psychology that we are beginning to understand to such a degree that manipulating them is possible: "it does not seem unreasonable to suppose that moral enhancement technologies which operate on relatively simple emotional drives could be developed in the medium term" (Douglas, 2008, p. 233).

Some authors, for example Robert Sparrow, voice concerns about reductionism, i.e. "the claim that whether an individual is a (morally) good person is a function of that person's neurochemistry and/or that person's genetics" (Sparrow, 2014b, p. 27) . Others, such as Hans-Joerg Ehni, Diana Aurenque (Ehni & Aurenque, 2012) and Chris Zarpentine (Zarpentine, 2013), warn us not to underestimate the importance of societal and cultural influences.

Unintended or undesirable side effects

Given the complexity of our moral psychology and biology, can we hope to influence it without also altering other crucial processes? Crockett (Crockett, 2014a) warns against

the unintended, and possibly undesirable, consequences of altering the function of a specific neurotransmitter, beyond the desired effects on moral behavior. Karim Jebari, for example, discusses findings that suggest that enhancing empathy may render individuals less fair and more partial rather than less fair and more impartial (Jebari, 2014). Agar provides the following example:

> What we recognize as the correct pattern of judgment strikes a particular balance between the call of empathy and the appeal of moral reasoning. (…) Unbalanced enhancement of empathy is likely to disrupt what we view as the morally correct trade-off between benefits conferred on those to whom we are socially bonded and costs experienced by those to whom we are not socially bonded. It tends to reinforce our tendency to endorse solutions that inflict suffering on strangers to protect our nearest and dearest from less significant suffering. (Agar, 2015b, p. 344)

In the case of genetic engineering, many more systems than just the targeted virtue or vice might be effected. For example, Bronstein asks: "What happens if selecting for virtuous genes also increases the likelihood of cancer, diabetes, heart disease, or even shyness or depression?" (Bronstein, 2010, p. 85).

Scientific rigor, standards

In addition to the scientific and philosophical uncertainties regarding our moral psychology and biology, some authors touch upon the issue of future scientific experimentation with respect to moral bioenhancement. Bronstein asks whether medical experimentation is permissible, given the fact that immorality is not a disease: "Walker's project design may also violate one of the great principles of human experimentation: that medical experimentation approved by proxy on behalf of those who cannot consent must benefit the patient. Lack of moral virtue is perhaps suboptimal, but we have not yet classified it as a disease" (Bronstein, 2010, p. 85). Bronstein further observes that, in the case of Walker's Genetic Virtue Program, many questions can be asked with respect to the research design:

> Will genetically modified humans be raised in controlled environments so that they can be more easily observed and exposed to uniform socialization? If not, how can we know whether these particular humans are indeed more virtuous, and that their virtue is indeed genetic? (Bronstein, 2010, p. 86)

4. Means and Processes of Arriving at Moral Improvement Matter Ethically

Under the fourth cluster, we discuss arguments that explore the question as to whether the difference between biomedical and non-biomedical means matters ethically.

Other (non-biomedical) means are preferable

From an intrinsic perspective, it is often considered whether it would be better, more praiseworthy, or more authentic, if a person betters herself without resorting to biomedical means? Or as Douglas puts it: is it the case that adopting "biomedical means to moral enhancement is objectionable not just relative to other alternative means, but in an absolute sense" (Douglas, 2008, p. 236)? Sprinkle thinks that even if a safe and "convincingly enduring" intervention would become available, people would likely prefer a non-genetic remedy (Sprinkle, 2010, p. 89). This argument implies that we have intrinsic reasons, such as authenticity and personal identity, to reject biomedical interventions in favor of traditional means for moral enhancement. It also assumes that there is a principled difference between biomedical and traditional, non-biomedical means.

Others argue that biomedical and non-biomedical means will be used in concert, rather than separately. According to Douglas (Douglas, 2008), we will likely regard biomedical enhancement and self-education as complementary and are likely to reinforce the desire for both by initially engaging in one or the other. Moreover, according to Kabasenche, "none of us achieve any measure of success in moral formation without significant assistance from others. If authentic moral formation is something you do completely by yourself, none of us has done it" (William Paul Kabasenche, 2012, p. 20).

There is no principled difference between traditional and biomedical means

Alternatively, some authors argue that there is no principled difference between using traditional and biomedical means to morally better oneself or others. DeGrazia, for example, argues that (many of the) arguments against biomedical means also apply to traditional, non-biomedical means: "one should not inculcate moral values that are wrong, so how can a parent be sure that she or he is justified in providing a particular type of moral instruction? Also facing this challenge are public school teachers who attempt to inculcate in students certain moral virtues such as civility, respect for differences and concern for the poor" (Douglas, 2014b, p. 363).

Likewise, according to Walker's 'companions in innocence' line of reasoning, any principled argument given against biomedical means for moral enhancement, such as those involved in his Genetic Virtue Program, equally applies to socialization and education efforts: "If the [Genetic Virtue Program] is wrong in attempting to promote virtue as a means of making people morally better, then much current socialization and education is mistaken as well" (Walker, 2009, pp. 35-36).

Sparrow however maintains that a significant disanalogy exists between traditional means of moral improvement and the biological manipulation of behavior and motivation. Whereas education is characterized by "a fundamental moral equality between educator and educated", biomedical interventions to reshape the agency of others "involve a subject acting towards an object and as such are fundamentally structured by a profound inequality" (Sparrow, 2014a, p. 26).

5. Arguments Related to the Freedom, Identity, and Autonomy of the Individual

Under cluster five, we discuss various arguments regarding the question as to whether moral bioenhancement would limit the individual in his or her opportunities to freely choose his or her behavior? We focus on concerns related to individual liberty and autonomy.

Moral bioenhancement might threaten the freedom of the individual

The concept of freedom takes a central role in the moral enhancement debate, with Harris being one of its most ardent defenders. He is of the opinion that individual liberty is of utmost importance, and should take priority over all other good ends that we might pursue. He explicitly opposes "any measures that make the freedom to do immoral things impossible, rather than simply making the doing of them wrong and giving us moral, legal and prudential reasons to refrain" (Harris, 2011, p. 105). William Simkulet argues that "when one is forced against one's will to do as the virtuous person in one's place would freely do" we should speak of moral compulsion instead of moral enhancement (Simkulet, 2012, p. 17).

Some authors, such as Jebari and Birgit Beck, call for conceptual clarification of a suitable concept of freedom in the debate on moral enhancement: what kind of freedom is at stake? (Jebari, 2014; Beck, 2015)

Moral bioenhancement might endanger our identity and autonomy

Related to the worries concerning individual liberty, concerns are voiced that moral bioenhancement could pose a threat to our true, autonomous self. Douglas argues that the counter-moral emotions that would be altered are at best part of a person's 'brute' self, and thus enhancement would be "allowing his true self *greater* freedom" (Douglas, 2008, p. 240). Our true or authentic self refers to our internal characteristics whereas our brute self refers to anything that is external to this (Douglas, 2008, p. 240). Certain enhancements could alter our brute self in such a way that it constrains our true self, thereby threatening our freedom and autonomy. At the same time, others argue that we should not overestimate an individual's capacity for full autonomous behavior as

exemplified by his/her 'true self'. For example, Russell Blackford warns against attributing to ourselves "a "spooky kind of autonomy all the way down" that does not exist in the real world" (Blackford, 2010, p. 83).

Some authors also worry about possible changes of identity as a result of moral bioenhancement. Douglas distinguishes a loss of identity in a strong sense, in which an individual would, post-enhancement, be a completely different person, and a weak sense, in that moral bioenhancement would change some of her most fundamental psychological characteristics (Douglas, 2008, p. 239). Yet Douglas stresses that we only have reason to preserve those psychological characteristics that have positive value for the individual in question. Whereas Douglas emphasizes that the individual is free to choose, others such as Agar (Agar, 2010) are less clear on how the positive value of particular psychological characteristics is to be determined: on the basis of an individual's own judgment, on someone else's judgment (e.g. within a criminal justice contexts), or on a specific moral theory?

Given the large diversity of potential moral enhancements, moral enhancement interventions will inevitably prioritize some moral values or character traits over others. This, some argue, will place the person who undergoes the intervention at the mercy of the person performing the intervention: "Someone who has been subjected to moral enhancement is likely to have a reduced sensitivity to moral reasons rejected by his or her enhancer" (Agar, 2010, p. 75).

Despite concerns about individual liberty and autonomy, a trade-off is justified

Acknowledging that moral bioenhancement might indeed negatively impact the freedom, autonomy, or identity of an individual, should this stop us from pursuing moral bioenhancement? Some authors, for example Douglas and DeGrazia, stress that although such a loss might be regrettable, if it is compensated by an increase of some other good, the loss can be justified (Douglas, 2013; DeGrazia, 2014). DeGrazia, for example, maintains that "we should not exaggerate the value of freedom. After all, moral behavior itself, the end product, is also extremely important—independently of how free it is" (DeGrazia, 2014, p. 367; see also Ingmar Persson & Julian Savulescu, 2014b, p. 252). Savulescu and Persson claim that:

> We are not free to commit serious crime even now – the laws prohibits it on pain of punishment. What we weren't free to do, the God Machine makes strictly impossible. If this is a loss, it would be outweighed by the fact that there are no victims suffering from serious crimes. (Savulescu & Persson, 2012, p. 13)

Bronstein is not convinced however that this trade-off is justified: "one might ask whether the goal of moral perfection is worth the trade-off for human autonomy on a large scale" (Bronstein, 2010, p. 86).

6. Arguments Related to Social/ Group Effects and Dynamics

Finally, under the sixth cluster, we discuss arguments that consider possible societal and group effects of moral bioenhancement. Would moral bioenhancement foster abuse? Should its use be mandatory, and who should decide? Is there overconfidence in the possibilities of biomedical solutions?

Moral bioenhancement benefits others

To start, Douglas situates questions about the desirability of moral bioenhancement in the context of the wider enhancement debate, and argues that the fact that (unlike other enhancements) moral enhancement primarily benefits others, neutralizes many of the objections often raised in the broader enhancement debate: "moral enhancement[s] could not easily be criticized on the ground that their use by some would disadvantage others" (Douglas, 2008, p. 230).

Bronstein however turns this argument around, and criticizes Walker's Genetic Virtue Project precisely because it appears to prioritize the benefits to others over the benefits to the agent (Bronstein, 2010, p. 85). He worries that in the design of the project, the interests of society will take priority over the interests of the individual.

Walker (Walker, 2009) raises the worry that the biologically unenhanced might be discriminated against in favor of the biologically enhanced, but immediately adds that it is not clear whether the same incentive for discrimination would arise in cases of moral enhancement as can be expected in cases of physical and/or cognitive enhancement.

Moral bioenhancement might foster abuse

Another concern that is put forward, is the fear that an altered ratio of moral people to immoral people might give rise to free-riding: the few morally unenhanced might more easily take advantage of the good intentions of the many morally enhanced. This dynamic might be visible not only between groups of individuals, but between countries as well, Shook suggests: "Depictions of entire societies or a whole planet undergoing empathetic moral enhancement will remain utopian fantasies. One country after another will decline moral enhancement until the "worse" countries have done it, and each country would want their neighbors to go first" (Shook, 2012, p. 11).

Walker however thinks that this worry underestimates the resilience of the morally enhanced: "it seems to assume that the virtuous are meek or compliant" (Walker, 2009, pp. 42-43).

Moral bioenhancement might undermine moral diversity and moral debate

Moral enhancement also raises questions with respect to which moral views or paradigms may benefit at the expense of others, and whether this may lead to a diminished diversity of views. Brooks argues that:

> The question is not only whether moral enhancement might lead to only one moral judgment, but also whether moral enhancement might benefit some reasonable moral, philosophical, and religious doctrines over others. If not all reasonable doctrines may benefit equally, then moral enhancement might violate the equality between citizens and fail to respect the reasonable pluralism that exists. (Brooks, 2012, p. 29)

This may hamper the quality of political and social debate: Shook worries that moral enhancers "could diminish opportunity, capacity, and responsibility for serious ethical thinking" (Shook, 2012, p. 8). Moreover, in a future dystopia, moral enhancements may be regarded as suitable quick fixes in case of moral ambiguities and dilemmas, thereby further reducing valuable opportunities for serious ethical reflection: "individuals thinking too hard about moral ambiguities and dilemmas are told that they simply need their enhancers adjusted" (Shook, 2012, p. 8).

Sparrow worries that the morally enhanced would gain important advantages, for example by their exclusive participation in social and political institutions, thereby generating or intensifying social inequalities (Sparrow, 2014b).

Risks of utopian derailing

Related to the risk of disrespect for reasonable pluralism, some authors express the worry that the promises and high hopes of moral bioenhancement projects (for example Walker's (Walker, 2009) proposal for a Genetic Virtue Project) will repeat many of the mistakes (such as mass regimentation and loss of autonomy) of what Bronstein calls 'High Modernist planning': "Walker's plan features all of the confidence, and many of the other signs, of High Modernist planning. The project that he proposes is transnational in scope; it seeks to transcend, or one might even say ignore, current political realities. Its emphasis is on the future, and it is enlivened by the discourse of the good" (Bronstein, 2010, p. 86).

Sprinkle recalls the excesses of the eugenics movement: "Walker's work should not be turned wrong-side-out. But neither can the lessons of the Eugenics Movement be taken as safely learned long ago. It was avowed to be a progressive movement, a product of progressive thought and an instrument for progressive action" (Sprinkle, 2010, p. 89).

Mandatory implementation or free/ parental choice

Despite the many reservations described above, if we assume that safe and effective moral bioenhancement would become available, would it be justifiable to make its use mandatory – for all, or for specific target groups? First, Bronstein raises doubts that many people will voluntarily seek moral bioenhancement: "how large is the distance between explaining what is good for you and imposing what we know to be good for you? (…) Genetic virtue is an idea that ought not to be imposed on an unwilling public—and seems unlikely to find a willing public" (Bronstein, 2010, p. 86).

Persson and Savulescu however argue that an unwilling public should not stop the program, and that safe, effective enhancements should be compulsory: "If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them" (Ingmar Persson & Savulescu, 2008, p. 174). Rakić (Rakić, 2014b, 2014a) however argues that making moral bioenhancement obligatory would deprive us of an essential part of our human existence, that is, the freedom "for us acting *intentionally* in a morally appropriate manner" (Rakić, 2014b, p. 248), and therefore advocates voluntary moral enhancement.

A specific version of this question arises with respect to children. Should parents be the ones to decide whether their children should undergo an intervention, or whether their future offspring should be genetically selected within the framework of Walker's (Walker, 2009, 2010) Genetic Virtue Project? Walker himself favors "some hybrid or conditional option to mediate between the state-mandated versus liberal (parental choice) implementation" in which parents are free to choose which virtuous characteristics they would like to see enhanced, but in which they are not free to choose the associated vices (Walker, 2009, p. 43). Arnhart wonders who will be responsible for setting and enforcing the standards for these virtues and vices, and fears the "threat of tyranny – either the tyranny of a few or the tyranny of the majority" (Arnhart, 2010, pp. 80-81).

Critical appraisal of the current debate

The debate on moral bioenhancement is a fairly recent phenomenon. Douglas, in the first article on the subject in 2008, discusses moral bioenhancement as a theoretical possibility in the context of the discussion on the permissibility of enhancement in general. Persson and Savulescu first discuss the possibility of moral bioenhancement in the context of the rapid developments in the field of cognitive enhancement (Ingmar Persson & Savulescu, 2008). Although they are doubtful as to whether moral bioenhancement is in fact feasible in the foreseeable future, they nevertheless call for intensified research efforts because they see moral bioenhancement as the only possible solution for a number of pressing problems. Later on, the debate moved on to fundamental philosophical questions about what human morality involves, and whether or not we would be able to reach enough consensus to transcend our current pluralistic moral reality. In what follows, we critically assess the arguments discussed under the six clusters presented above, and identify those issues and concerns that have been neglected so far. We identified four topics of concern: (1) the distinction between treatment and enhancement; (2) an overestimation of the possibilities/feasibility of moral bioenhancement; (3) insufficient attention to side-effects, risks and safety; and (4) identity changes.

The Distinction Between Treatment and Enhancement

There has been surprisingly little discussion concerning the criteria we should use to identify specific interventions as moral enhancement rather than as therapy or moral education. In the debate so far, it remains unclear whether 'moral enhancement' should be taken to include treating those with a pathological lack of certain moral capacities. Horskötter and colleagues argue that those interventions aimed at restoring normal moral functioning in subjects whose moral functioning is somehow pathologically impaired, should be called medical treatment, rather than enhancement (Horstkötter, Berghmans, & de Wert, 2012, p. 27). Agar is one of the few authors in the debate who do make this distinction. He claims that targeted interventions aimed at "therapeutic ends" can possibly amend specific deficiencies, but that these same interventions "can produce unbalanced excesses when used to enhance beyond human norms" (Agar, 2014, p. 369). Other authors, for example Douglas and DeGrazia, appear to use examples of general moral enhancement and specific mental pathologies, such as psychopathy and antisocial traits (e.g. violent aggression), interchangeably (Douglas, 2008; DeGrazia, 2014).

Although some, for example DeGrazia (2014), reject the distinction between treatment and enhancement, and while it is clear that the exact boundaries will of course

always be up for dispute, we believe that distinguishing between treatment of those with pathologies and enhancement of normal people will greatly benefit the debate. Shifting the focus to treatment of pathological deficiencies in morality raises new and interesting questions that are different from the ones raised in the debate on moral enhancement for normal people. What is to be considered normal moral functioning, and who is to determine whether a subject functions 'normally'? Should subjects who lack certain capacities or who show 'abnormal' moral functioning be considered to have a disease or disorder? In other words: when should diminished moral functioning or immoral behavior be considered to be pathological? How should society deal with common moral deficiencies? As mentioned above, Persson and Savulescu argue that safe and effective moral enhancement should be compulsory since those individuals that need them will be least willing and/or likely to use them (Ingmar Persson & Savulescu, 2008). Such safe and effective moral bio-enhancements are not available at present and difficult questions remain to be answered. This brings us to our final remark regarding the distinction between treatment and enhancement: more debate is urgently needed on what would be the right kind of response towards those who behave immorally due to pathological deficiencies of their moral capacities: treatment or punishment?

Overestimation of the Possibilities/ Feasibility of moral bioenhancement

Although some authors, such as Bronstein and Sprinkle, are very cautious and even skeptical (Bronstein, 2010; Sprinkle, 2010), many if not most authors in the current debate voice an overwhelming enthusiasm concerning the feasibility and future applications of moral bioenhancement. However, based on the empirical possibilities available today and in the near future, this enthusiasm seems somewhat misguided. The lack of scientific scrutiny is particularly striking when the possibility of genetic screening and modification to morally enhance individuals and potentially reduce criminal behavior is put forward (e.g. Walker's Genetic Virtue Project). Although genetic findings may improve our understanding of the risk factors associated with criminal behavior, we are far from identifying genetic risk factors for crime that could predict with reasonable certainty which individuals are at greater risk of engaging in criminal behavior. There is no one-to-one relationship between biological factors and criminal behavior. Indeed, depending on the environmental context, many individuals that carry biological risk factors for such behavior will not develop it, while others who do not show these risk factors might (van Goozen & Fairchild, 2008; Glenn & Raine, 2014; Glenn et al., 2015). Genetic modification that could lead to reliable moral enhancement is extremely far removed from our present-day knowledge and capacities, and it is doubtful whether it will ever be successfully achieved.

Morality and moral behavior are associated with so many different areas of the brain, that it has been claimed that morality is everywhere and maybe nowhere in the brain (Young & Dungan, 2012). Offenders with impaired moral decision-making, such as individuals with antisocial personality disorder, show overlapping abnormalities in several of these brain areas. Morality and moral behavior are very complex human traits and this is reflected at the developmental, experiential and neuroanatomical level, and most likely at the genetic level as well. It is misguided to think that we will be able to identify single genes or a single combination of genes that underlie morality and moral behavior. As most psychiatric and personality disorders are polygenic (i.e. involve a set of genes) and genetically heterogeneous (i.e. different sets of genes underlie the same diagnosis), it is highly likely that complex cognitive-emotional processes such as morality and (im)moral behavior are similarly associated with a myriad of genes and substantial genetic heterogeneity among different individuals. Similarly, not all antisocial individuals show the same biological deficits and a wide range of biological and environmental factors may contribute to antisocial behavior in a variety of ways (Glenn et al., 2015).

Other potential interventions for moral enhancement could be neurofeedback, transcranial brain stimulation (e.g. magnetic stimulation, and direct current stimulation), electrical stimulation of the brain via electrode implants (e.g. deep brain stimulation or DBS), and neuron replacement therapy (currently investigational). Although some of these neurotechnologies have reached the stage of demonstrated clinical effectiveness for certain disorders (e.g. neurofeedback for ADHD, TMS for depression, and DBS for Parkinson's), few have reached that level regarding phenotypic traits that likely contribute to (im)moral behavior. Nevertheless, preliminary studies indicate that transcranial brain stimulation might be effective for addiction, isolated case-studies exist where electrical brain stimulation via electrode implants is applied for chronic aggression or addiction, and preliminary experimental research suggests that functional magnetic resonance imaging neurofeedback could hold some potential for addressing addictions, antisocial personality disorder, psychopathy and sexual disorders. Improving 'normal' moral traits or behavior with such means is even further away from being practically feasible at present.

The idea of reliably bioengineering complex cognitive-emotional processes such as altruism or virtues is not feasible in the near future and it is highly unlikely even in the distant future. Because so many different biological and environmental factors influence an individual's (im)moral behavior, we are convinced that biomedical means alone will not suffice for moral enhancement. Hence, the debate on moral enhance-

ment should take as its starting point a combined approach in which traditional methods and emerging biomedical methods are used in tandem.

Little Attention to Side-effects, Risks and Safety

Except for psychopharmacological or hormonal treatments, few actual or potential interventions for moral bio-enhancement have been discussed in the moral enhancement debate. Moreover, whenever specific biomedical interventions are discussed, it is particularly worrisome that surprisingly little attention is given to side-effects, risks and safety-issues. Every biomedical intervention in the brain will likely have unintended, unwanted or unexpected side effects, especially so in cases where the underlying mechanisms of action are not well-understood and/or the procedure is invasive.

While neurofeedback and transcranial stimulation are non-invasive procedures, electrical stimulation of the brain via electrode implants and neuron replacement therapy are highly invasive. For example, aside from risks associated with brain surgery and electrode placement (e.g. brain hemorrhage, infection, death), DBS for Parkinson's disease carries a 1.1-33% risk of cognitive side effects (e.g. speech disturbance), a 1.3-10.2% risk of behavioral side effects (e.g. hypomania), a 0.5-25% risk of psychiatric side effects (e.g. depression) and a 50-71% risk of familial side effects (e.g. marital problems) (Clausen, 2010). While electrical stimulation of the brain via electrode implants is essentially reversible, neuron replacement therapy is a non-reversible procedure involving the injection of stem cells into the brain or spinal cord. Aside from the risks associated with surgery and stem cell injections, this carries the risk of tumor growth, seizures or intractable pain (Nuffield Council on Bioethics, 2013). Finally, even more familiar interventions such as pharmaceuticals and hormones have risks and side-effects, and their long term effects are not always known. In sum, in contrast to what has been the case so far, safety should be a key issue in the debate on moral enhancement.

Identity Changes

It is particularly surprising from a philosophical point of view that so little attention is given in the debate to unintended, unwanted or unexpected identity changes and the huge impact these changes may have on one's self-understanding, well-being and social and familial relationships. This problem has been discussed quite extensively in relation to psychopharmaceuticals (especially SSRI's) and DBS, but potential identity changes due to changes in one's moral dispositions or behavior are not often touched upon in the moral enhancement debate.

Within the broader biomedical treatment and enhancement debate, several authors have convincingly argued that whereas drastic identity changes are problematic from

an ethical perspective, typically requiring the discontinuation of the treatment or intervention, mild or moderate changes are not necessarily ethically problematic (e.g. Baylis, 2013). The key philosophical concept that is at stake is the concept of narrative identity rather than numerical identity. Narrative identity reflects an individual's most central and salient characteristics (e.g., motivations, beliefs, values, desires, character traits) that together comprise their self, and needs to be understood within the dynamics of psychological change. These characteristics may and often do change throughout one's life in response to various life events. It is important for the continuity of narrative identity that such changes are or can be incorporated into one's life story in a coherent way, without compromising one's sense of self. Since many if not most of such 'naturally' or 'traditionally' occurring changes are experienced in a non-problematic and identity preserving way throughout our life, one can, by comparison, argue for the ethical acceptability of narrative identity changes resulting from biomedical treatment or enhancement interventions (DeGrazia, 2005; Schechtman, 2009). At the same time however, it could be the case that unforeseen, instantaneous, uncontrollable and/ or drastic changes in one's *moral* dispositions or behavior are more likely to disrupt one's narrative identity, giving rise to a whole new array of personal and even societal queries. In agreement with the latter view, in a public opinion study on biomedical enhancement, individuals reported to be most reluctant to undergo enhancements of traits that are more fundamental to the self (e.g., morally relevant traits such as empathy and kindness) and the most frequently voiced reasons for resisting those enhancements were concerns of changes to their fundamental self (Riis et al., 2008).

Summary

In this paper, we have categorized and discussed the arguments in the published debate on the ethical desirability of moral bioenhancement. We have organized these arguments under the following headings: (1) why we (don't) need moral bioenhancement, (2) it will (not) be possible to reach consensus on what moral bioenhancement should purport, (3) the feasibility of moral bioenhancement and the status of current scientific research, (4) means and processes of arriving at moral improvement matter ethically, (5) arguments related to the freedom, identity and autonomy of the individual, and (6) arguments related to social/ group effects and dynamics.

After discussing each argument in more detail, we have identified a number of issues that in our view merit greater attention. First, we observed that, in the debate so far, discussions about the moral enhancement of 'humanity as a whole' and the targeted treatment of specific mental health disorders (such as psychopathy) are not sufficiently

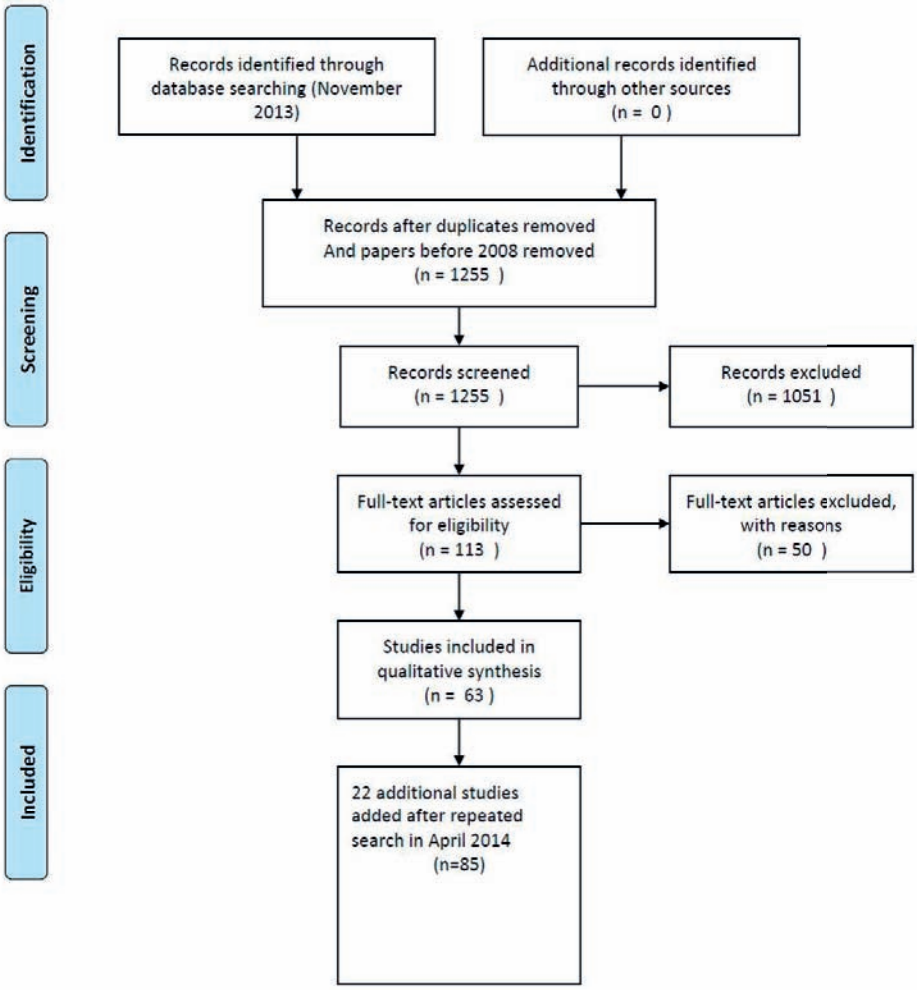
distinguished. Many authors overestimate the scientific as well as the practical feasibility of the interventions they discuss, rendering the debate too speculative. Related to this is our observation that insufficient attention is devoted to possible side-effects, risks and safety. There is also remarkably little attention to questions about identity and identity change.

We believe that the debate on moral enhancement is extremely interesting from a meta-ethical point of view, since it triggers important questions about the nature of morality, moral thought and moral behavior. However, the normative ethical question as to whether moral bioenhancement *as such* is good or bad, desirable or not, is not a very fruitful question for further debate. We therefore believe that, instead of speculating about non-realistic scenarios like the genetic engineering of morality, or other imaginary forms of biomedical moral enhancement of 'the whole of humanity', it would be much more useful to discuss novel and emerging biomedical interventions that may improve moral capacities or moral behavior in specific target groups and in relation to particular mental health problems. For the near future, biomedical treatment of moral pathologies may be a more realistic option than moral enhancement, and may raise more concrete moral questions.

In order for the moral enhancement debate to move beyond its focus on speculative philosophical theorizing and discussion, we need in-depth analyses of both the practical feasibility of existing or novel biomedical interventions for moral therapy (and perhaps eventually enhancement); and the ethical acceptability of such interventions, including safety concerns.

We conclude that although the discussion on moral enhancement so far raises interesting questions on an abstract, philosophical level, it often appears to be too remote from real (and realistic) contexts and applications to do justice to the specific ethical questions raised by such practices. We therefore urge for a more focused debate on realistic options of biomedical treatment of moral pathologies and the concrete moral questions these treatments raise.

Appendix A



Appendix A PRISMA 2009 Flow Diagram

3

Imagining moral bioenhancement practices. Drawing inspiration from moral education, public health ethics, and forensic psychiatry

Specker, J., & Schermer, M. H. N.
(2017). *Cambridge Quarterly of Healthcare Ethics, 26*(3), 415-426.

Abstract

In this article, we consider contexts or domains in which (future) moral bioenhancement interventions possibly or most likely will be implemented. By looking closely at similar or related existing practices and their relevant ethical frameworks, we hope to identify ethical considerations that are relevant for evaluating potential moral bioenhancement interventions. We examine, first, debates on the proper scope of moral education; second, proposals for identifying early risk factors for antisocial behaviour; and third, the difficult balancing of individual freedom and third party concerns in (forensic) psychiatry. In imagining moral bioenhancement in practice, we observe that unlike other forms of enhancement, moral enhancement fundamentally asks how the interests and preferences of the individual and the interests of others should be weighed (in view of public safety and managing public risk). Highly diverse domains such as education, mental health, and the judicial domain might be involved, and moral bioenhancement might challenge existing institutional settings. Given these highly varied contexts and domains, it appears unlikely that there will be a distinct set of practices that will be referred to as "moral bioenhancement."

Introduction

In the literature on the ethical desirability of moral bioenhancement, a range of different potential interventions and applications has been proposed. For example, following the growing understanding of biological (neurological, genetic) underpinnings of aggression, some argue for morally enhancing (potential) violent offenders. Others have speculated about enhancing our capacities for fairness, cooperation, and empathy, and for countering (supposedly innate, evolutionary hard-wired) xenophobic and racist biases (Douglas, 2008).

Recurring themes in the debate on the ethical desirability of moral bioenhancement are questions about the need for moral bioenhancement; the possibility of reaching consensus on what moral bioenhancement should realize; its scientific and practical feasibility; questions whether the difference between traditional and biomedical means of arriving at moral enhancement matters ethically; arguments related to the freedom, identity, and autonomy of the individual; and arguments related to social/group effects and dynamics (Specker et al., 2014). More recent discussions have focused, among other concerns, on the ethical relevance of the distinction between direct and indirect moral bioenhancement (Focquaert & Schermer, 2015; Schaefer, 2015), the effect that prescription drugs could already be having on moral agency (Levy et al., 2014a, 2014b), the question of whether mandatory moral bioenhancement would be permissible or even required, and how people could be encouraged to participate in moral bioenhancement voluntarily (Ingmar Persson & Julian Savulescu, 2014b; Rakić, 2014b; S. Carter, 2015).

In an earlier article, we argued that although discussions on moral enhancement so far raise important philosophical questions about the nature of human morality, moral behavior, and moral thought, these debates often appear to be too remote from real (and realistic) contexts and applications to do justice to the specific ethical questions raised by such practices (Specker et al., 2014). We therefore advocated a more focused debate on realistic options of biomedical treatment of moral pathologies and the concrete moral questions these treatments raise.

As the science on moral bioenhancement is "in its infancy" (Crockett, 2014c), it is not known which ethical issues will play out once potential moral bioenhancement technologies are more fully developed. In what Alfred Nordmann has termed the "if and then syndrome," a possible technological development is presented as inevitable and as something that demands immediate attention (Nordmann, 2007, p. 31). We agree that speculating about potential applications of technologies currently under

development—such as moral bioenhancement—without specifying intended users, target groups, and contexts of implementation runs the risk of "exploring a misguided or irrelevant set of ethical issues" (Brey, 2012, p. 2), while at the same time overlooking issues that would be important to consider once these technologies are implemented. This is why we think that the rather speculative and abstract debate on moral bioenhancement so far can benefit from closely examining either a number of contexts in which interventions under the heading of moral bioenhancement might first be implemented, or domains that are in one or more aspects importantly similar to potential moral bioenhancement practices, and therefore can inform our ethical thinking.

In this article, we start from the assumption that it is helpful to consider in which contexts or domains (future) moral bioenhancement interventions possibly or most likely will be implemented. By looking closely at similar or related existing practices and their relevant ethical frameworks, we hope to identify ethical considerations that are relevant for evaluating potential moral bioenhancement interventions. Focusing on how we deliberate ethically within these domains now will help identify ethical considerations specific to those domains. We think the moral bioenhancement literature can and should build on the ethical literature about these related practices. We also believe that it might point to ways in which moral bioenhancement differs from other types of enhancements such as sports, beauty, and cognitive enhancement. And finally, we think that because of these differences, ethical evaluation of moral bioenhancement asks for partially different evaluative frameworks. With this article we hope to contribute to the development of such frameworks.

Given these considerations, in this article, we have chosen to examine, first, debates on the proper scope of moral education; second, proposals for identifying early risk factors for antisocial behavior, and third, the difficult balancing of individual freedom and third party concerns in (forensic) psychiatry.[13] If and when moral bioenhancement technologies would become available, we expect that they would first be implemented within these domains. Moreover, we think that a number of central themes that are currently discussed within the debate on moral bioenhancement are the subject of debate within these three domains as well. For example, what moral enhancement should purport in contexts of moral pluralism is already being discussed extensively within the literature on moral education. Likewise, whether and under what conditions treatment should be made mandatory is a topic of debate in the literature on (forensic) psychiatry. In the following paragraphs, we intend to show that ethical considerations

13 This list is not exhaustive, and closer examination of other domains might be of interest as well; for example, the domain of artificial intelligence (Savulescu & Maslen, 2015).

that are central within these existing practices can inform and add to current debates on moral bioenhancement.

Moral Bioenhancement in Context

Moral Education and the Right to an Open Future

One often-cited reason why the possibilities of moral bioenhancement should be explored is the supposed lack of effectiveness of so-called traditional methods of moral enhancement, such as upbringing, socialization, and education. Another recurring argument is that there are little principled differences between employing traditional and potential biomedical methods of moral betterment in terms of their ethical acceptability. David DeGrazia, for example, argues that (many of the) arguments against biomedical means also apply to traditional, non-biomedical means: "one should not inculcate moral values that are wrong, so how can a parent be sure that she or he is justified in providing a particular type of moral instruction? Also facing this challenge are public school teachers who attempt to inculcate in students certain moral virtues such as civility, respect for differences and concern for the poor" (DeGrazia, 2014, p. 363).

An example of a domain in which these same questions are explicitly discussed is the field of so-called "moral education." What is the proper aim of (moral) education and what is the appropriate role of parents and educators in guiding the moral development of children? What particular moral codes and values do we want to nurture or even instil in our children, and how much pluralism are we willing to tolerate? What aspects of the family, social, and schooling environment impact on the development of prosocial and moral competencies of children?

The theoretical and empirical study of moral functioning—of moral cognition, moral emotions, moral behavior, moral motivation, moral character, and moral development—has attracted a lot of interest from a wide range of scientific disciplines in recent years (Young & Dungan, 2012; Darragh et al., 2015; Decety & Wheatley, 2015). At what Daniel Lapsley and Gustavo Carlo have termed 'the new crossroad' of moral development research, questions arise about the relevance of these different lines of research for "addressing the pressing social issues of our times" (Lapsley & Carlo, 2014, p. 3).[14] They argue that in order to adequately address these issues, broad, integrative

14 The diverse social issues they mention are frequently discussed in the debate on moral bioenhancement as well: violence, genocide, and war; concerns about environmental degradation; poverty and famine;

and multidisciplinary theoretical and methodological approaches are needed. One important similarity between discussions on moral education and those on moral bioenhancement is a search not only for a scientifically adequate model of human moral psychology, for linking "substantive philosophical issues with relevant (neuro) psychological data and in providing the empirical paradigm with which to do so" (Frimer & Walker, 2008, p. 334), but also an explicit focus on the practical implications of this growing body of knowledge.

Traditionally, within the literature on moral education, two main dominant educational approaches to the moral formation of children can be distinguished: traditional character education—focusing on the inculcation of virtuous traits of character—and rational moral education—seeking to facilitate the development of autonomous moral judgment and the ability to resolve disputes and reach consensus (Narvaez, 2006, p. 703). The latter approach focuses on nurturing skills such as perspective taking and deliberation more than on direct instruction on what to value and how to act (Narvaez, 2006, p. 708), and has, therefore, been characterized as a more progressive approach to moral education. Traditional character education on the other hand focuses on internalizing values inherent in the tradition and culture of society (Schuitema et al., 2008). Darcia Narvaez formulates a number of familiar strengths and weaknesses of both approaches:

> Whereas rational moral education adopts constructivism and adult coaching, fosters reasoning for civic engagement, and avoids relativism, it can be criticized for a narrow emphasis on moral reasoning, whether in dilemma discussion or a just community, which is insufficient for moral action and misses the centrality of moral identity in moral behaviour. Traditional character education rightly emphasizes the importance of content and demonstrated some insight into the impact of environments. However, it can be faulted for a changing set of core virtues open to the charge of relativism, for downplaying the importance of autonomy, and for a problematic pedagogy (Narvaez, 2006, p. 712).

More integrative and comprehensive perspectives on moral functioning and the implication for moral development and education are explicitly being sought, for example in a special issue of the *Journal of Moral Education* in 2008 (Turiel, 2008). Dissatisfaction with "the near exclusive focus on verbal reasoning about justice in a large portion of the research in the field" (Reed & Stoermer, 2008, p. 418) has inspired attempts to formulate unifying, integrative models that encompass "not only personality but,

---

and the persistence of racism and discrimination.

also, on the one hand, the brain and central nervous system, and on the other hand, interaction and culture" (Reed & Stoermer, 2008, p. 419; Kim & Sankey, 2009), and that acknowledge the importance of both reflective reasoning (in the Kantian or utilitarian traditions), moral emotions (based in the Humean tradition), and virtues and good practical judgement (following the Aristotelian tradition) (Narvaez, 2006, p. 703; Reed & Stoermer, 2008, p. 419; Wren, 2014). This literature could inform the moral enhancement debate.

Both in the debate on moral bioenhancement and within moral education, a central and much discussed topic of debate concerns the question what moral bioenhancement should focus on given the fact that ethical systems and theories differ considerably (Schaefer, 2011; Shook, 2012; DeGrazia, 2014; Hauskeller, 2014; Sparrow, 2014a; Ingmar Persson & Savulescu, 2015a). Given a plurality of moral values and outlooks, a question that has been asked in the philosophy of education is whether parents and educators are justified in raising and shaping children according to their own preferences: "Should our goal be to raise our children so that they will have, as adults, as many options as possible, to give them, insofar as we can, a maximally 'open' future? Or should our goal be more directive, to lead our children toward a more specifically shaped future that we ourselves endorse?" (Mills, 2003, p. 499)

The "right to an open future" argument concerns the balancing of the liberty of parents to raise their children according to their own lights, and the "anticipatory autonomy rights" of children to leave important life choices open. Joel Feinberg who introduced the right in 1980, discusses adults' rights to exercise their religious beliefs, a right that a child is not yet capable of exercising, but that still holds "prematurely" as a so-called "right-in-trust" (Feinberg, 1998, p. 251). The argument has become commonplace in applied ethics in debates on genetic reproductive technologies, through which disclosure of certain types of genetic information could infringe on childrens' right to have future's options kept open until they are capable of making their own decisions (Lotz, 2006, p. 537).

In the context of education, a child's right to have important life choices kept open includes "restrictions on what parents (and others) are allowed to do to children, and, on some interpretations,…what parents (and others) ought to provide children" (Millum, 2014, p. 522). Besides an argument for refraining from imposing important life choices, the right to an open future might also be interpreted as a "positive claim right" with a corresponding obligation for parents and educators to ensure that a child has a good starting position (Lotz, 2006). This right to an open future argument could therefore also be used within the context of moral enhancement: on the one hand, to argue for

limitations on what parents or others may do to "morally enhance" children, and on the other to argue in favor of instilling capacities that would help them to become full moral subjects with their own moral autonomy.

Early Identification of "Risky Children" and the Moral Pitfalls of Screening

It has been argued that "early childhood is probably the optimal starting point for moral enhancement" (Christen & Narvaez, 2012, p. 26), and that neuroscientific knowledge may inform traditional forms of moral education as well as enable early identification and prevention of antisocial behavior. Research is underway to determine which biomarkers—genetic, neurobiological, and physiological—give the most accurate risk predictions (van Goozen & Fairchild, 2008; Singh & Rose, 2009; Ferguson, 2010; Liu, 2011), and to determine which preventive interventions would be effective (Rocque et al., 2012; Cornet et al., 2013; Glenn & Raine, 2014; Glenn et al., 2015). "Biological information may provide useful information about which individuals may be at somewhat greater risk for antisocial behaviour, and thus may provide for the opportunity to intervene with programs designed to reduce this risk" (Glenn et al., 2015, p. 1690).

Early detection of children "at risk" for developing antisocial personality disorder or for becoming violent criminal offenders is ultimately aimed at prevention of antisocial behavior and violent crime.[15] Envisaged biomedical interventions aimed at prevention of such behavior can be understood as moral enhancements, in the sense of improving the moral capacities and/or behavior of these children. It aims to make them "morally better" persons than they would otherwise be. An (as yet largely hypothetical) comprehensive program for screening and early intervention to prevent antisocial behavior can, therefore, be understood as a program for moral enhancement. Some may oppose this framing, as consensus on the definition of moral enhancement is lacking (Raus et al., 2014). However, even if such screening and prevention practices do not meet the criteria of one's preferred definition of moral bioenhancement, the similarities are obvious. The emerging ethical discussion around early detection and prevention of antisocial behavior may therefore inform the moral enhancement debate and introduce new and important considerations and arguments.

15 As Nikolas Rose (2010) rightly observes about the aims of such programs: "The first, is the desire to identify risky individuals—that is to say, those who will present a future risk to others—before the actual harm is committed. The second is the hope that one might be able to identify individuals at risk—those whose particular combination of biology and life history makes them themselves susceptible to some future condition" (Rose, 2010, p. 80). Hence two notions of risk—risk to self and risk to others—tend to become conflated.

First, in the moral enhancement debate so far, the idea that a suitable target population will need to be determined for any moral enhancement intervention that may be developed has not yet been considered. Targeting interventions at specific groups; for example, those who already exhibit immoral behavior, or those who are at risk of doing so, requires a way to distinguish those people from others. Some form of screening will be necessary to determine who is in need of moral enhancement, who will benefit most from a specific intervention, or who need to be targeted to optimize societal benefit. A first—moral—question is who the target population should be; a second, more practical, question is how these subjects can be identified.

When such a hypothetical screening program for children "at risk" for antisocial behavior is considered, a number of ethical issues come up that have so far hardly been addressed in the moral enhancement debate (Schermer, 2006; Pieri & Levitt, 2008; Horstkötter & de Wert, 2013; Glenn & Raine, 2014; Horstkötter, Berghmans, & de Wert, 2014; Munthe & Radovic, 2015). First of all, screening tests may give false negative or false positive outcomes, resulting in unjustified reassurance and failure to intervene, or in unjustified labeling and unnecessary interventions. Limitations to the predictive value of screening tests are, therefore, morally problematic. Of many identified risk predictors, it is unclear whether they are causally relevant factors or mere correlates, making their exact meaning doubtful (Wikström, 2010). Moreover, screening tests themselves or the follow-up interventions could have side effects or carry risks; therefore, issues of safety should be considered. Next, it has been argued that such a screening program may have unwanted social effects. It may lead to individualization of problem behavior and lead to neglect of social and environmental factors contributing to this behavior. It may also lead to stigmatization and discrimination of children who are labeled (correctly or falsely) as "at risk." This may also have negative effects on their self-identity or even become a self-fulfilling prophecy: if one is looked on and treated as a potential violent offender, this may cause one to start behaving as such. One could even say that this threatens the child's "open future." At the same time, too strong a focus on antisocial behavior "might disadvantage those in greatest need of good youth care who are not also, for example, genetically at risk to translate early personal into later social problems" (Horstkötter & de Wert, 2013, p. 20). Finally, it has been brought forward that such screening programs could easily lead to oppression and exclusion of the identified individuals, especially in the absence of effective preventive or therapeutic interventions (Munthe & Radovic, 2015).

All these considerations are highly relevant for understanding what it would mean to single out certain individuals or groups as suitable candidates for moral improvement,

but they have so far been neglected in the debate on moral enhancement, because this debate has hardly concerned itself with potential "real world" practices.

A third lesson to be drawn from the discussion about early detection and prevention of criminal or antisocial behavior concerns the tension between the interests of the individual and those of society. This is mentioned as a very important issue by almost all commentators in the screening debate. Briefly stated: is the intended purpose helping the detected children, or is it protecting society? What if both aims cannot be served at the same time? Which one should take precedence? "At the core, perhaps, is the challenge of ethically balancing public protection and individual autonomy, privacy, and liberties" (Giordano et al., 2014, pp. 81-82). This issue will be taken up further in the next section.

Mandatory Treatment in (Forensic) Psychiatry and the Dual Role Dilemma

We can imagine examples of people who freely choose to morally enhance themselves using biomedical technologies, in which case an individual choice frame that emphasizes the freedom of the individual and related concepts such as autonomy and self-determination appears the most suitable ethical framework. Nonetheless, in the debate on moral bioenhancement so far, the theme of "mandatory moral bioenhancement" has been a central one. Examples that are frequently discussed range from preventing individuals from initiating "ultimate harm," to preventing violence or violent crime in the population at large (Ingmar Persson & Savulescu, 2008; Ingmar Persson & Julian Savulescu, 2014b). The argument is made that those who are most in need of moral enhancement are least likely to be inclined to use it, and that mandatory use is therefore required. Subsequently, there has been explicit and quite extensive discussion on whether or not (safe and effective) moral bioenhancement technologies should be mandated either for the general population or for specific target groups (Curtis, 2012; Rakić, 2014b, 2014a; S. Carter, 2015; Rakić & Hughes, 2015).

This raises a number of questions with respect to who decides, implements, and monitors the mandatory intervention in terms of expected and appropriate contexts, patient rights, and professional roles and responsibilities. These questions are reminiscent of a number of central ethical dilemmas in psychiatry, and in forensic psychiatry in particular, surrounding involuntary or coercive treatment, domains that are arguably a likely setting for implementing potential moral bioenhancement interventions.

In general, the conditions under which coercive treatment is warranted in mental health now involve direct prevention of suicide or other forms of self-inflicted harm, or situations when an individual poses a direct risk to others. Generally speaking, within

mental health settings "protection of the community is typically a side constraint, something that moderates the treatment of the client or patient" (Ward, 2013, p. 95), whereas "a fundamental obligation to the best interests of one's patient" remains central (Robertson & Walter, 2008, p. 233). However, when focusing on forensic contexts[16] ethical issues become more complex and controversial (Sharma & Sharma, 2006, p. 98), as these contexts are much more motivated by non-offender considerations (Ward, 2013, p. 95).

The so-called "dual-role dilemma," central in the ethics of forensic psychiatry, refers to possible tension between psychiatrists' obligations of beneficence toward their patients, and conflicting obligations to the community or third parties (Robertson & Walter, 2008). This tension or dissonance is something an individual psychiatrist might experience when "attempting to adhere to the conflicting ethical requirements associated with client well-being and community protection" (Ward et al., 2015). This tension might also be inextricably linked to the profession, as forensic psychiatry "involves an interaction between two distinct state institutions, the criminal justice and mental health systems" (Ward, 2013, p. 92) that each carry varying sets of ethical norms that can conflict.

In a classic text from 1984, Alan Stone formulated fundamental criticism on the lack of clear ethical boundaries for the professional roles of forensic psychiatrists (Stone, 1984). His main worry was that the role of evaluator moves the forensic psychiatrist away from the role of physician and the fundamental notion of nonmaleficence: "helping the patient, which is the ethical thesis of the practitioner, becomes the ethical temptation in the legal context" (Stone, 1984, p. 171). Paul Appelbaum on the other hand argued that beneficence and nonmaleficence are not central notions in forensic psychiatry, and that forensic psychiatry is instead guided by a distinct set of ethics. The ethics of the "forensicist" is directed toward the benefit of society, and, therefore, the central responsibility of the forensicist is to justice, not the patient (Appelbaum, 1990, 1997). A notable example is forensic risk assessment, where the welfare of the patient is not necessarily the immediate object of concern, but rather the immediate concern is about public safety (Roychowdhury & Adshead, 2014).

16 Contexts that involve court-mandated forensic psychiatric evaluation and assessment, risk management, and forensic treatment. There is great variation in what the profession of forensic psychiatrist entails. There are large differences in the range of forensic psychiatric services available, there are no unified standards, and forensic psychiatry is not recognized as distinct subspecialty everywhere (Velinov & Marinov, 2006).

This distinction suggests three directions for mandatory moral bioenhancement interventions. Either moral bioenhancement will be situated firmly within a health context, in which the interests and well-being of the patient are central and the patient's right to treatment and other forms of support is emphasized, and in which in principle medical doctors prescribe a potential moral bioenhancement intervention, and monitor and weigh potential risks. An alternative option is to position mandatory moral bioenhancement within correctional, criminal justice contexts, where non-offender considerations, such as community protection and risk assessment, are primary. Whereas the first option presupposes the primacy of medical ethical considerations, the latter would function within a more general ethical framework, in which respect for persons, justice, and human rights are important considerations. A third, hybrid model would explicitly acknowledge that "more than one ethical theory will be required to justify and guide offender treatment" (Ward, 2013, p. 97), and would seek a pragmatic, procedural solution to solve potential conflicts between the different sets of norms and values. The Dutch Entrustment Act for example represents a rather exceptional combination of punishment and treatment (van Marle, 2002), a system that is criticized for (possible infinite) preventive detention (Petrila & de Ruiter, 2011), but that is also praised for the quality of the care provided to prisoners and the lower recidivism rates. However, the fact that "individuals may be unable to ethically justify their professional actions and could move somewhat erratically between different courses of action" (Ward, 2013, p. 98) remains a concern, one that should also be addressed with respect to those who would implement potential moral bioenhancement interventions.

Imagining Moral Bioenhancement Practices

To distil a number of ethical considerations that are relevant for the moral evaluation of proposals for moral bioenhancement but have not yet been sufficiently recognized in the discussion on moral bioenhancement yet, we have imagined practices and domains in which potential moral bioenhancement might first be implemented.

In the domain of moral education, a similar focus on the practical application of the scientific study of morality is visible. In seeking practical applicability and effectiveness, it transcends the separation of moral functioning into cognitive, emotional, and affective domains, and instead aims to formulate integrative models and interventions. Within moral education debates, discussions on the appropriate role of educators and parents, and a commitment to securing the (future) autonomy rights of children, are visible. In addition to providing a reason to be cautious in forcing on children certain choices

(based on the negative claim right that follows from the right to an open future), this principle might under certain circumstances also provide reason to intervene in order to secure better chances in life for the child in question.

This connects to the second cluster: ethical concerns with respect to potential early detection of children at risk of exhibiting antisocial behavior (either now or in later life). On what basis the intended target population for potential preventive interventions should be determined (and what the intended population should be) is something that needs to be discussed, especially in view of potential unjustified labeling resulting in stigmatization and potentially unnecessary interventions that may carry risks and burdens. Moreover, a comprehensive screening program for antisocial behavior needs to address possible tensions between the interests of society and the interests of the individual, and be clear on its intended purpose.

In the field of forensic psychiatry as well, safety and risk management concerns often need to be balanced with the interests and well-being of the individual. The dual role dilemma describes the tension between distinct sets of norms that guide professional conduct, which may be "personal, universal, agency related, or professional in nature" (Ward, 2013, p. 98). If neither a purely medical nor a purely criminal justice frame is appropriate, professionals in forensic psychiatry risk being bound to navigate a complex ethical landscape without clear ethical frameworks to justify their professional conduct.

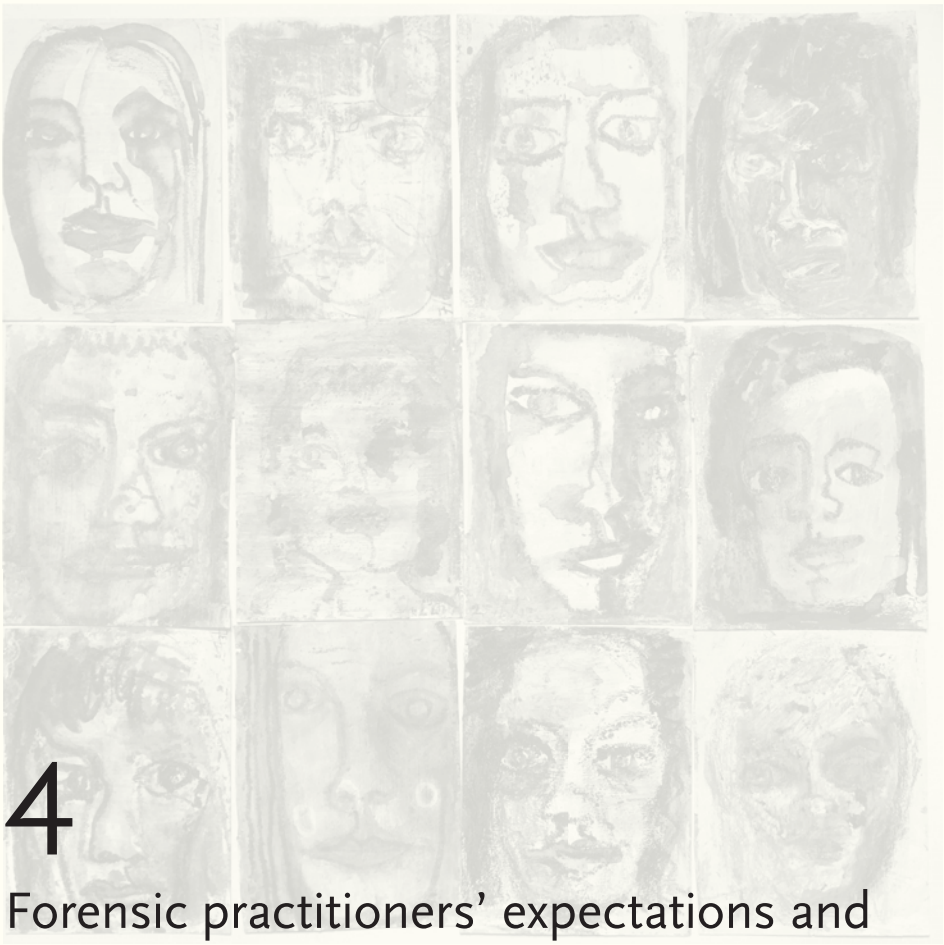
In imagining moral bioenhancement in practice, we have seen that unlike other forms of enhancement, moral enhancement fundamentally asks how the interests and preferences of the individual and the interests of others should be weighed (in view of public safety and managing public risk). Whereas enhancement in general is advocated for contributing to both the individual and collective good, moral enhancement appears to be predominantly aimed at the collective good. Because morality is fundamentally something relational, and hence asks for concern for others, a focus on interventions that would merely benefit the individual and not others is not a very likely perspective for *moral* enhancement. The opposite—moral enhancement solely in light of some greater good—might be possible, but would run the risk of neglecting the freedom and preferences of the individual. Especially in the genetic enhancement debate, such a one-sided focus on the benefits for society has been fiercely opposed because of its—historically evident—potential for abuse. Therefore, the need for balancing individual and collective interests may well be inherent in moral enhancement, and something that ought to be considered in more depth. Likewise, whereas it is an open question whether many people would consider moral bioenhancement voluntarily, the condi-

tions for offering or even mandating potential moral bioenhancement interventions should be discussed as well (E. Shaw, 2014).

Mandatory treatment is subject to strict conditions, and liberal democratic societies generally share a commitment to tolerance and value pluralism, and respect the (future) autonomy rights of individuals. Although value pluralism is considered a challenge within moral education debates, it is not considered a reason ("principled" or "practical") to stop moral education programs from being implemented. In line with the views from moral progressivists, an important "guiding concern" in implementing potential moral bioenhancement interventions might be to focus on enhancing morally relevant capacities, without imposing one particular comprehensive moral doctrine. However, it should be noted that in societies that do not share this commitment to pluralism and tolerance, any possibilities for moral bioenhancement might lead to risks.

In asking which occupational group would be charged with implementing moral bioenhancement, the particular institutional setting in which they work becomes highly relevant. With respect to cognitive enhancement, the options discussed in this respect are rather limited: either enhancement will be obtained from the free market, or it will be prescribed by physicians, or possibly by so-called "schmocters" (Parens, 1998, p. 11). With moral enhancement, highly diverse domains such as education, mental health, and the judicial domain might be involved, and moral bioenhancement might challenge existing institutional settings.

Given these highly variable contexts and domains, it appears unlikely that there will be a distinct set of practices that will be referred to as moral bioenhancement. Rather than attempting to formulate a single evaluative framework, the current debate on the ethical desirability of moral bioenhancement might point to aspects of moral betterment within our current practices, and might invite us to consider a number of ethical issues anew.



4

Forensic practitioners' expectations and moral views regarding neurobiological interventions in offenders with mental disorders

Specker, J., Focquaert, F., Sterckx, S., & Schermer, M. H. N. (2018). *BioSocieties, 13*(1), 304-321.

Abstract

Neurobiological and behavioural genetic research give rise to speculations about potential biomedical interventions to prevent, contain, or treat violent and anti-social behaviour. These developments have stirred considerable ethical debate on the prospects, threats, and limitations of integrating neurobiological and behavioural genetic interventions in forensic psychiatric practices, yet little is known about how forensic practitioners perceive these potential interventions. We conducted a qualitative study to examine: (i) the extent to which forensic practitioners expect that effective biomedical interventions will be developed and integrated in their daily work practice; and (ii) their normative views concerning those potential biomedically informed interventions. We focused on: potential biomedical possibilities to lower aggression, the possible usage of neuroimaging in assessing legal responsibility, and the potential use of biomarkers in assessing risk for future violent and antisocial behaviour. Forensic practitioners expect novel biomedical interventions to be developed and display a general openness towards them. At the same time, they express concern that integration of neurobiological and behavioural genetic elements in explanatory models of violence and antisocial behaviour may lead to misinterpretations, especially when implemented in the forensic field.

Introduction

Scientific developments in neurobiology and behavioural genetics give rise to speculations about potential biomedical interventions to prevent, contain, or treat violent and anti-social behaviour (de Ridder et al., 2009; Rocque et al., 2012; Glenn & Raine, 2014; van der Gronde et al., 2014; Eichelberger & Barnes, 2015; Lee, 2015; Fozdar, 2016). These potential developments have stirred considerable ethical debate on the prospects, threats, and limitations of integrating neurobiological and behavioural genetic interventions in forensic psychiatric practices (Rose, 2000; Pustilnik, 2009; Rose & Abi-Rached, 2013; Chhangur et al., 2015; Horstkötter, 2015; Munthe & Radovic, 2015; Hübner & White, 2016).

In this study, we intend to broaden the ethical debate by asking how practitioners in forensic mental health contexts (both forensic psychiatrists and clinical psychologists and therapists) view these potential interventions. Do they expect that effective biomedical interventions aimed at preventing, containing, and treating violent and anti-social behaviour will be developed? How do they normatively view the potential integration of such interventions within current treatment practices? For the present study, we interviewed forensic practitioners about their expectations as well as moral views regarding potential applications of current neurobiological and behavioural genetic research aiming to understand (and possibly help prevent, contain, or treat) violent and antisocial behaviour.

An earlier study asked people who are professionally engaged in in the criminal justice system and who work with young people and families judged to be 'at risk' of criminal or antisocial behaviour about their views on the causes of violent and antisocial behaviour, about their response to the idea of a genetic susceptibility to aggressive behaviour, and possible implications for their own work (Pieri & Levitt, 2008). Other studies explored the views of juveniles in juvenile justice institutions on biological explanations of anti-social behaviour (Horstkötter, Berghmans, Feron, et al., 2014) and the views of parents and (non-forensic) healthcare professionals (genetic professionals and paediatricians) on genetic tests for violent behaviour (Campbell & Ross, 2004). To the best of our knowledge, this is the first interview study to probe the normative views of forensic mental health practitioners concerning potential future biomedical interventions for forensic patients and offenders with mental health issues.

We first asked forensic practitioners, with reference to their role in treating forensic patients and offenders with mental health issues, about their expectations regarding potential future biomedical possibilities to lower aggression. Current research on the

neurobiology and (epi)genetics of aggression may inform new treatments for pathological, maladaptive, or antisocial aggressive behaviour (Asherson & Cormand, 2016). Aggressive behaviours are often partially refractory to medication (Raine, 2013; Gurnani et al., 2016; Temel et al., 2016). Research on novel medications and neuro-stimulation may provide new ways to deal with aggressive behaviours. Non-invasive interventions may involve biofeedback, transcranial magnetic stimulation, and transcranial direct current stimulation. More invasive potential interventions such as deep brain stimulation involve electrical stimulation of the brain through electrode implants (Temel et al., 2016).

Second, with respect to forensic practitioners' role as court-appointed experts, we probed their views on the possible uses of neuroimaging in assessing legal responsibility within court settings. Further developments in neuroimaging research may be used in criminal justice settings to inform forensic psychiatric and neurological expert reports (Witzel et al., 2008). Currently, the use of structural or functional brain images in court is more widespread in US courts as compared to European courts and it is mostly limited to the sentencing phase to acquire sentence mitigation (de Kogel & Westgeest, 2015; Ginther, 2016; Morse, 2016; McSwiggan et al., 2017). Future usage of brain imaging in court settings may involve establishing the presence or absence of psychiatric or neurological disorders (Roskies et al., 2013), and scientifically assessing intentionality, degrees of legal responsibility, or recidivism risk (Aharoni et al., 2008).

Finally, with relation to their role in prevention and public protection, we asked them about their views concerning the possibilities and desirability of early detection and intervention based on biomarkers. Preliminary research is underway to determine which biomarkers - genetic, neurobiological, and physiological – might give accurate risk assessments of future violent and antisocial behaviour and which preventive interventions would be effective in reducing those risks (van Goozen & Fairchild, 2008; Singh & Rose, 2009; Liu, 2011; Rocque et al., 2012; Singh et al., 2013; Glenn & Raine, 2014; Glenn et al., 2015; Gaudet et al., 2016). Other uses of biomarkers that are being investigated are biological predictors of inpatient violence, length of stay, and reoffending (Aharoni et al., 2013; Sedgwick et al., 2016).

Methods

Sample and recruitment

We recruited 18 forensic practitioners (forensic psychiatrists, clinical psychologists and therapists) in The Netherlands and in Belgium. Subjects were recruited via professional

organizations and by snowballing, and were not incentivized. Our sample consists of seven females and 11 males, ranging in age from 32 to 68 years. At the time of the interviews, nine participants were employed in The Netherlands and nine participants were employed in Belgium.

We conducted eight interviews with forensic psychiatrists (FP) (at the time of the interview, one participant worked as a general psychiatrist, but had worked in forensic settings in the past). We conducted 10 interviews with clinical psychologists (CP) or therapists (T) (at the time of the interview, one participant - a therapist - was primarily involved in research and did not consult patients, but had done so in the past). 10 participants (five psychiatrists and five psychologists) are involved in scientific research, alongside their clinical or therapeutic work.

Qualitative interviews

Participants took part in an individual semi-structured interview lasting approximately one hour. During one interview, two respondents were present and interviewed together. The interviews were held in Belgium and The Netherlands, and took place between July 2015 and July 2016. The interviews were conducted by JS, FF and MS. JS attended 14 interviews, FF attended seven interviews, and MS attended one interview.

The interview schedule included open-ended questions about participants' expectations towards and normative views on: potential future options for treating aggression; potential usage of neuroimaging in determining legal responsibility within court settings; and the question as to whether insight into the neurobiological and biosocial predictors of criminal behaviour might broaden opportunities for early detection and prevention of children and adolescents who might be considered at risk of future violent or antisocial behaviour. As forensic psychiatrists and psychologists, and psychiatrists and psychologists in general, are trained to understand and treat human behaviour from a medical and/or healthcare perspective, and faced with the many variations that may exist between people, with respect to both non-deviant and deviant behaviour on a daily basis, we expected to receive highly nuanced and diverse responses to our questions.

In addition to the open-ended questions, the interview schedule contained a number of examples (prompts) to enable the interviewer to gain a more detailed answer. For example, regarding potential future options for treating aggression, we would add: "One might well look at interventions such as TMS, DBS, or potential new pharmacological interventions. Do you expect interventions such as these to find entrance

in your professional practice?" In this paper, we report on our findings and discuss possible ethical implications.

Coding

All interviews were conducted in Dutch, audio-recorded and transcribed verbatim. Interviews were coded in QSR NVivo version 10 using descriptive theme analysis (Bazeley & Jackson, 2013). Because our methodological orientation was content analysis, and our research was not hypothesis-driven but data-driven, we aimed at exploring the broad, varied perspectives and meanings that participants hold (Tong et al., 2007; Creswell, 2013).

All transcripts were independently read by all members of the research team (JS, FF, MS, SS) and discussed with the purpose of drafting a preliminary analytic framework. JS independently coded the transcripts by labelling sections and text units referring to one or multiple concepts relevant for the study purpose. An iterative approach was used in which new data that challenged the existing coding structure were used to revise the themes until no new themes emerged. Interpretative bias of data was avoided by means of investigator triangulation, which involved all researchers checking the codes for consistency. Two researchers (FF and MS) independently coded three transcripts and compared their coding and categorizing with that of JS. Illustrative quotes that were included in the results section were translated by JS. All members of the research team checked the translations for accuracy.

Results

Aggression

Do forensic practitioners expect that effective biomedical interventions to lower aggression will be developed, and how do they morally view those potential interventions? A large majority of participants expect that novel interventions will be developed. At the same time, however, nearly all participants stress that violent and antisocial behaviour can have many different (kinds of) causes, that at present there is lack of fundamental knowledge, and that no simple, mono-causal model of aggression is (and will be) available. In view of this complexity, several participants question the effectiveness of one single (biomedical) intervention and urge for an approach combining biomedical and psychosocial interventions instead. Forensic practitioners' moral views of potential biomedical interventions are quite diverse, and range from outright rejection, to cautious embracement, and more positive and welcoming attitudes.

Complexity of aggression

Before offering an opinion on potential neurobiological and genetic interventions for treating aggression, nearly all participants discuss at length the considerable difficulty of formulating a correct and comprehensive model of aggression. Forensic psychiatrists particularly tend to emphasise the importance of accurate diagnosis, of understanding different types of aggression (e.g., recurring and constant, or incidental, premeditated or impulsive), and of unravelling possible associations with mental or personality disorders, intellectual disability, and addiction and substance abuse. This can be illustrated with the following quote from a participant:

> Apparently, with compulsive behaviours, behaviour regulation can be influenced on a fundamental level. With aggression, we need to ask, is it constant or incidental? Oftentimes, it is incidental and triggered in a specific situation. That is different from, for example, OCD, which is more or less permanent. (FP4)

Many participants emphasise the current absence of a comprehensive explanatory model of aggression. One example of a knowledge gap that is often mentioned is the lack of understanding of the relation between psychiatric disorders and (aggressive, deviant) behaviour.

> We know very little about the relationship, the causal link between disorders and behaviour. This applies to problems as we observe them and as we now classify them according to the DSM – let alone how that plays out physiologically. […] You can look at a brain, and you can say, 'well, maybe those neural pathways are less developed, or that lobe is a bit smaller or bigger' – but by saying that, you have not explained why someone is aggressive. Not everyone who has that brain-abnormality is aggressive. The reverse does not apply either. We are still very far from being able to intervene and remedy it [aggression]. (T1)

> If I try to understand how different factors interact, the only fair answer is that we do not know. We know a number of codetermining factors, but we have no idea how they relate to each other, how they might reinforce, stop, trigger other factors, set up a cascade or not. We do not understand why many people in spite of these factors continue to develop in a healthy and well-behaved way. (CP3)

Most participants (forensic psychiatrists as well as psychologists) discuss the ways in which societal norms and definitions of aggression differ and the ways in which aggression is often dependent on personal and social circumstances. Many participants stress that neither our understanding of aggression nor potential interventions to counter

aggression should focus exclusively on either individual or environmental factors, but that, instead, both should be taken into account.

> A general trend can be observed towards intervening at the biological or somatic level. We see this for example in sexual offenders, in a quite rigorous way. This entails a risk of taking a fairly one-sided approach. With aggression, it might be that treatment will develop in the same direction. Although I have yet to see if that would be possible as specifically as with sexual disorders. (FP6)

A number of participants identify barriers to the introduction of biomedical interventions, such as the need to build consensus and support among professionals with diverse disciplinary backgrounds, who work from within different paradigms. Some participants mention that, historically, research in forensic populations has not been a priority.

> Once you enter prison, you find yourself in a criminal justice context. Here, medicine stops, and research stops as well. That really struck me. It is an interesting, special pathology that you do not see in regular medicine. Yet hardly any research is conducted there. (FP2)

Aggression: Moral views

Several participants refer to the harmful side effects of many current (typically pharmacological) biomedical interventions and their negative impact on treatment adherence. Some participants express the hope that future biomedical interventions will have fewer side effects. Others, however, are concerned that interventions must be highly invasive in order to counter aggression effectively.

> I think to do this, for this to be really effective, I feel the interference with brain function needs to be so high impact, that things such as identity and personal performance would become radically standardised. (CP3)

Some participants express the concern that wider opportunities to intervene will instigate less tolerance towards deviant traits or towards acting out (both in society at large and in forensic contexts). They stress that deviant or seemingly unwanted traits (such as those associated with psychopathy) can be advantageous in some contexts, and that altering 'bad traits' or enhancing 'good traits' could have unforeseen and unwanted consequences (e.g., too much empathy).

Several participants question whether potential interventions may be mandatory (for example as a precondition for rehabilitation). Many stress the importance of voluntariness and proper informed consent procedures, even if – as a number of participants discussed – treatment in forensic contexts may often involve coercive offers. A number of participants speculate on how future biomedical interventions may offer options for patients who currently are considered treatment refractory, yet at the same time argue that invasive interventions should preferably (or only) be offered as a last resort.

> The more invasive and irreversible it is, the longer you should wait to consider it, I think. However, it need not be ruled out completely. (CP1)

Several participants prefer alternative, less invasive ways to limit aggressive behaviours and promote impulse control and self-regulation, e.g. cameras, gaming, apps, electronic monitoring, virtual reality training, etc.

Several participants discuss whether biomedical interventions would render the subject's role in treatment too passive. Some participants feel that passivity in this context is morally problematic in itself. For example because biomedical treatments tend not to consider the reasons why someone engaged in aggressive behaviour, or because they feel treatment should focus on enhancing a patient's capacity to manage herself or on fostering self-awareness and self-knowledge. Others are worried that passivity would reduce the treatment's long-term effectiveness, for example because they expect the aggression reducing effects to last only as long as the intervention is given.

Neuroimaging in the courtroom

Will enhanced options for neuroimaging improve the identification of criminal responsibility? Expectations vary widely; some participants indicate that they indeed expect that developments in neuroimaging will be rapid, and that neuroimaging will be used more often in court settings, while others situate such options in the distant future. By and large, most participants were highly sceptical of its potential effectiveness, and mentioned numerous moral concerns regarding its implementation.

Neuroimaging in the courtroom: Confrontation of legal and medical terminology

Also on this issue, many participants reflect on the conceptual framework before going into the expected use and moral desirability of using neuroimaging to assess legal responsibility. Several participants mention the difficulties associated with operationalizing responsibility, which is essentially a legal and not a medical concept, and argue that legal concepts fall outside of their expertise as forensic practitioners. Some even indicate that they feel they should not say anything at all on this issue.

> The whole concept of legal responsibility is almost a philosophical concept. As psychiatrists, and behavioural experts in general, we are limited in what we can say about it. (FP6)

Some problematize the concept of legal responsibility itself, and express a preference for a system in which forensic risk assessment instead of assessment of legal responsibility determines whether someone is referred to prison, treatment, or both.

Neuroimaging in the courtroom: Moral views

Many participants are concerned that imaging data will be misinterpreted and oversimplified. They refer to the difficulty of correctly interpreting neuroimaging results and the lack of expertise of, for example, judges or juries. They talk about the power of images and the illusionary appearance of objectivity they might convey. They discuss the false sense of security neuroimaging might offer, and how this might negatively impact legal processes.

> At one point, the image starts to determine who someone is, well, yes almost determining who someone is. And therefore also, how far his responsibilities, to what extent he is responsible for his behaviour. (…) I even think it's a risky development, because it claims that it can classify people very precisely. And taking into account the ways things tend to go in the confrontation between behavioural experts and the judiciary, the behavioural expert suddenly pretends to hold the truth. And if you're not careful, the judiciary will go along with that. (FP6)

Some participants are concerned that the use of neuroimaging results would reinforce ideas of incorrigibility, of 'violent brains', 'born criminals' or 'born monsters' who differ fundamentally from normal people, thus underestimating plasticity and individual adaptability and undermining ideas of individual agency and responsibility.

Several participants indicate that they feel that presenting imaging data unequivocally would fulfil a societal demand, for example because judges welcome it or because it would make it easier for many people to tolerate and accept crime.

> This man said 'Show me a brain and I can tell you if he is a paedophile or not.' That is something society begs for, I think, a doctor who clearly indicates based on an image, 'This is a paedophile and he is dangerous, he's a predator, lock him up.' Or 'Remove that part of his brain, break his skull open.' That is something society would really welcome. If only it would be so simple and easy to address. I think this

> is a serious danger. In this regard, I feel more than a professional resistance; I think we should not go there. (CP4)

Early identification and prevention of violent and antisocial behaviour using biomarkers

Finally, we asked participants whether insight into neurobiological and biosocial predictors of violent and antisocial behaviour might broaden opportunities for early detection and prevention of children and adolescents who might be considered at risk of future violent or antisocial behaviour. All participants (except for one psychiatrist) indeed expected biomarkers to be integrated in preventive screening programs and even considered this a largely unstoppable development. At the same time, however, participants expressed a strong professional resistance, discussed the substantial implications of such programs, and mentioned a range of moral concerns.

Many participants stressed the importance of prevention rather than cure, of being able to prevent pathology, and of intervening before people become involved with the criminal justice system. Some expressed a particular interest in identifying extreme cases as early as possible.

> The earlier we can intervene, the better. Because we often see people with such difficult pasts and with histories of failing assistance. This is due to the person in question, but also due to us. We do too little, or we fail to persevere. Legal frameworks also provide us insufficient opportunities to intervene, measures such as involuntary commitment for example. (FP3)

Pessimism about or reluctance towards prospects of using biomarkers in early prevention and intervention programs sprang from the following main considerations: the low predictive value of biomarkers; the lack of feasibility of large and comprehensive screening programs; and moral concerns about labelling and stigmatization.

Early identification and prevention of violent and antisocial behaviour using biomarkers: Feasibility of accurate prediction

Several participants stress difficulties associated with (or even the impossibility of) accurately predicting risk of future violent or antisocial behaviour. For example, they assert that hardly any direct, causal relations between physiological characteristics and future antisocial or criminal behaviour are established.

> It seems that a range of factors – that are also associated with each other – correlate with crime. I cannot see how you could intervene solely based on physiological

> characteristics. There are hardly any cases where such a one-on-one relation exists. (T1)

> Of course there is always the problem that there never is a one-on-one relation, and that all those factors that contribute to you going off track in such a way that you end up in civil commitment, are all, single-handedly, very prevalent. (CP6)

> Everyone is at risk for aggression. Because aggression is the most human thing there is. (CP5)

Some also mention the difficulty of determining the proper moment to intervene. They caution against intervening too early and note the problem of potential false positives and negatives.

Several participants draw a parallel with screening and prevention programs within general psychiatry, aimed at identifying people who are at a higher risk of experiencing psychoses, where a large group needs to be screened and monitored in order to identify the relatively small sub-group that is actually at risk and that could benefit from interventions. Participants contemplate the (cost-) effectiveness of such large and comprehensive programs, and mention difficulties associated with correctly interpreting biomarkers and formulating appropriate selection and inclusion criteria. A few participants problematize underlying conceptions of crime, delinquency, antisocial behaviour, and violence.

> Delinquency has to do with transgressing norms, which are nowhere to be found in the brain but are defined by a social group. (CP3)

Early identification and prevention of violent and antisocial behaviour using biomarkers: Moral views

In addition to concerns about feasibility, many participants express moral concerns regarding early prevention and detection programs. They are worried that such programs would be focused primarily at preventing future crime instead of helping children or relieving suffering, that such approaches would tend to be paternalistic and coercive in nature, and that they would target and single out individual children as "risky children". Many participants mention and strongly reject the likely stigmatizing and discriminatory (and even self-fulfilling) effects of labelling – especially young children – as being at risk for future antisocial behaviour.

> If I need to choose between labelling and prevention, I would say that it is more important not to label, especially young people. Labelling is based solely on risk, and so little on opportunities and possibilities for positive development. (FP1)

Although many participants underline the importance of prevention in general, they see important ethical differences between prevention programs aimed at precluding impaired development of the child or individual in question and programs that focus on preventing crime.

> What you strive for, is helping children, when they are hindering themselves or others, to develop in a positive way. Once you identify a child as a potential offender, you are already treading on very thin ice. (FP6)

Some participants welcome biomedically informed prevention efforts, provided that they aim to relieve current and future suffering, aid development, and offer opportunities to develop a full life. Some participants speculate about biomarker research offering insights into underlying factors, how such research may offer new, more precise treatment options, and how it could be used as a source of information in treatment settings.

> What I find more interesting is whether it [e.g. skin conductance] may provide indications for doing something different than we had imagined by looking at behaviour alone. That could give some sort of insight; 'You have been tense for a couple of days.' And that this would help prevent escalation. I find that very interesting, that you can use it as a source of information, to gain insight. (FP8)

A number of participants prefer general awareness and education campaigns (for example about impulsiveness for all school-aged children) or (psycho) education aimed specifically at individuals at risk:

> I support everything that has to do with self-awareness and with being able to deal with those personal characteristics that increase the likelihood of undesirable outcomes. (CP3)

However, some participants worry that other, non-biological markers (such as social and environmental factors) will disappear from sight, which would negatively affect the therapeutic relationship.

> I hope that it won't become self-evident, that my job in several years will be reduced to being confronted with a blood level and to determine on that basis which treatment to initiate, and whether treatment should be outpatient or inpatient. I hope I will still be able to be in contact with people. (CP4)

Some participants situate speculation about biomedically informed early detection and prevention programs in the context of broader societal trends, such as increased focus on early prevention, higher societal risk aversion, and decreasing tolerance for deviance and abnormality.

> I think these issues are very topical, because much appears to be controllable. At many levels, not only with regard to people but also in business, we assume everything manageable. I don't believe such malleability exists – we are keeping this illusion alive, but it is a kind of fake truth. (FP5)

> Everyone wants us to do more prevention. Authorities are asking, what should we do with those children in the neighbourhoods? Which signals must we take seriously and which not? When should we scale up? That has nothing to do with innovative technology, but rather with how far we should go in actively tracking and detecting. (FP8)

> At that moment, we will have to ask what kind of society we want to live in. North Korea will have a different perspective than European countries. If you decide as a society that you prioritise zero-risk policy, where the collective rather than the individual is the starting point, then I can imagine that politics will embrace such policies. Yes, I find that prospect quite disturbing. (FP5)

Discussion

In this study, we interviewed forensic practitioners in The Netherlands and in Flanders, Belgium about their expectations and moral views regarding potential biomedical interventions to prevent, contain, or treat violent and anti-social behaviour, explicitly focusing not only on potential efficacy but on moral desirability as well. Little is known about forensic mental health practitioners' perceptions of potential biomedical interventions for forensic patients and offenders with mental health issues.

Overall, this study indicates that forensic practitioners expect novel biomedical interventions to find entry in their professional practice, either because they anticipate

that rapid scientific developments will generate effective biomedical interventions, or, alternatively, because they observe a growing (societal, professional) demand for biomedical interventions (for example because of their perceived greater objectiveness and effectiveness as compared to non-biomedical interventions). This is a reason why, in our analyses, we have been careful to distinguish between four aspects: forensic practitioners' expectations that potential new biomedical interventions will become available; their expectations regarding the effectiveness of these potential interventions; whether they generally welcome biomedical interventions and hope new biomedical interventions will become available, or whether, on the contrary, they fundamentally ethically reject the introduction of potential future biomedical interventions in their professional practice.

Few forensic practitioners in our study seem to oppose biomedical interventions per se. Yet, when discussing their potential efficacy, most forensic practitioners strongly advocate epistemic caution. Participants underline that mono-causal explanatory models are not available and point to the complex (non-direct/ non-causal) relations between individual (neuro)biology, genetic susceptibilities, mental disorders, environment, SES, and violent and antisocial behaviour. They discuss difficulties associated with correctly interpreting neuroimaging results (see for example Rose & Abi-Rached, 2013), and warn that genetic susceptibility to violence (Pieri & Levitt, 2008) and predictions based on biomarkers lack individual predictive value (Starr, 2014).

Participants who in general welcome novel biomedical interventions anticipated or hoped for less severe side effects than those of current (mostly pharmacological) treatments. Moreover, they hoped that these interventions will provide insight into underlying causes of violent and antisocial behaviour as well as increased opportunities to tailor interventions to the individual in question (personalised medicine). Participants also expressed hope that such interventions might provide opportunities for increased self-awareness and self-regulation, and will make patients less dependent on contingent factors such as the particular care provider and the preferred form of (psychosocial) therapy.

Our study reveals a mixed picture with respect to forensic practitioners' moral concerns about potential biomedical treatments for aggression. Major scepticism and considerable moral concerns were expressed with regard to the use of neuroimaging in court. As to early detection and prevention using biomarkers, these were almost entirely rejected by our respondents.

By and large, participants seem to be aware of major ethical issues as they are being discussed in current ethical debates on potential biomedical interventions. These are: concerns about reductionism and disregard for socio-political circumstances in case of novel treatment options for violent and antisocial behaviour (Goldberg, 2011), the difficulty of reaching consensus on the meaning and significance of neuroimaging in the courtroom (Aggarwal, 2009; Glannon, 2014), as well as major ethical concerns with respect to early detection and prevention of violent and antisocial behaviour such as labelling and stigmatization, societal control, and undue focus on risk (compare Wolpe, 2013; Horstkötter, Berghmans, & de Wert, 2014).

Regarding potential novel treatment options for violence and antisocial behaviour, several participants stressed the need to obtain informed consent when offering biomedical interventions. The legitimacy of coercion and of mandating biomedical interventions has recently been a topic of discussion in the ethical literature as well. Interventions involving coerced or semi-coerced drug and/or hormonal treatments that may involve very serious side effects and affect an individual's mental liberty are considered ethically problematic (Bublitz & Merkel, 2014; Focquaert, 2014). However, some philosophers have recently argued that an offender's mental liberty or 'freedom of thought' is potentially equally violated by forced incarceration practices as by forced biomedical interventions (e.g., Douglas, 2014a; S. Carter, 2016; Petersen & Kragh, 2017). In our view, important differences exist between forcing a biomedical intervention upon an offender and depriving an offender of the right to free movement. Even though current incarceration practices are often ethically problematic, incarceration does not violate an offender's *mental* liberty in the same way forced neuro-interventions can violate it, because biomedical interventions are more likely to bypass our capacity to reflect upon the changes they bring about, and can overrule the ability to gradually endorse, reject or object to the alterations of our self (Focquaert & Schermer, 2015; Focquaert, 2017).

Most direct biomedical interventions (e.g. pharmacological interventions, transcranial magnetic stimulation, transcranial direct current stimulation) exert temporary effects and are essentially reversible in nature. This means that the beneficial treatment effects (e.g., symptom reduction in case of aggressive thoughts and behaviour) only last as long as one takes the medication or gets the stimulation. Nevertheless, it is not unlikely that the long-term use of biomedical interventions can have irreversible effects on one's cognitive-emotional functioning, especially when the intervention is started and an early age (e.g., the use of Ritalin for attention deficit hyperactivity disorder (ADHD)). Non-invasive, indirect biomedical interventions such as environmental enrichment and food supplements on the other hand supposedly have lasting, positive

effects on children's brain development. Although more invasive interventions such as deep brain stimulation are considered to have the potential to permanently rewire the brain after long-term use as well (e.g. in case of obsessive compulsive disorder (OCD)), such findings have not been reported.

With respect to the use of neuroimaging data in court, participants in our study were particularly concerned about mis- and overinterpretation and potential detrimental effects on legal processes. Similar worries are voiced in a recent paper on the use of neuroscientific evidence in Canadian criminal proceedings (Chandler, 2016). For example, an experimental study by Aspinwall and colleagues showed that judges that are confronted with neuroimaging data tend to impose lower sentences and list more mitigating factors (Aspinwall et al., 2012). A recent empirical analysis examining the use of neurological and behavioural genetic evidence in US criminal law (between 2005 and 2012) reports that neurobiological data was most commonly used by criminal defence attorneys to mitigate responsibility and punishment (Farahany, 2016). At the same time however, this study showed that prosecutors have argued for higher sentences referring to 'hard-wired' incorrigibility or future dangerousness of perpetrators, confirming the idea that neuroimaging data can function as a double-edged sword (Aspinwall et al., 2012; Chandler, 2016).

A significant added value of our study is that it shows forensic practitioners to be aware and cautious of various misunderstandings and misrepresentations that might result from the specific dynamics of the forensic field. These misinterpretations may arise, first, from the ways in which scientific findings are translated to forensic psychiatric practice. Second, they may arise when forensic practitioners interact with other legal (judges, jurors, etc.) and political (e.g. Department of Health or Justice) professionals who operate in different institutional frameworks (law, medicine, politics), and who may lack the expertise to correctly interpret biomedical interventions or findings. Third and finally, in confrontation with wider society, forensic practitioners are concerned about how biomedical interventions will be perceived by the general public and the media (Berryessa et al., 2016). Of particular concern is the risk that biomarkers for future violent or antisocial behaviour will reinforce perceptions of children as at risk, or risky[17] before they have actually engaged in harmful behaviour, and the substantive

17 "Two senses of risk are brought into alignment. The first is the desire to identify risky individuals – that is to say, those who will present a future risk to others – before the actual harm is committed. The second is the hope that one might be able to identify individuals at risk – those whose particular combination of biology and life history makes them themselves susceptible to some future condition – here personality disorder, impulsivity, aggressivity, or whatever, but more generally susceptibility for any psychiatric disorder" (Rose & Abi-Rached, 2013, p. 197).

negative effects of labelling and stigmatization (Pieri & Levitt, 2008; Levitt & Pieri, 2009; Rocque et al., 2012; Wasserman, 2014a; Chhangur et al., 2015; Horstkötter, 2015).

Notably, a range of ethical concerns voiced by forensic practitioners in this study may not be exclusive to bio-interventions. Some participants even explicitly mentioned that many of their moral concerns apply to some non-biomedically informed interventions as well:

> I think dilemmas of this type apply equally to behavioural therapy for example. (FP8)

Likewise, in ethical analyses regarding early detection and prevention programs, the argument has been made that many (but not all) objections voiced against biomedical approaches apply to psychosocial approaches as well, and that therefore, from an ethical point of view, "it is more important to determine how to deal responsibly with possible risks of early detection and prevention than asking whether this is based on a social scientific, a psychological, a biological or a mixed approach" (Horstkötter, Berghmans, De Ruiter, et al., 2012, p. 295; Horstkötter, Berghmans, & de Wert, 2014). On the other hand, our interviews suggest that forensic practitioners think that some issues may nevertheless be more relevant to bio-interventions (notably concerns about invasiveness, mental freedom, and irreversible or long-term side-effects).

In conclusion, forensic practitioners mostly appear to endorse and reason based on bio-psycho-social models of violent and antisocial behaviour. They stress that biomedical approaches will not (or should not) make current psychosocial approaches obsolete and that they should work in concert instead. This is consistent with current scientific literature in which there is a growing consensus that bio-psycho-social explanatory models of violent and antisocial behaviour are the most promising ones (Eichelberger & Barnes, 2015; Lee, 2015). Yet, our interviews show that forensic practitioners are also concerned about and acutely aware of the fact that the integration of neurobiological and behavioural genetic elements in explanatory models of violence and antisocial behaviour may be misinterpreted in various ways and by various actors, especially when implemented in the forensic field.

Limitations and suggestions for future research

Because the study is qualitative and the sample size is limited, we cannot draw conclusions about the minor differences we observed between psychiatrists and psycholo-

gists in terms of their general openness towards potential biomedical interventions. Whereas the majority of forensic psychiatrists demonstrated a basic openness to these developments, we observed a more critical attitude among psychologists. Drawing a comparison between these groups was not a central study aim, as our focus was on the expectations and moral views regarding potential biomedical interventions of forensic practitioners generally. The differences we observed reflect and are consistent with the various disciplinary backgrounds of forensic practitioners, i.e. medicine in the case of forensic psychiatrists and clinical and social psychology and criminology in the case of forensic psychologists and therapists, as well as the slightly different institutional contexts they tend to work in. Future research might explicitly focus on comparing expectations and moral views of different forensic practitioners working in different institutional settings, and in addition might take into account differences with respect to institutional architectures of forensic care in for example European countries compared to the United States.

5

Forensic practitioners' views on stimulating moral development and moral growth in forensic psychiatric care

Specker, J., Focquaert, F., Sterckx, S., & Schermer, M.H.N. (2018). *Neuroethics*, 10.1007/s12152-018-9363-x.

Abstract

In the context of debates on (forensic) psychiatry issues pertaining to moral dimensions of (forensic) psychiatric health care are frequently discussed. These debates invite reflection on the question whether forensic practitioners have a role in stimulating patients' moral development and moral growth in the context of forensic psychiatric and psychological treatment and care. We conducted a qualitative study to examine to what extent forensic practitioners consider moral development and moral growth to be a part of their current professional practices and to what extent they think that stimulating moral development is a legitimate objective in the context of forensic psychiatric treatment. In addition, we asked how forensic practitioners balance pubic safety and risk management concerns with the interests and wellbeing of the individual patient. We conclude that: (i) elements of moral development and moral growth in forensic psychiatric care practices are to a certain extent inevitable and not necessarily questionable or undesirable; (ii) yet, as in similar debates these elements need to be made explicit in order to discuss the accompanying ethical challenges and boundaries. An open academic, professional and public debate on aspects of stimulating moral betterment within current practices is therefore desirable.

Introduction

A number of separate debates invite reflection on the question whether forensic practitioners have a role in stimulating patients' moral development and moral growth in the context of forensic psychiatric and psychological treatment and care.

Psychiatrist Sean Spence has raised the question whether moral improvement (in the sense of being a better person, or a better *behaving* person) is an implicit, or even explicit, goal of psychiatric treatment: "Can pharmacology help us enhance human morality? (...) I argue that we are already deploying certain medications in a way not totally dissimilar to the foregoing proposal: whenever humans knowingly use drugs as a means to improving their future conduct." (Spence, 2008, p. 179). Apart from the – what Spence calls – 'Promethean project' of "specifically designing drugs that target and increase a pro-social feeling and behaviour such as 'kindness'" (Spence, 2008, p. 179), treatment can have morally relevant side-effects or consequences. Spence discusses the example of "a man prone to psychosis, who can be violent when ill, takes his medication reliably, thereby reducing his risk to others)" (Spence, 2008, p. 179). Here, Spence argues, the well-being of others has improved as a direct result of pharmacological treatment of a mental disorder. A number of commentators have discussed potentially morally relevant "side-effects" of existing drugs that may (or already) have altering effects on moral decision making or on morally significant behaviour, and urge more research to able to better distinguish between desirable and less desirable effects (Levy et al., 2014a, 2014b). One example discussed by the authors concerns selective serotonin reuptake inhibitors (SSRIs) that are prescribed to treat depression and anxiety disorders, but as a possible side-effect may increase aversion to directly causing harm in others (Crockett et al., 2010b).

Psychiatrist Steve Pearce and philosopher Hanna Pickard make a similar point when they argue that psychiatric treatment can foster moral growth in various ways: "First, they can lead to the emergence of new moral motives and intentions. Second, they can lead to the acquisition or development of cognitive skills such as empathy, which are central planks of moral action. Third, they can enhance the ability to apply moral understanding and skills in particular circumstances" (Pearce & Pickard, 2009, p. 281). They take it as a given that interventions that can foster moral growth occur routinely within psychiatric settings, most notably in the treatment of personality disorders. In this context the question is posed as to whether forensic psychiatric disorders should partly be understood as moral disorders, and forensic psychiatric treatment as moral therapy. Diagnostic criteria for personality disorders involve traits that involve failings of morality or virtue, such as lack of empathy in the case of narcissistic personality

disorder, or anger and impulsivity in the case of borderline personality disorder (Pearce & Pickard, 2009; Pickard, 2011). Pickard has discussed this in terms of the inherent 'Janus-faced nature' of personality disorders (PD): "The fact that the characteristics and traits that cause distress and impairment to the individual often involve harm to others. (…) Although harm to others, broadly conceived, is not part of the DSM-IV-TR definition of PD, it is part of how particular kinds of PD are diagnosed: via characteristics or traits that count as failures of morality or virtue and thus impair social, occupational, or other areas of interpersonal functioning." (Pickard, 2011, pp. 182-183).[18]

In sum, in the context of debates on (forensic) psychiatry issues pertaining to moral dimensions of (forensic) psychiatric care are frequently discussed. Although some experts have argued that moral betterment is or should be a goal within forensic psychiatry and psychology practices (Pearce & Pickard, 2009), it is unclear to what extent stimulating moral development and moral growth is a goal within current forensic mental health settings and much less so whether it should be.

In this article, we explicitly focus on questions related to the moral dimensions of forensic psychiatric practice. The main objective of this study is to explore the question whether forensic practitioners consider stimulating moral development and moral growth to be a part of their current professional practices, and to what extent they think that stimulating moral development is a legitimate objective in the context of forensic psychiatric treatment. In addition, we ask how forensic practitioners balance public safety and risk management concerns with the interests and wellbeing of the individual patient. In the discussion, we discuss whether, and of so in what ways, our findings relate to and can be informative for the bioethical debate on moral bioenhancement.

Methods

Sample and recruitment

We recruited 21 forensic practitioners (forensic psychiatrists, clinical psychologists and therapists) in The Netherlands and in Belgium. Subjects were recruited via professional organizations and by snowball sampling, meaning that initial research subjects

18 See also: "wrongfulness-laden disorders should be investigated to determine whether the disorder involves a moral incapacity (a disability in the moral sphere or "faculty") or is simply a matter of wrongful moral choice" (Sadler, 2014, p. 167); and "some psychopaths do, in fact, appear to have deficits that distinguish them from responsible offenders. These deficits appear to undermine psychopaths' ability to understand morality" (Shaw, 2016).

suggested potential future subjects from their network (Atkinson & Flint, 2004). Our sample consists of nine females and 13 males, ranging in age from 32 to 68 years. At the time of the interviews, 12 participants were employed in The Netherlands and nine were employed in Belgium.

We conducted 11 interviews with forensic psychiatrists (FP) (at the time of the interview, one participant worked as a general psychiatrist, but had worked in forensic settings in the past). We conducted 10 interviews with clinical psychologists (CP) or therapists (T) (at the time of the interview, two participants - a therapist and a clinical psychologist - were primarily involved in research and did not consult patients, but had done so in the past). Twelve participants (seven psychiatrists and five psychologists) are involved in scientific research, alongside their clinical or therapeutic work.

Qualitative interviews

Participants took part in an individual semi-structured interview lasting approximately one hour. During one interview, two respondents were present and interviewed together. The interviews were held in Belgium and The Netherlands, and took place between January 2014 and July 2016. The interview guide was developed by JS in consultation with MS, FF and SS. The interviews were conducted by JS, FF and MS. JS attended 17 interviews, FF attended seven interviews, and MS attended three interviews.

The interview schedule included open-ended questions about the moral dimensions of forensic psychiatric practice, about participants' views on the question whether they consider stimulating "moral improvement" or "moral development" part of their current work practise and as a legitimate part of their professional responsibilities, and about how to balance and prioritize public safety and risk management concerns with the interests and wellbeing of patients. The interview schedule also contained a separate part with questions on forensic practitioners' expectations and moral views regarding potential applications of current neurobiological and behavioural genetic research aiming to understand (and possibly help prevent, contain, or treat) violent and antisocial behaviour. We have reported on that topic elsewhere (Specker et al., 2018).

Coding

All interviews were transcribed verbatim and coded in QSR NVivo version 11, using descriptive theme analysis (Bazeley & Jackson, 2013). All transcripts were independently read by all members of the research team (JS, FF, MS, SS). All transcripts were independently read by all members of the research team (JS, FF, MS, SS). JS and FF discussed a random selection of transcripts with the purpose of drafting a preliminary

analytic framework. JS independently coded the transcripts by labelling sections and text units referring to one or multiple concepts relevant for the study purpose. An iterative approach was used in which new data that challenged the existing coding structure were used to revise the themes until no new themes emerged. Interpretative bias of data was avoided by means of investigator triangulation, which involved all researchers (JS, FF, MS, SS) checking the codes for consistency.

Results

Do stimulating moral development and moral growth play a role in treatment?

The first set of questions offered to participants raised the – deliberately broadly formulated – issue of whether forensic psychiatric treatment and care involve, in one way or another, elements of stimulating moral development and moral growth. In their responses, participants did not only differ in their opinion on whether these elements should or shouldn't be part of treatment, but also in their understanding of what morality entails. Nearly all participants started their response with discussing how to understand moral development and moral growth, and *what kind of* morally relevant aspects are, or potentially can be, targeted in treatment. Before outlining the different aspects of morality participants mentioned in the subsequent section, below we discuss the reasons participants offered why they do or do not think stimulating moral development and moral growth is part of treatment.

Whereas only a few participants indicated that stimulating moral development and moral growth were not part of treatment, most participants appeared to be more ambivalent in their answers. Participants who indicated that stimulating moral development and moral growth play no role in treatment, mentioned that their treatment plans do not involve aspects of stimulating moral development, and that their medical training did not involve a focus on the moral aspects of behaviour. Instead, they underlined their medical rather than 'moral' expertise. This can be illustrated with the following quote from a participant:

> Look, our task is not to create 'better people'. We want them to stop doing awful things, we want to lower the risk factors, and I think that, to create better people, that is a very big step. CP8

> If someone says to me; 'Generally, I'm quickly aroused, high in blood so to say (…) And then I feel rejected very soon', if I can improve that in any way, in how he relates to higher values and the world surrounding him, by intervening by giving him

> a beta blocker, for example, to make sure he is less quickly aroused – I will do that. But my goal is to increase his quality of life, my goal is not to improve someone's morality. Because, I actually think that that does not belong to my expertise, to my profession, and is actually not part of my assignment. FP1

A number of participants argued that stimulating moral development or moral growth should not be part of treatment, and emphasized the importance of maintaining a clinical stance towards their patients, if only to provide a safe place to discuss sensitive subjects. They argued that the primary task of a forensic psychiatrists or therapist should be to treat disorders and to improve the quality of life of their patients – not to *moralize* – and that moral condemnation and judgment, if applicable, should happen not in the consulting room, but elsewhere (in court for example, or perhaps in society at large).

> I do not like the idea that this would be a required task of a forensic psychiatrist. I think, I can only speak for myself, I think, well, we are not to judge about good and evil. I mean, we can only observe. And the only thing we are trained to do is to see if we can find a way to improve the quality of life of patients, preferably in a holistic way. And that is the only thing we can do, anything else, we cannot. FP1

> Is it the aim for psychologists to become priests? To become moralists? Please, no. There must be a place for someone, who of course is condemned everywhere else within society, to find shelter and to not be judged. If such a place is no longer available (…) that person will no longer dare to share her most immoral thoughts. CP6

> And by the way, who am I to lecture that person? Because presumably they would be able to mention a few things they disapprove of about me, right? PF5

However, other participants indicated that moral development is indeed part of forensic psychiatric treatment. These participants often mentioned that they considered improving patients' capacities for empathy (both cognitive and affective) and moral reasoning (in terms of correcting cognitions and logical errors) as explicit treatment goals.

> Empathy, for example, is certainly a goal for us. We focus on impulse control, relapse prevention, empathy enhancement, and responsabilization. CP4

> Very often with people who do not behave morally, I feel it is about logical errors. And then I try – but I am necessarily limited in this regard – I try to determine

> whether there are any thinking errors involved, and whether I can test their flawed ways of thinking, and possibly correct or adjust them, pharmacologically or psychotherapeutically. FP1

Interestingly, as the interviews progressed, several participants kept coming back to this subject and wondered whether, even though moral development is not an explicit treatment goal, this might be an implicit part of forensic treatment:

> However, as far as I am concerned, not to change him as a person, no. In the sense of trying to impose a certain kind of moral awareness, no.
> Interviewer: That is not one of the goals of treatment?
> Interviewee: No. Not explicitly, and perhaps also not implicitly, but I'm not entirely sure about that. FP8

In this context, several participants referred to the inherently normsetting and prescriptive nature of their profession and the challenge of not placing one's own moral convictions and moral values at center. Participants thus appeared hesitant to moralize, but at the same time discussed that to a certain extent, this might also be inevitable:

> Well, at least I think many psychiatrists, unknowingly, very much approach and also treat their patients on the basis of a certain moral idea, that is, with their own norms, values, and morality, which simply pervades everything you're trying to convey to your patients. So implicitly I would certainly agree. I think explicitly, also - many people, hm… Well, to provide a very concrete example, we offer a training that consists of three parts: social skills, emotion regulation, and moral reasoning. Moral reasoning is about casuistry: 'What does this mean for the other?; What do you do merely for your own advantage?; What if this would happen to you?'. So in training, for example, we do call it, 'to learn to reason better morally'. So apparently we have some kind of idea about what good moral reasoning is, and apparently it is also something we want to teach. FP10

> Coming from general psychiatry, it does indeed strike me that, the, hm, the moral framework is implicitly present – much more than outside of forensic psychiatry. It is not made explicit, but it does play a role. If you would put it bluntly: 'To what extent do we want the people whom we are treating here to be good citizens?' Yes, I'm afraid that it does, that it does play a role beneath the surface, but that we do not talk about it. On a superficial level, we aim to make sure that people no longer pose a risk, or as little risk as possible to themselves or others. But of course that

> has a very strong moral component. So, yes: it plays a big role. And no: it is not expressed as such. FP8

> So it is not my job to socialize people. Although somehow it is, but I will never say this out loud, because otherwise people will interpret socializing as re-educating, in the sense of 'becoming like us'. I embrace a socialization that takes place from within a subject's own coordinates. If someone regains a place within society – without necessarily actively participating in society, but also without wandering and suffering; if someone is able to make life bearable for himself/oneself, in a very discrete manner, without experiencing others as threatening and so forth – that for me is already a successful socialization. Whether others will share that perspective? The prevailing norms of others claiming that a normal individual should be like this or like that. I don't care about such norms. As long as that person no longer poses a physical threat to others or to himself, that's okay for me. Regardless of what that person is like at that time. CP6

Participants describe how particular patients and types of offences can elicit moral outrage or even abhorrence, and stress the importance of a clinical stance or attitude in order to overcome or distance themselves from these negative emotions. Some reflect on the ways their profession has forced them to reflect on their own moral framework and commitments.

> And of course, I experience these thoughts as well: "Come here, you boor, and I will beat you up". Apparently, that is part of us as human beings. But then I realize that this would satisfy my own frustration more than anything else. FP3

> That subcategory evokes repugnance in almost everyone. And the difficult thing is – and that is true for medicine generally of course – that we are trying to disconnect this from the disorder. So we see pedosexual offenders primarily as people with a problem, with a disorder that we should help them to get rid of as much as possible. And in interactions with patients, the moral dimension is not addressed, right? So you never say to someone; "What a horrible thing you have done!" FP8

What aspects of human morality do forensic practitioners deem relevant for treatment?

In their responses, participants identified and reflected on various aspects of human morality that are, or potentially can be, targeted by treatment: patients' remorse, conscience, or guilt; self-regulation and self-awareness; motivation and will to change; moral responsibility; capacities for moral reasoning; and moral emotions (such as em-

pathy). Participants differed in the way they conceptualized morality, and in what they understand morality to be. Whereas some focused more on capacities to be moral (such as having the capacity to empathize with others, or to reflect on one's own behaviour), others focused on more symbolic elements, such as restoration with society.

> With respect to expressing regret and restoration and such things: a chance of recovery, a bond with the victim, restoring the bond with his own family (because they are affected as well), restoration with society – these are things we certainly address. CP4

> Interviewer: "So morality is not addressed at all in treatment?"
> Interviewee: "Yes, it is, but not in terms of morality, but in terms of reciprocity. To the extent that one can build a reciprocal relationship with someone and is capable to handle and sustain that relationship and to anticipate the other's position, and best-case scenario even to mentalize it. And that he is able to take the position of the other." FP11

One of the aspects that participants deemed especcially relevant in the context of treatment and that was mentioned frequently concerns self-awareness and the capacity and will to control oneself.

> You teach people to analyze themselves. You offer a kind of frame to pay attention to what they think, feel, and do. Also in the case of sexual offenders. To make them aware of the kinds of things they tell themselves when they start to commit and continue to engage in an offence, what they tell themselves while they are doing it, and what they tell themselves afterwards, to be able to say that it wasn't that bad. And so on. So we actually give people such grids and tools to get to know themselves better and to pay attention to feelings and thoughts and actions that they did not pay attention to before, and without being fully aware, they would proceed to commit a crime. FP2

> Really, it is inhibition that is actually our core business. You try to teach these people to keep that under control. CP7

> To get a chance to, well, create a motivation, the will to control oneself. CP7

Many participants raised the question of whether focussing on one aspect or capacity would accomplish a genuine improvement in the sense of someone being 'a genuinely better person'. For example, several participants reflected on whether it is enough for

someone to stop a particular behaviour, for example by enhancing inhibition, without accompanying changes in beliefs or thought patterns.

> If someone says, 'I will not perform that behaviour anymore,' then you could say that, because of that, you have become a better person, right? CP7

> In the core? I don't know. It is also possible, you may also have a different motive to stop doing it. To prevent relapse, in your own interest. CP8

> I will quote Freud here: We are all rapists and thugs in the depths of our thoughts, but the bad ones are those who act on them, and the good ones are those who think about it, but don't act. FP9

Several participants discussed responsibility as an important part of forensic psychiatric treatment. Both in the sense of looking back (*I* was the one who did these things) and in the sense of looking forward (*I* need to make changes in order to prevent myself from doing the same thing again). Several practitioners stressed that, frequently, the first is needed to achieve the second:

> As long as patients say 'I could not do anything about it', I will tell them: 'Well, yes, if you really couldn't, if you really feel that it was because the sun was shining or it was raining, you could not do anything about it, that it was the weather; well, then you cannot go outside, can you? That must be terrible; it could happen to you tomorrow again, couldn't it?' Well, of course they do not think along those lines. (…) But when framed like that, that's not what they want for themselves. So we need to address what's possible. 'Well, then we do have to figure out what you can do about it. For all I care, you bring both your umbrella and your sunglasses, to make sure that you… But you must address it.' T2

> Even very seriously disordered people are, at a certain level, accountable. And that also makes it possible to achieve progress with them, do you understand? That's the space you need of course, because if you have the extreme, 'I cannot do anything about it' – yes, and then what? FP5

> That has to do with giving responsibility. Because it is you who makes that choice, despite the feelings you may have; you are the one who makes the choice to act. CP8

Nevertheless, several practitioners expressed reservations with respect to the importance of addressing responsibility in treatment. For example, because behaviour change

is far more difficult to achieve than was once thought, and this places limits on the degree to which moral development can be addressed in treatment. Many participants discussed the degree to which you can hold people accountable for past behaviour, for example because of societal and situational factors, or expressed a more general skepticism with respect to human free will.

> I have been working in this field for thirty years of course, the forensic field. And I started at a time when we were thinking very much about 'Malleable Man'. A period in which self-regulation and free will were values that we were holding dear. And of course, over the course of time, that optimism has diminished, with all the consequences that this entails. So we were thinking, when I started more than thirty years ago, that as long as people would be willing, and we would be motivating and stimulating them, they would move in the right direction. And now with developments, also neurobiological developments, you think that there is more to it than simply the idea: as long as you want to, you will succeed. CP1

> My belief in free will is limited. If you observe those boys – and I've really seen hundreds, also intensively – they are almost all friendly fools who fell victim to their own life, their own environment, their upbringing, their lack of intelligence, and so on. FP3

> I would say: I take human free will as a starting point. That is, ultimately, a hypothesis, a subjective truth, yes. (…) But also for the court it is a basic starting point: it's assumed that people are responsible for what they do, what they think, and so on, for their actions, until the opposite is proven. FP2

How do forensic practitioners' balance the wellbeing of the patient with public safety concerns?

Why would stimulating moral development and moral growth in fact be part of forensic treatment? Many participants stressed that their primary objective is to lower the risk that someone will harm others.

> That is a problem we have with sexual delinquents: that the bodily integrity of others is potentially in danger. (…) If I treat a serial rapist, I cannot say: 'He relapsed, but that is already an improvement, because nothing happened in the three months before.' FP2

> But within this profession, I always say to patients: 'You can remain as crazy as you are now, I am not saying that you have to change at all. I just need to change this one thing, that is, that you will never do it again.' T2

In contrast, several participants indicated that stimulating moral development may be part of treatment, when it can help to manage stress relief or to reduce the suffering of the patient. Several participants conceptualized this as an 'egocentric perspective', in that they try to refer to the patient's interests.

> Moral outrage about paedophiles etcera that's, that's very intense in society. Society demands that we do something about it. But the moment I… Opportunities to work with these people will not grow the moment I start talking about morality. Perhaps when I talk with them about empathizing with victims – maybe you could classify that under that heading? Which is of course part of the treatment of sexual offences. To empathize. But you empathize with the other, in order to enhance your own inhibition. It is not about feeling sorry for those people – do you understand that? The victim, that is merely, that is actually only just, actually only just instrumental for the patient himself. The more you empathize, the more the resistance grows, the resistance to act on it. FP5

> But what I can say is that if someone, because of his moral deficiencies so to say, gets in to a lot of trouble with his environment, and if he is rejected a lot, and because of that is acting very hostile, and so on – I will point out that mechanism to him. And I would say to him, 'I would advise you to do some tests, to take a look at what we can do, maybe that will help.' Yes, that I will do. FP1

Some participants characterized stimulating specific morally relevant aspects as a means rather than a goal; moral development may help achieve some other goal of forensic psychiatric treatment (such as lowering recidivism), but it is not an end in itself.

> "I would say, it is a collateral advantage" CP1

These different potential objectives of stimulating moral development and moral growth – a focus on safety and harm reduction to protect others versus a focus on the patients' wellbeing and treatment goals - are mirrored in two different sets of professional roles and responsibilities of forensic practitioners: on the one hand their medical background as doctors, and on the other hand their responsibilities regarding

public safety. We asked participants in our study to reflect on these roles and potential tensions between them, and to indicate which of the two, if any, they consider primary.

In general, most participants acknowledged both responsibilities. They differed however, in their views on which of these professional roles they considered primary. Whereas some participants explicitly positioned themselves as a medical doctor first, most participants also stressed their responsibilities in preventing harmful crimes from being performed again.

> If I would have to choose, I would be inclined to favor the protection of society, because the civil commitment of one disturbed forensic patient can prevent the victimization of several victims. FP9

> If we would make a list of the ten things we are doing here, that would be number one: No new victims. And this is also clearly defined in terms of professional secrecy. We have professional secrecy pertaining to all, everything that is discussed here, until we estimate that there is a real danger with an identifiable future victim. CP4

> I think the task of forensic psychiatry is, primarily, to minimize recidivism. That's really primary, because that makes the profession what it is. That doesn't mean I am blind to people's suffering of course, but that is primary, that is absolutely paramount. And then, hm, I would say, secondly, can I, can we maybe, make people suffer less, have fewer problems, improve their quality of life. Also for their environment, I think that is often forgotten; for the children and for family members, that is very important. (…) That is the system within which we operate, and that also allows the patient a certain degree of autonomy. FP5

Several participants discussed various tensions between, on the one hand, their medical responsibilities, and, on the other hand, public safety concerns.

> Sometimes these people experience profound suffering. Sometimes there is no suffering. Those are fundamentally different situations. (…) To put it bluntly, someone who does want help in preventing making the same mistakes again, and someone who refuses that help – you do have different options available. FP6

> You must adhere to the rules of medicine. And that is a danger, I think, for forensic psychiatrists. Actually, that is a danger in many disciplines in which you specialize, that you have to think carefully where you came from, where your foundations lie, to not stray from one's subject field. (…) Because a forensic psychiatrist is first and

> foremost a doctor. And must also work from those foundations, and according to the oath and principles of proportionality and subsidiarity. FP6

Some participants offered pragmatic rather than principled arguments for not focusing too much on future risk in treatment:

> Yes, both of course. But when it comes to initiating and achieving successful forensic treatment, I don't think that the focus should be on that risk. Because if you want to motivate people for their own treatment, because that is necessary for treatment success, you have to start from their own suffering. And sometimes, that is a different suffering than how society sees it, but that needs to be the starting point for treatment. Because otherwise, you will not have any commitment of your clients. CP2

Several participants drew parallels with regular psychiatry, where their medical expertise and authority solely function within a therapeutic care setting, and forensic psychiatry, where their medical expertise and prognosis become embedded within a legal framework, and non-medical or non-therapeutic considerations come into play. Participants also discussed different settings forensic practitioners can work in, ranging from outpatient care, to providing mental health care in prison, to specialized long term residental secure care, and how these different settings influence the degree to which they are able to assert their medical authority.

> Perhaps that is specific to forensic psychiatry, that this power [to extend imprisonment] does not come to lie with you, but that you are able to function within a kind of triangular relationship. But that also entails that you must be able to tolerate that someone else is watching along. And that is different compared to a dialogue in regular psychiatry. Perhaps therein lays the uniqueness of forensic psychiatry. Interviewer: In this third factor? Interviewee: Yes. FP11

> I have also worked in prison, there you have nothing to say. That is a prison, it is the warden who calls the shots. It is not a medically protected domain, with a healthcare logic. So we, as healthcare professionals, can build in safety – but it has to be on my own territory. FP2

Discussion

Forensic practitioners' views on potential moral dimensions of forensic psychiatric treatment and care are highly diverse, as these interviews show. Whereas several practitioners rejected the idea that stimulating moral development or moral growth is or should be part of forensic psychiatric treatment, other practitioners appeared to be more open to reflecting on potential elements of stimulating moral development in their work practice. And although current forensic practices do not (explicitly) seek the moral development and moral growth of forensic patients, elements of stimulating moral development and moral growth might be part of forensic psychiatric care implicitly, as this study suggests. Yet, forensic psychiatric treatment is hardly ever discussed in those terms. As discussed in the discussion, forensic experts Pearce and Pickard argue that psychiatry is both a moral and a medical science, and that a convenient blindness to the moral content of psychiatry opens the door to potential abuse (Pearce & Pickard, 2009). They conclude that our best defence against abuse in forensic psychiatry is honesty and ever-vigilant self-reflection.

In general, the forensic practitioners we interviewed appear to be cautious about moralizing and imposing particular moral views and values, and often stress the importance of a professional, clinical stance to counter this. In line with this, Marga Reimer mentions that it is widely agreed that moral judgment should play no role in the practice of medicine due to its capacity to impair clinical judgements and especially so in the case of psychiatric conditions (Reimer, 2010).

Participants identified and discussed a range of morally relevant aspects that are or potentially can be addressed in the context of forensic psychiatric treatment, ranging from stimulating empathetic concern, improving cognitive skills and correcting cognitions, strengthening protective factors to prevent recidivism, to lowering risk factors for future problem behaviour. Future research might study in a more systematic manner whether there is a relation between what respondents understand morality to be with their views on the appropriateness of stimulating moral growth in treatment.

Participants mentioned different potential objectives for stimulating moral development and moral growth in treatment: to treat mental disorders and alleviate suffering of their patient, and/or to reduce the risk of reoffending and prevent future harm to others. Our study suggests that forensic practitioners are both security-oriented (in terms of risk reduction and recidivism prevention) *and* concerned about patient care, with some individuals focusing more strongly on the care aspect and others more strongly on the security aspect. Several participants discussed the importance of maintaining

a clinical stance and relying primarily on their medical expertise and patient-centred responsibilities, although most practitioners also discussed their role in promoting public safety, as well as potential tensions between these two responsibilities.

Professionals working in forensic psychiatric mental health care are said to indeed have diverse, and potentially conflicting, roles and duties, as they need to balance responsibilities towards patients (individual offenders), towards the legal system, and towards broader society (Day & Casey, 2009). Yet, this study also indicates that a clear code of ethics on how to manage potential tensions between promoting public safety on the one hand and the wellbeing of individual offenders on the other hand is largely lacking.[19]

Professionals may encounter a range of ethical conflicts between these two roles or sets of tasks, often discussed in terms of a 'dual role', 'dual relationship', or 'dual loyalty' dilemma (Robertson & Walter, 2008; Jörg et al., 2012; Ward, 2013). This dilemma is discussed, first and foremost, in the context of debates about potential conflicts between a psychiatrist's duties as 'healer/caretaker' and as 'evaluator', for example in the USA when forensic psychiatrists are involved in evaluations that may lead to administration of the death penalty (Robertson & Walter, 2008). In other legislations, for example in the UK, a similar conflict may occur when a forensic psychiatrist's evaluation of dangerousness may lead to a person's pre-emptive detention. Also in treatment contexts, individual forensic practitioners may face the ethical demands of two roles, one prioritizing the needs and interests of the community, the other the (medical and therapeutic) needs and interests of the offender (Ward, 2013).

Choice and consent, for example to consent to or refuse treatment, is particularly complex in a secure psychiatric care context, as Gwen Adhead and Teresa Davies discuss: "There is a sense in which the medication is fulfilling a penal role in reducing the risk of re-offending, in addition to the therapeutic role. Patients may not be allowed to refuse to take medication if professionals think that taking medication will reduce their risk" (Adshead & Davies, 2016, p. 78). Sex offender therapy might serve as an example here. According to forensic psychiatrist Bill Glaser, sex offender therapy should be characterized as *treatment-as-punishment* rather than *treatment of the punished*, for the reason that "this type of therapy does not have the interests of the offender

19 "If there are different ethical codes or systems of norms available to guide offender assessment and treatment, it could be hard to agree on a subsequent course of action. One forensic expert might justify his or her actions by appealing to obligations to the court while another could refer to the needs of patients or offenders, and an obligation to ease suffering whenever possible. The problem of ethical *incommensurability* raises its head here." (Ward, 2013, p. 94)

as its primary focus" (Glaser, 2009, p. 251; 2010). Glaser advocates that in a treatment context, it should be made clear to offenders that "the goals of treatment not always coincide with their own interests: (Glaser, 2009) as this satisfies the requirements of both minimizing distress (caused by otherwise deceitful disguising of the true purpose of treatment) and promoting equality (by providing offenders with the same amount of knowledge regarding treatment goals as that possessed by therapists)" (Glaser, 2010, p. 267). Although evidence about the effectiveness of pharmacological agents in treating sex offenders is inconclusive (Långström et al., 2013; Khan et al., 2015), Daniel Turner and colleagues nevertheless discuss that "clinical experience suggests that, for some paraphilic patients, medication is a useful addition to psychotherapeutic interventions and, as such, its use is being recommended by both clinicians and the WFSBP [World Federation of Societies of Biological Psychiatry] guidelines" (Thibaut et al., 2010; Turner et al., 2017).

Adshead and Davies also argue that forensic patients should be included in the medical decision-making as much as possible to avoid feelings of humiliation, despair, emotional isolation and stigmatization and to stimulate the long-term recovery of patients (Adshead & Davies, 2016). Forensic practitioners and forensic patients need to be able to rely on each other for support and safety. This may include that forensic psychiatrists at times where patients are unable to make fully competent decision support patients in the decision-making process to reinstate full autonomous decision-making on behalf of the patient and achieve maximal long-term rehabilitation. Liégois and Eneman similarly argue that shared decision-making should always be the desired goal within a psychiatric context (Liégeois & Eneman, 2008). Coercion should never be self-evident and should always be normatively defended.

The diversity of ideological and theoretical justifications of penal strategies and criminal justice institutional frameworks worldwide (more focused on rehabilitation versus more focused on retribution) arguably reflect these same tensions.[20] Significant differences exist between forensic mental health systems globally (Dressing et al., 2007), yet literature on international comparisons of forensic psychiatric care is scarce (Ogloff et al., 2000). In terms of legal demands, admission criteria, the concept of criminal responsibility, service provision and treatment philosophy, large differences exist, even between Western European countries (Dressing et al., 2007; Salize et al., 2007;

20 A recent comparison of forensic psychiatric care in England, Germany and The Netherlands confirms the presence of the dual role or dual relationship dilemma in Western European contexts: "Clearly, all three countries are in the process of significant challenges and changes in care provision reflecting the tensions between the two key values of forensic psychiatry: Care for the individual and protection of the public" (Edworthy et al., 2016, pp. 24-25).

Edworthy et al., 2016; Sampson et al., 2016). These differences between national legislations shape the particular ways in which forensic practitioners may experience dual role or dual relationship dilemmas. We urge for more awareness of the historical, ideological and political rationales behind particular institutional settings. If only because studies have shown that the context in which health care takes place, influences and potentially compromises the provision and ethics of health care (White et al., 2014).

Forensic psychiatrists need to be able to fulfill their therapeutic role without feeling pressured to give precedence to public safety in ways that harm or are likely to harm their patients. We agree that although forensic psychiatry "can contribute significantly to the protection of the public in individual cases, crime prevention cannot be its primary purpose. In a social climate that places increasing emphasis on the management of risk, the pressure to do so is substantial" (Buchanan & Grounds, 2011, p. 422). Forensic psychiatry as a medical discipline needs to be wary of attempts to use psychiatry as a means to impose the state's interests on the lives of offenders.

Several authors have linked the treatment of offenders, especially with neuro-biological interventions, to the debate on moral enhancement. In this debate, the main question is whether biomedical interventions that enhance prosocial tendencies and emotions and/or inhibit anti-social tendencies and emotions may – or should – be used to improve morality and moral conduct, in order to solve pressing societal problems such as crime and violence, or even terrorism and climate change. Commentators have discussed the use of neuro-interventions for offenders or forensic patients who are suffering from various cognitive, motivational and emotional impairments as examples of moral enhancement. As such impairments may involve risk factors for various kinds of immoral behaviour (e.g., sexual crimes, violence, racism), proponents argue that moral bioenhancement could provide new ways to achieve successful recidivism reduction and rehabilitation (Douglas, 2008; S. Carter, 2016).

Several commentators in this debate, including the present authors, have discussed whether psychiatric treatments that address neurobiological risk factors for deviant behaviour should indeed be understood as proper instances of moral enhancement (Specker et al., 2014; Reichlin, 2017; Specker & Schermer, 2017). Discussing the treatment of forensic mental health disorders in terms of the overall practice of moral enhancement might have undesirable consequences. One potential negative consequence of doing so might be that framing forensic mental health treatment as 'mere' moral enhancement could bring the public to disregard the seriousness of the mental health problems forensic patients may face (Focquaert & Raine, 2012). Even more problematic is the possibility that conceptualizing certain risky, invasive and non-voluntary forensic

interventions under the umbrella of moral enhancement could inadvertedly promote the acceptance of criminal justice practices that are ethically troubling. Examples of which would be coerced drug and/or hormonal treatments that may involve very serious side effects and/or affect an individual's mental liberty (Bublitz & Merkel, 2014).

Nevertheless, outspoken proponents such as Ingmar Persson and Julian Savulescu have argued that a number of psychiatric disorders can be characterized as "moral defects", and therefore, that treating these disorders should indeed be understood as moral enhancement:

> The opposite of promoting another's interests is damaging another's interests. Traits which increase harm to others cause immoral behaviour. The paradigm is psychopathic personality disorder, but other personality disorders such as antisocial personality disorders, borderline personality disorder and narcissistic personality disorder can cause great harm to those who come into contact with these individuals. The reduction in these tendencies are thus moral enhancements. (Savulescu & Persson, 2012, p. 410)

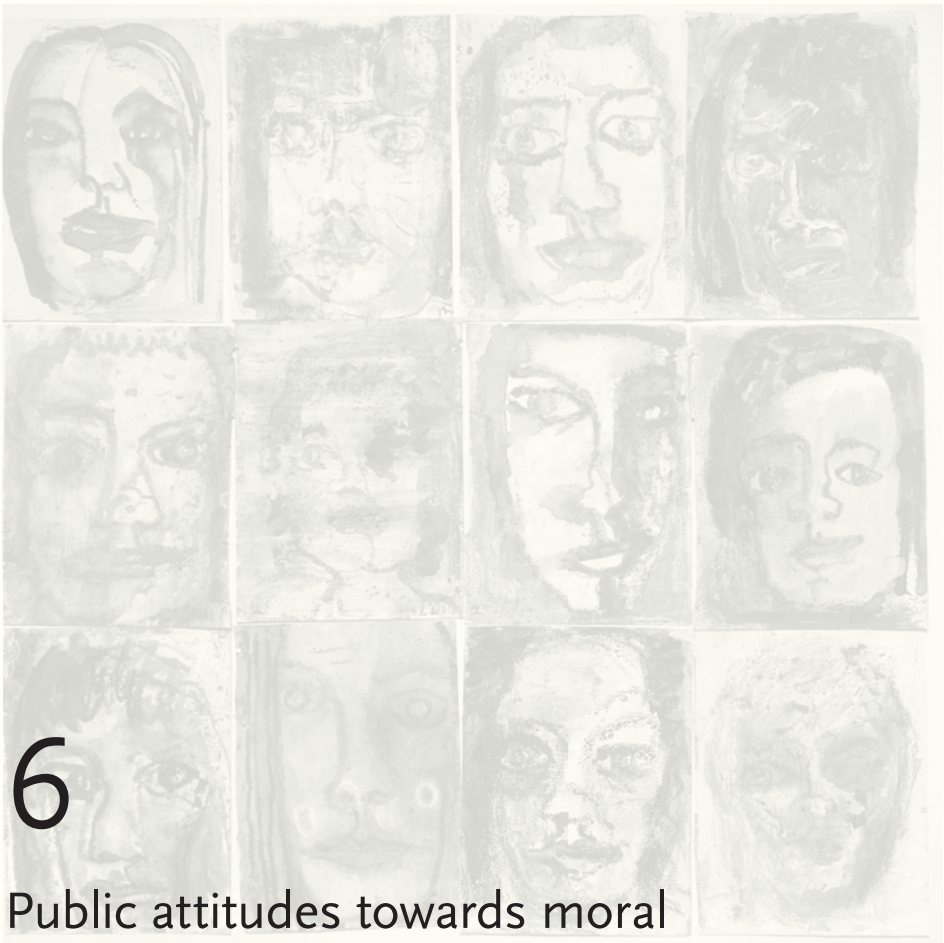
Likewise, David DeGrazia has characterized the treatment (or prevention) of antisocial personality disorder as a uncontroversial example of moral enhancement (DeGrazia, 2014), and Thomas Douglas has discussed "institutions of criminal justice'" as institutions that are arguably "already engaged in a kind of moral enhancement" (Douglas, 2014c, p. 1245).

A reason in favour of discussing certain aspects of forensic psychiatric care practices in the context of the debate on moral enhancement is therefore that it enables explicit debate on moral dimensions of forensic psychiatric care practices, and fosters professional dialogue and transparency. As Wiseman notes:

> if we are already getting moral enhancement by proxy, and this is to some extent inevitable, the best solution may be to drag the whole thing out into the open and critically inspect the process in the full light of day. If some forms of medical and mental health treatments will always have morally related aspects or societal judgments embedded within them, let us make these judgments explicit and attempt to find some way of integrating them within an acceptable code of practice – something which ensures that the therapeutic context is appropriately person-centered in nature and nonreductive, and that the healthcare professionals involved are appropriately directed and sufficiently well-armed against the dangers raised above (Wiseman, 2016, p. 219).

Moreover, the moral enhancement debate has proceeded without much attention for the specific institutional contexts in which potential moral enhancement interventions will be implemented. By exploring views of forensic practitioners on elements of moral development and moral growth in current practices, we hope to open up space for discussion about where and how 'moral enhancement' may – or may not – be brought into practice. Without adhering to the view that treatment of psychiatric disorders should be understood as moral enhancement, this exploration of views on potential moral dimensions of forensic psychiatric care can, in our view, inform the debate on moral enhancement.

In conclusion, we would submit that: (i) Elements of stimulating moral development and moral growth in forensic psychiatric care practices are to a certain extent inevitable and not necessarily questionable or undesirable; (ii) yet, as in similar debates, these elements need to be made explicit in order to discuss the accompanying ethical challenges and boundaries. The history of concepts like deviance and mental disorder has led to a wide array of "muddled concepts, systems, values, and priorities" within current psychiatry (Sadler, 2013). There is a need for philosophical reflection on the aims of criminal justice and how these relate to forensic psychiatric practices. How far should the authority of the legal system extend within forensic psychiatric practices and how should psychiatrists approach and deal with the ethical difficulties that are specific to their field? Without such reflections, forensic practitioners risk having to navigate a "moral minefield" (Sadler, 2013). Especially in view of the growing interest in neurobiological interventions, an open academic, professional and public debate on the (un)desirability of stimulating moral development and moral growth within current practices is therefore needed.



6

Public attitudes towards moral enhancement. Evidence that means matter morally

Specker, J., Schermer, M.H.N., & Reiner, P.B.
(2017). *Neuroethics, 10* (3): 405-417.

Abstract

To gain insight into the reasons that the public may have for endorsing or eschewing pharmacological moral enhancement for themselves or for others, we used empirical tools to explore public attitudes towards these issues. Participants (*N*= 293) from the United States were recruited via Amazon's Mechanical Turk and were randomly assigned to read one of several contrastive vignettes in which a 13-year-old child is described as bullying another student in school and then is offered an empathy-enhancing program. The empathy-enhancing program is described as either involving *taking a pill* or *playing a video game* on a daily basis for four weeks. In addition, participants were asked to imagine either their own child *bullying* another student at school, or their own child *being bullied* by another student. This resulted in a 2 x 2 between-subjects design. In an escalating series of morally challenging questions, we asked participants to rate their *overall* support for the program; whether they would support *requiring* participation; whether they would support requiring participation of children who are *at higher risk* to become bullies in the future; whether they would support requiring participation of *all children* or even *the entire population*; and whether they would be willing to participate in the program themselves. We found that people were significantly more troubled by pharmacological as opposed to non-pharmacological moral enhancement interventions. The results indicate that members of the public for the greater part oppose pharmacological moral bioenhancement, yet are open to non-biomedical means to attain moral enhancement.

Introduction

Moral competence is universally valued. Religious texts and the virtue ethics traditions all valorize the attainment of moral fluency. The Enlightenment brought its own contributions to the project, with deontology and consequentialism imparting 'rational' means of defining what it means to be a moral person (Kitcher, 2011).

Recently, a debate has emerged regarding the propriety of moral *bioenhancement* (Douglas, 2008; Ingmar Persson & Savulescu, 2012; Harris, 2016a; Wiseman, 2016). The suggestion is that we are on the cusp of understanding the neurological and genetic underpinnings of moral (and immoral) behaviour, and that we should use that knowledge to develop technologies that enhance human morality. However, what constitutes moral enhancement is highly contested (Raus et al., 2014; Beck, 2015): "clear and precise definitions of "moral enhancement" are not to be found; what has been called "moral" enhancement ranges from encouraging empathic concern to increasing personal responsibility all the way to heightening respect for global fairness" (Shook, 2012, p. 3). The debate is highly speculative. As the science of moral enhancement is "in its infancy", neuroscientist Molly Crockett has warned to "be careful not to draw premature conclusions about potential avenues for moral bioenhancement" (Crockett, 2014a, p. 370). Most contentious of all has been the suggestion that moral bioenhancement ought to be compulsory.[21] The debate has been vigorous but is at somewhat of an impasse (Harris, 2011; Douglas, 2013, 2014c; Harris, 2014; Hauskeller, 2014; Rakić, 2014b, 2014a; Sparrow, 2014a, 2014b; Agar, 2015b, 2015a; Ingmar Persson & Savulescu, 2015a; Rakić & Hughes, 2015; Bublitz, 2016; Harris, 2016b; Hauskeller, 2016; Ingmar Persson & Savulescu, 2016).

In earlier articles we therefore advocated a more focused debate on the potential domains for which moral bioenhancement interventions will most likely will be implemented (Specker et al., 2014; Specker & Schermer, 2017). Similarly, Harris Wiseman has advocated a 'practical-realities first' approach to potential moral bioenhancement interventions, implying that speculation about moral bioenhancement should account for "the specific practical realities to be found on the ground level, which are not at

21 In their first publication on moral enhancement, Ingmar Persson and Julian Savulescu famously argued: "If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them" (Ingmar Persson & Savulescu, 2008, p. 174), and in subsequent publications they have reaffirmed this position: "we do not rule out that moral bioenhancement could be justifiably imposed without the informed consent of the subjects" (Ingmar Persson & Savulescu, 2017).

all incidental but the very realities around which the abstractions of the debate must be made to shape themselves (not the other way around)" (Wiseman, 2016, p. 13). Given that the issue is of interest not just to philosophers but to the public at large, we explored public attitudes towards moral bioenhancement using both quantitative and mixed methods techniques.[22]

We carefully considered what sort of immoral behaviour would be best to evaluate in our studies. We previously found that the public was generally supportive of using pharmacological means to alter criminal behaviour so long as safety was assured (Berryessa et al., 2016), but empirical studies have found criminals to be outside of the circle of moral concern (Crimston et al., 2016). It is unclear whether members of the public are similarly supportive of biological interventions aimed at persons who engage in immoral yet legal behaviour. We therefore narrowed our focus to morally contentious behaviours that are not unlawful. We settled on bullying, an act that is generally condemned as immoral but usually does not cross the line to illegality.

Bullying is an act that is intended to harm, takes place repeatedly, and is characterized by a systematic abuse of the imbalance of power between the aggressor and target (Smith et al., 2002). Bullying takes place in schools, between siblings, in prisons, and in the workplace (Monks et al., 2009), as well as online (Dooley et al., 2009). The observation that school bullying (both perpetration and victimization) predicts aggression and violence later in life has prompted calls for early prevention efforts (Ttofi et al., 2012; Wolke & Lereya, 2015). Although bullying *per se* is not the primary target of the moral bioenhancement debate, commentators have argued that "early childhood is probably the optimal starting point for moral enhancement" (Christen & Narvaez, 2012, p. 26; Savulescu & Persson, 2012).

Our objective was to test a range of issues that are central to the debate on moral bioenhancement. We hypothesize that the degree to which members of the public support an empathy-enhancing moral enhancement program depends on whether or not the means employed were pharmacological or non-pharmacological. We expect people to be less supportive of pharmacological than of non-pharmacological programs, even when safety and efficacy are held constant. In addition, we hypothesize that people are less supportive of pharmacological moral enhancement of their own children than they are of other people's children. Second, we hypothesize that the degree to which

22 Previous studies on attitudes towards enhancement of a range of different traits suggested that people are least open to enhancing traits they believe to be more fundamental to the self/ identity, which included morally relevant traits (such as empathy and kindness) (Riis et al., 2008; Cabrera et al., 2014).

respondents support these programs depends on whether they imagine themselves or someone outside their immediate circle of concern to participate. In other words, we expect that the distinction between self and other is relevant (E. F. Williams & Steffel, 2014). We expect that this distinction matters more for the pharmacological program than the pedagogical program. Finally, we hypothesize that the public is uncomfortable with mandating moral enhancement interventions, and particularly averse to mandatory *pharmacological* moral interventions.

Methodology

Experimental methods

In order to explore attitudes of members of the public towards moral enhancement, we used the contrastive vignette technique (Burstin et al., 1980). The key outcome measure was the difference in group means between contrastive conditions rather than individual stated preferences. In addition to these quantitative measures, we also employed a novel mixed-methods design in which content analysis of free-response answers were quantitized and assessed in a contrastive fashion (Cabrera & Reiner, 2016).

Vignette design strategy

Participants were presented with one (and only one) of several contrastive vignettes in which a 13-year-old child is described as bullying another student in school and then is offered an empathy-enhancing program (see Appendix B, p. 143-5). The vignettes were designed to be minimally contrastive, plausible, and to ensure that the results would be responsive to the hypothesis under consideration. Vignettes were analyzed using the Flesch-Kincaid Reading Ease and Grade Level readability tests, and we confirmed that a 15- to 21-year-old would easily understand the text of the vignettes.

One form of contrast involved the *means* of moral enhancement: the empathy-enhancing program was described as either involving *taking a pill* or *playing a video game* on a daily basis for four weeks [means: pharmacological or non-pharmacological]. Both programs were described as being equally safe and effective. We took pains to insure that the pharmacological moral bioenhancement was as innocuous as possible, describing it as a pill "based on the natural hormone oxytocin", as we did not want to bias our results with off-putting interventions such as genetic modification or deep brain stimulation. A second form of contrast built into the vignettes compared the *closeness to the subject* of the individual who is under consideration for moral bioenhancement: participants were asked to imagine either their own child *bullying* another student at school, or their

own child *being bullied* by another student at school [closeness to subject: other's child or own child]. This resulted in a 2 x 2 between-subjects design.

In an escalating series of morally challenging questions, we asked participants to what degree they thought that it would be a good idea for the bully to participate in the program (question 1; anchors ranging from 0: bad idea to 100: good idea); to what degree they thought that it would be a good idea for the bully to be required to participate in the program (question 3); to what degree they thought that it would be a good idea for children, identified by a test to be at higher risk of being bullies in the future, to be required to participate in the program (question 5); to what degree they thought it would be a good idea for all children to be required in the program, given that the program increases empathy (question 6); to what degree they thought society would be better off if the general population was required to participate in the program (question 7; anchors ranging from 0: much worse off to 100: much better off); and to what degree they would be willing to participate in the program themselves (question 8; anchors ranging from 0: entirely unwilling to 100: entirely willing).

In the second part of the experiment, we asked participants to read a second vignette describing the same 13-year-old child bullying another student at school (either their child as bully, or their child being bullied), but this time is being required to participate in an the *alternative empathy-enhancing program*: participants who were initially presented with a program that involved taking a pill daily for four weeks, were now reading about an alternative program that involves playing a video game for four weeks – and vice versa. We presented them with the rating they gave in response to question 3, and asked them to what degree they thought that it would be a good idea for the bully to be required to participate in this alternative program.

We asked participants to explain the rationale for the answers they had given to questions 1, 3, and 9, in open response boxes (questions 2, 4, and 10), which were coded using Contrastive Quantitized Content Analysis (Cabrera & Reiner, 2016). A comprehension check probed whether participants remembered whether the vignette described their own child engaged in bullying of their own child being bullied at school. Finally, a pair of questions asked participants to *optionally* tell us whether they or their family members had been bullied or been bullies themselves.

The vignettes and questions can be found in ix B. The University of British Columbia's Behavioural Research Ethics Board approved the study.

Sample population and survey format

Participants were recruited via Amazon's Mechanical Turk (Paolacci et al., 2010; Rouse, 2015). Participants provided informed consent, and after completion of the survey were compensated $0.50. Surveys were administered using Fluid Surveys, and survey responses were collected online on June 6, 2016.

Statistical analysis

Data were analyzed using SPSS. Quantitative questions were analyzed using a two way independent Analysis of Variance (ANOVA) to identify significant main and interaction effects of the two independent variables on vignette measures. We analyzed significant effects with independent sample t-tests. Descriptive statistics were used to characterize the composition and properties of the sample. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Answers entered into the free-response box in question 2 were analyzed using the mixed-methods strategy called Contrastive Quantitized Content Analysis (CQCA) (Cabrera & Reiner, 2016). The technique provides a mechanism for quantifying the content of participants' answers and comparing them across contrastive conditions. Answers to the open-ended questions following questions 3 and 9 were included, but the data did not appear to be different so analysis of these responses is not presented here. In order to mitigate experimenter bias, we first randomized the full set of comments and blinded the coder to the particular experimental vignette read by the participant who offered a given comment. We then carried out traditional content analysis of the blinded comments, developing themes iteratively. An initial subset of ~50 comments was analyzed by two coders, and disagreements were discussed until consensus was reached. Each theme was treated as a binary variable, and each comment received either a 1 when the theme was present or 0 when the theme was absent. Once all comments were coded, the data were unblinded and the frequency with which any theme emerged in the comments was compared across contrastive conditions, with inferential statistics (Pearson Chi-Square) used to explore if any observed differences were meaningful. The code sheet used in contrastive quantitative content analysis for question 2 can be found in Table 4 (Appendix C, p. 146).

Results

Sample and demographics

A total of 384 participants from the United States completed the survey; 91 participants failed the comprehension test resulting in a final sample of 293 respondents from 38 states and the District of Columbia (missing were: Alaska, Arizona, Maine, Mississippi, Nebraska, North Dakota, Oklahoma, South Dakota, Utah, West Virginia, Wyoming). The mean age was 35.7 years old (with a standard deviation of 11.8 years). Frequencies of sample demographics are summarized in Table 5 (Appendix D, p. 147).

Do means matter morally?

An often-cited argument as to why we should explore the possibilities of moral bio-enhancement is the lack of effectiveness of so-called traditional methods of moral enhancement, such as upbringing, socialization, and education (Ingmar Persson & Savulescu, 2008; but see Harris, 2011; Zarpentine, 2013; DeGrazia, 2014). A related argument is that there are little principled differences between employing traditional and potential biomedical methods of moral betterment in terms of their ethical acceptability (Levy, 2007; Pugh, 2017). David DeGrazia, for example, contends that many arguments against biomedical means also apply to traditional, non-biomedical means: "one should not inculcate moral values that are wrong, so how can a parent be sure that she or he is justified in providing a particular type of moral instruction? Also facing this challenge are public school teachers who attempt to inculcate in students certain moral virtues such as civility, respect for differences and concern for the poor" (DeGrazia, 2014, p. 363). Likewise, according to the so-called 'companions in innocence' line of reasoning (Walker, 2009, 2010), any principled argument against the attempt to making people morally better using genetic means, will also apply to educational and socialization efforts.

Other commentators have argued that there are in fact morally relevant differences between traditional and biomedical moral enhancement, for example because education is characterized by a fundamental moral equality between educator and educated, an equality that is lacking in the case of biomedical interventions aimed at reshaping the moral agency of others (Sparrow, 2014a, p. 26). Or, along these same lines, because the distinction between (direct) biomedical (neurological, pharmacological) interventions and (indirect) traditional interventions tracks a more fundamental distinction between reason-responsive and reason-bypassing interventions, or between interventions that allow for active involvement of the person undergoing the intervention and those interventions that do not (Focquaert & Schermer, 2015).

In this study, we hypothesized that members of the public would be less supportive of pharmacological than of non-pharmacological programs, even if in the programs were described as being equally safe and effective. Moreover, we expected that the public would be uncomfortable with mandating moral enhancement interventions and with mandatory *pharmacological* moral interventions in particular. We tested these hypotheses with an escalating series of morally challenging questions.

The first question we posed asked whether it was a good idea for the bully to *participate* in the program as described in the vignette. A two way independent ANOVA showed a significant main effect of *version of program* ($F(1,289)=141,57$, $p<0.001$). An independent sample t-test revealed that respondents were significantly more supportive of an anti-bullying program that involved playing a video game than one that involved taking a pill ($M_{diff}=42.39$, 95%CI [35.32, 49.45], $p<0.001$, $d=1.38$) (Fig. 1, Participation).

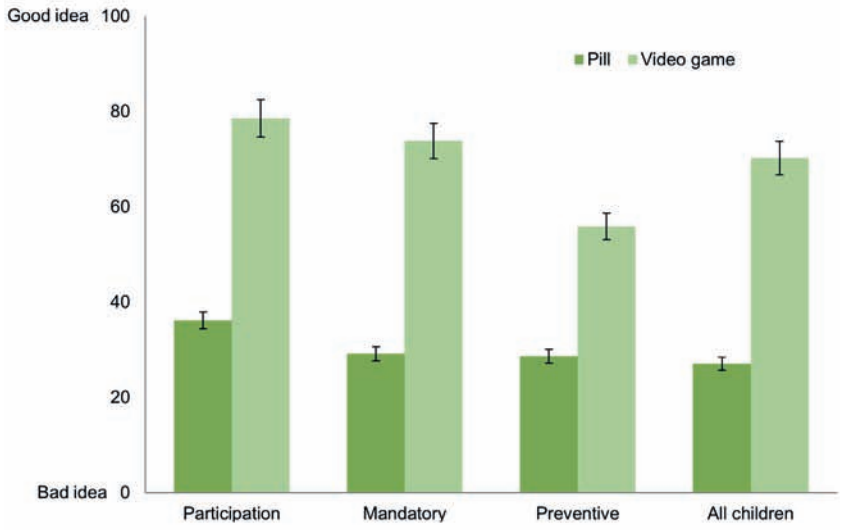


Fig. 1 Mean ratings for participation, mandatory, and preventive empathy-enhancing anti-bullying programs and for empathy-enhancing programs for all children. Error bars represent 95% confidence intervals

We then asked respondents whether it would be a good idea for the bully to be *required* to participate in the program, using the same 101-point scale. A two way independent ANOVA showed a significant main effect of version of program ($F(1,289)=147.26$, $p<0.001$). An independent sample t-test revealed that respondents were more supportive of a mandatory anti-bullying program that involved playing a video game than of a mandatory program that involved taking a pill; $M_{diff}= 44.68$, 95%CI [37.30, 52.05], $p<0.001$, $d=1.39$ (Fig. 1, Mandatory). Thus, respondents were significantly less

supportive of a mandatory pharmacological than a mandatory non-pharmacological anti-bullying program.

Next, we asked respondents to rate their support for a mandatory *preventive* program for children who were identified as being at higher risk of being bullies in the future. A two way independent ANOVA showed a significant main effect of version of program on the support rates for the preventive anti-bullying program ($F(1, 289)=51.702$, $p<0.001$). An independent sample t-test revealed that respondents supported a preventive anti-bullying program that involved playing a video game more than a program that involved taking a pill; $M_{diff}= 27.21$, 95%CI [19.77, 34.65], $p<0.001$, $d=0.84$ (Fig. 1, Preventive). People were less supportive of empathy enhancement within the context of prevention of future immoral behaviour as compared to support for empathy enhancement in cases where immoral behaviour (bullying) has already manifested itself. This is of interest for debates on "public health approaches to preventing crime" and growing attention for early identification and prevention of antisocial behaviour (Glenn & Raine, 2014; Horstkötter, 2015; Munthe & Radovic, 2015).

Subsequently we asked respondents about empathy enhancing programs that go beyond bullying in schools, specifically whether it would be a good idea for *all children* (not just bullies or potential bullies) to be required to participate in the empathy-enhancing program. A two way independent ANOVA showed a significant main effect of *version of program* on the support rates for a mandatory empathy-enhancing program for all children ($F(1,289)=131.005$, $p<0.001$). An independent sample t-test showed that respondents supported a mandatory preventive empathy-enhancing program for all children that involved playing a video game more than one that involved taking a pill; $M_{diff}= 43.13$, 95%CI [35.70, 50.57], $p<0.001$, $d=1.33$ (Fig. 1, All Children). Thus, respondents were significantly less supportive of requiring all children to participate in a mandatory pharmacological empathy-enhancing program compared to their support for required participation in a non-pharmacological program.

Taken together these results indicate that across a range of questions people were consistently more troubled by pharmacological than non-pharmacological moral enhancement interventions.

The distinction between self and other

The second hypothesis driving this study is that people rate an empathy-enhancing program differently depending on whether they are imagining that their own child or someone else's child is participating in the program. A previous study found that people employ double standards when thinking about the fairness of cognitive en-

hancement in situations where they would cognitively enhance themselves versus situations where others would do so: people perceive the same enhancing interventions as less ethically acceptable when other people use them than when they themselves use them (E. F. Williams & Steffel, 2014). Because a similar asymmetry may influence people's reasoning about moral enhancement interventions, we explored the effects of vignettes which compared the distinction between self and other.

We tested this hypothesis in two ways. First, in contrastive versions of the vignettes respondents were asked to imagine either that their own child is bullying another student in school (own child), or that their own child is being bullied by another student in school (other's child). Second, we compared responses to two questions in which we asked respondents about their support for a population-wide empathy-enhancing program (everyone else) and their willingness to participate in such a program themselves (self).

We first analyzed the data to see if there was a difference between vignettes in which the empathy-enhancing program was to be administered to one's own child who had been a bully (own child) versus those in which someone else's child had been bullying the respondent's child (other's child). Two way independent ANOVAs demonstrate that there was no significant main effect of closeness to subject on the support rates for either *participation* ($F(1,289)=.451$, $p=.502$), *mandatory* participation ($F(1,289)=1.473$, $p=.226$), or *preventive* approaches to the anti-bullying program ($F(1,289) =.994$, $p=.320$) (Fig. 2).

There was a significant interaction ($F(1,289)=4.4611$, $p=.033$) between version of program and closeness to subject when the respondents were probed on children being *required* to participate in the program, indicating that the mean difference between other's child and own child differs depending on whether the program involves a pill or a video game. Pairwise comparisons using an independent sample t-test revealed that people were more supportive of a mandatory empathy-enhancing program that involved taking a pill when they imagined the child participating in the program to be an other's child rather their own child; $M_{diff}= 12.55$, 95%CI [1.77, 23.32], $p=.023$, $d=0.38$ (Fig. 2, Mandatory).

In one of the first articles discussing moral enhancement, Thomas Douglas argued that unlike other types of enhancement, moral enhancement primarily benefits others: "on any plausible moral theory, a person's having morally better motives will tend to be to the advantage of others" (Douglas, 2008, p. 230). Others have speculated about potential societal benefits of moral bioenhancement, arguing that, "they may, through

contributing to civic virtue, help to secure the good functioning of our political institutions and processes. One way they could do this is by facilitating the dispositions towards cooperativeness and trust that plausibly underpin social solidarity" (Jefferson et al., 2014).

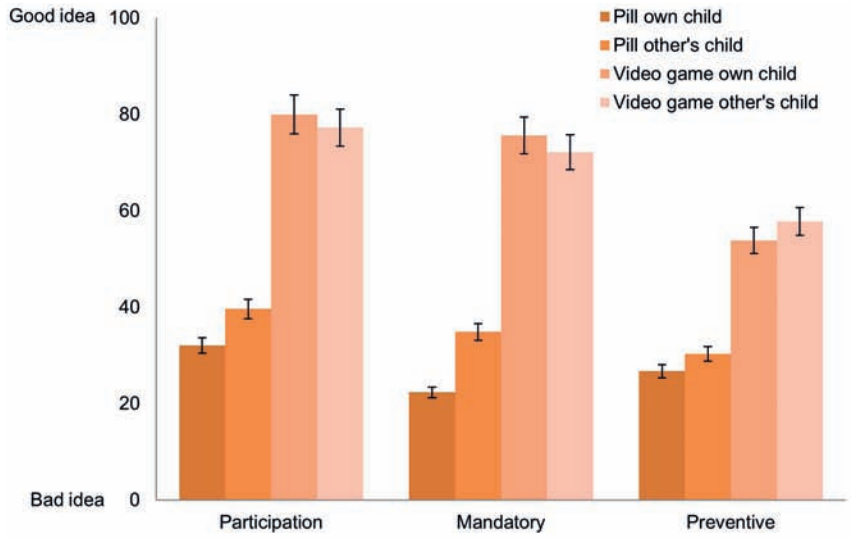


Fig. 2 Mean rating for participation, mandatory, and preventive pharmacological and non-pharmacological empathy-enhancing anti-bullying programs for own child or other's child. Error bars represent (95 %) confidence intervals

However, as "the advantages of moral enhancement may fall upon society rather than on those who are enhanced" (Focquaert & Schermer, 2015, p. 140), the need to balance potential risks to the one subjected to the program with benefits to others is arguably a central challenge when discussing *moral* enhancement. A fundamental question is how to weigh the interests and preferences of the individual and the interests of others (in view of public safety and managing public risk) (Specker & Schermer, 2017).

To address this, we asked participants whether they thought that society would be better off if the general population was required to participate in an empathy-enhancing program. As with the results presented earlier, a two way independent ANOVA showed a significant main effect of version of program on the support rates for a mandatory population-wide empathy-enhancing program ($F(1,289)=67.808$, $p<0.001$). An independent sample t-test revealed that respondents supported a mandatory population-wide empathy-enhancing program that involved playing a video game more than one that involved taking a pill; $M_{diff}= 30.72$, 95%CI [23.27, 38.20], $p<0.001$, $d=0.82$ (Fig. 3a, General population). Subsequently, we asked whether respondents would be willing to partici-

pate in the empathy-enhancing program themselves. A two way independent ANOVA showed a significant main effect of version of program on willingness to participate ($F(1,289)=93.432$, $p<0.001$). An independent sample t-test revealed that respondents were more willing to participate in an empathy-enhancing program that involved playing a video game than one that involved taking a pill; $M_{diff}= 39.02$, 95%CI [31.11, 43.51], $p<0.001$, $d=1.13$ (Fig. 3b, Self).

Finally, we asked whether participants had a history of bullying, either as perpetrators or victims of bullying. We found that 33.8% (n = 99) of the respondents indicated that they had been bullied in the past to such a degree that it interfered with their daily activities. 64.8% (n = 190) indicated that they hadn't been bullied (to such a degree), and 1.4% (n = 4) chose to not answer this question. 7.8% (n = 23) of the respondents indicated that they had bullied in the past to such a degree that it interfered with someone else's daily activities. 89.8% (n = 263) indicated that they hadn't bullied someone else (to such a degree), and 2.4% (n = 7) chose to not answer this question. Age, gender, and whether participants had been bullied or were bullies themselves had no influence on any of our parameters.

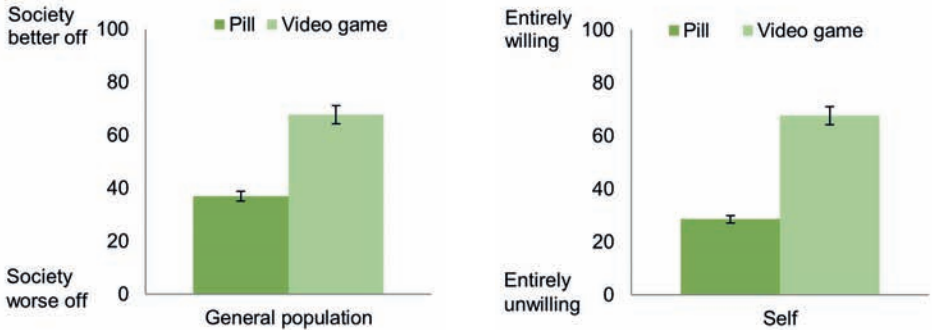


Fig. 3a and 3b Mean ratings for pharmacological and non-pharmacological empathy-enhancing programs, either with respect to the "degree society would be better off if the general population was required to participate in the program" and "degree you would be willing to participate in the program yourself." Error bars represent (95%) confidence intervals

Reasons offered for attitudes towards empathy enhancement

After rating the empathy-enhancing anti-bullying program described in the vignette on a sliding scale ranging from *good* to *bad idea* (question 1), participants were asked to explain why they answered as they did in a free-response format (question 2). Responses were analyzed using contrastive qualitative content analysis (see methods). The themes that emerged represented reasons that fell into four main categories: *Good idea, Bad idea, Ambivalent,* and *Appropriate reaction to bullying* (see Table 4 in Appendix C for the code sheet (p. 146) and Table 6 in Appendix E for the overall coding results (p. 148)).

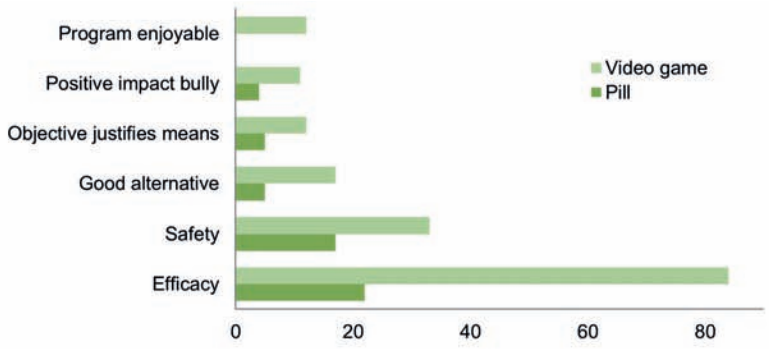


Fig. 4 Reasons why the empathy-enhancing anti-bullying program is a *Good idea* (*n*=230). Frequency of the theme as mentioned in comments (total number of comments, *n*=620)

Good idea

In the *Good idea* category, the following reasons in support of the empathy-enhancing anti-bullying program were given most frequently (cumulatively, across all versions of the vignette) (Fig. 4): the program's EFFICACY (*n*=106, 46.1 %), the program's SAFETY (*n*=50, 21.7 %), the notion that the program provides a GOOD ALTERNATIVE to other approaches (such as ignoring the problem or punishment) (*n*=22, 9.6 %), the idea that the OBJECTIVE JUSTIFIES MEANS (in spite of potential negative effects) (*n*=17, 7.4 %), the program's POSITIVE IMPACT ON THE BULLY (as it will give him better chances in life, make him a better person, or will increase his flourishing) (*n*=15, 6 %), and the idea that the PROGRAM IS ENJOYABLE (and that this will motivate the bully to participate) (*n*=12, 5.2 %).

The program's EFFICACY ($\chi 2$=57.49, df=1, p<0.001) and SAFETY ($\chi 2$=2.239, df=1, p=0.012), the notion that the program provides a GOOD ALTERNATIVE to current interventions ($\chi 2$=7.166, df=1, p=0.007), and the notion that the PROGRAM IS ENJOYABLE ($\chi 2$=12.598, df=1, p<0.001) were significantly more commonly mentioned in support of the non-pharmacological than the pharmacological program. No difference was found between version of the program and the following reasons in support of the program: the notion that the OBJECTIVE JUSTIFIES MEANS ($\chi 2$=3.111, df=1, p=0.078), the program's POSITIVE IMPACT on the BULLY ($\chi 2$=3.494, df=1, p=0.062). We found no relationship between reasons in support of the program and closeness to subject (own child or other's child), even when accounting for version of the program (pill or video game).

Bad Idea

In the *Bad idea* category, the following reasons against the program were given most frequently (Fig. 5): the notion that DRUGS SHOULD NOT BE USED (because they are artificial, because behavioural problems should not be remedied by taking drugs, or because there is nothing medically wrong with the child) (n=68, 33.7 %), the program's SUPERFICIALITY (as it addresses symptoms and not underlying causes, or because it offers no durable solution) (n=36, 17.8 %), EFFICACY DISBELIEF (disbelief that the program will effectively lower the bullying, or disbelief that increasing empathy will lower the bullying) (n=32, 15.8 %), the notion that ALTERNATIVE APPROACHES should be tried FIRST (and that the program should be a last resort) (n=31, 15.3 %), SAFETY CONCERNS (concerns about side-effects, long-term effects, and concerns about addiction) (n=16, 7.9 %).

The following concerns were more commonly brought up against the *pharmacological program* compared to the non-pharmacological program: the notion that DRUGS SHOULD NOT BE USED ($\chi 2$ = 87.949, df=1, p<0.001), the SUPERFICIALITY of the program ($\chi 2$=12.513, df=1, p<0.001), the notion that ALTERNATIVES should be tried FIRST ($\chi 2$=18.909, df=1, p<0.001), and SAFETY CONCERNS ($\chi 2$ = 16.809, df=1, p<0.001). No difference was found between versions of the program for EFFICACY DISBELIEF ($\chi 2$=2.307, df=1, p=0.129).

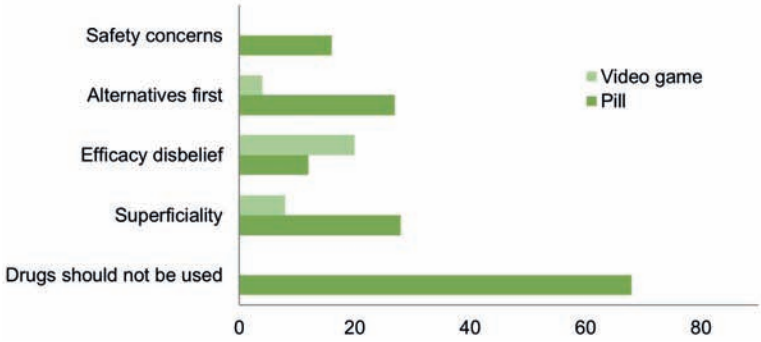


Fig. 5 Reasons why the empathy-enhancing anti-bullying program is a *Bad idea* (n=202). Frequency of the theme as mentioned in comments (total number of comments, n=620)

Chi-squared tests were performed to determine whether there was a relation between reasons why respondents rated the program as a bad idea and their imagined closeness to the bully (own child or other's child). We found no significant relationship between closeness to subject (own child or other's child) and particular reasons provided against the program without stratifying for version of program (pharmacological or non-pharmacological). However, the notion that ALTERNATIVES should be tried FIRST ($\chi 2$=5.930,

df=1, p=0.015) and SAFETY CONCERNS ($\chi 2$=36.886, df=1, p=0.049) were more commonly mentioned as reasons against the program when the program described in the vignette was a *pharmacological* program intended for one's *own child*. Moreover, we found that comments by respondents imaging their *own child* participating in a *pharmacological program* were more often coded as AMBIVALENT ($\chi 2$=4.848, df=1, p=0.028).

Respondents appear to perceive a difference in safety between non-pharmacological and pharmacological programs, in spite of the fact that the empathy-enhancing program was presented as equally safe and effective in contrasting versions of the vignettes. When confronted with a non-pharmacological program, many respondents indicated that the program's safety was an important reason for their support of the program, whereas respondents who had read a vignette that described a pharmacological program mentioned concerns about safety as a reason against the program. Moreover, concerns about safety were more commonly mentioned in response to pharmacological programs intended for own child than for other's child.

Appropriate reactions to bullying

In their answers, respondents oftentimes not only provided reasons for or against the program, but also explained what they considered to be an appropriate reaction to bullying. Appropriate strategies to alleviate bullying that were mentioned by respondents were: TEACHING (n=48, 31.4 %), EMPATHY (n = 41, 26.8 %), UNDERSTANDING (n=35, 22.9 %), HELP (n=19, 12.4 %), PUNISHMENT (n=8, 5.2 %), and moral AGENCY (n=2, 1.3 %) (Fig. 6).

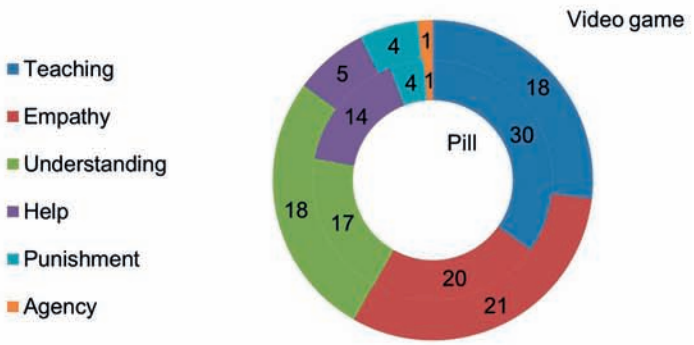


Fig. 6 Appropriate reactions to bullying (n=153). Frequency of the theme as mentioned in comments (total number of comments, n=620). (Inner circle = pill, outer circle = video game).

Moreover, respondents frequently made explicit whether they thought the program as described in the vignette indeed consisted of the strategy to alleviate bullying that they preferred or found to be most promising (Fig. 7). Chi-square tests were performed

to determine whether there was a relationship between what respondents considered an appropriate reaction to bullying, and version of the program (pill or video game). Respondents more commonly reasoned that the bully needed HELP when confronted with a pharmacological rather than a non-pharmacological program ($\chi 2$=4.493, df=1, p<0.034) (data not shown). We found no significant relation between version of program and respondents indicating that the bully needed TEACHING, EMPATHY, UNDERSTANDING, PUNISHMENT, or AGENCY.

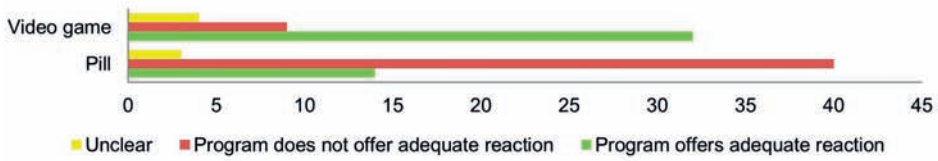


Fig. 7 Does the program provides an adequate reaction to bullying or not (n=153). Frequency of the theme as mentioned in comments (total number of comments, n=620).

When the program described in the vignette consisted of playing a video game, respondents more commonly expressed that they considered the program an adequate reaction to bullying ($\chi 2$=8.502, df=1, p=0.004), while when the program involved taking a pill, respondents more commonly indicated that they did not consider the program an adequate reaction to bullying ($\chi 2$=23.298, df=1, p<0.001) (Fig. 7).

In conclusion, the difference between self and other (either between one's own child and an other's child, or between the general population and oneself) influences the public's support for moral enhancement. For the pharmacological empathy-enhancing program, respondents were more critical when they imagined their own child rather than an other's child to be the one subjected to the program. They more often expressed ambivalence, mentioned safety concerns, and argued that alternatives should be tried first. And as regards the pharmacological program, respondents were more open to requiring the general population to participate in the program, but were less willing to participate themselves. However, the difference between self and other was largely absent for vignettes describing a non-pharmacological program, suggesting that this distinction is morally salient only when safety and risk concerns come into play.

Discussion

This study provides empirical evidence that *means matter morally*. For when it comes to moral enhancement, members of the public generally eschew pharmacological moral

bioenhancement yet are open to non-biomedical means to attain moral enhancement. Both the quantitative and the qualitative data confirm that the public disapprove of biomedical interventions for moral enhancement. These findings were confirmed convincingly across a range of questions and in all versions of the vignettes.

These findings are in line with previous research demonstrating a considerable bias against or mistrust of "pills" in general (Bergström & Lynöe, 2008; Partridge et al., 2014; Schermer, 2016). The added value of this study is that it sheds light on what kind of reasons members of the public have for their dislike of pharmacological interventions. Interestingly, when reflecting on the non-pharmacological program, many respondents explicated their support by reference to the fact that the program is described as safe and effective in the vignette. However, respondents who had read a vignette about the pharmacological program were often sceptical about the program's safety and effectiveness, even though the program was described as equally safe and effective in the vignette. Again, as might be expected based on earlier research, respondents argued that pills are bad because they are artificial or unnatural, expressed concerns about safety and undue medicalization (over-medicalization) of behavioural problems (Summers & Caplan, 1987; Mcleod et al., 2004). In addition however, many respondents reasoned that the pharmacological program offered no "real" solution to the problem; they were sceptical about the long-term effectiveness of the program and expressed concerns about the pharmacological program being too superficial and not adequately addressing underlying causes of the bullying behaviour. With the video game, respondents were more optimistic about its lasting effects.

Moreover, many respondents explained that the reason why they did or did not support the empathy-enhancing program in the vignette was related to whether or not they were under the impression that the program offered an appropriate pedagogical response to the problem behaviour (bullying). This might be interpreted as an indication that the public values fostering in children deeper understanding of and insight into why certain behaviour is morally wrong rather than mere conformity to (moral) rules. The public appears concerned not only about effectiveness (will an intervention reliably lower the immoral behaviour?) but also about whether the one participating in the program will, as a result of the intervention, have *learned* something on a deeper level and in the longer run.

One important consideration in interpreting these results is the fact that the scenario in our vignettes concerns bullying by a 13-year-old child. People might be particularly resistant to giving pharmacological substances to children, and different ethical considerations may come into play, for example about the responsibility of parents and

schools. However, comments given in the open-response boxes indicate that the concerns respondents have go beyond mere 'pills are bad'-considerations. Moreover, next to empathy enhancement in children, we also asked respondents about their support for empathy enhancement for the general population and for themselves.

Concerns about the importance of *doing the right thing, period* as opposed to *doing the right thing for the right reasons* are also raised in the ethical debate on moral bioenhancement. Douglas (Douglas, 2014c, pp. 1241-1243) discusses so-called "superficiality concerns" about forms of non-cognitive or reason-bypassing interventions that directly alter emotions. These interventions can be considered *brute* as opposed to *deliberate* (Danaher, 2013), because they directly alter emotions without requiring the exercise of deliberative faculties, such as "moral reasoning, introspective reflection on one's moral failures, or calm moral discussion with others" (Douglas, 2014b, p. 79). Douglas argues that these kinds of interventions are sometimes permissible (Douglas, 2008); Persson and Savulescu even argue that moral bioenhancement might be morally obligatory (Ingmar Persson & Savulescu, 2017). Other commentators disagree and reason that pharmacological or neuro-scientific interventions fail to produce *deep* moral understanding and *deep* moral improvement because these kinds of interventions fail to provide any moral *content.*[23]

Harris maintains that to be a moral agent is to consider moral reasons for action. He argues that direct, reason-bypassing interventions "might well take the conduct of the affected individual beyond moral review and certainly out of the realm of things that might be right all things considered" (Harris, 2012, p. 269). Fabrice Jotterand stresses that on a virtue ethical account, both moral emotions and moral reasoning are essential for autonomy and true moral agency:

> On my analysis, I conclude that moral neuroenhancement is unlikely to morally enhance people in the true meaning of the word. The development of neurotechnologies will allow us to control moral emotions but not to generate any content for moral reasons for actions. Without a systematic reflection on the nature of the good, the right and the just, one would end up, using MacIntyre's language, in bad character because of intellectual blindness. Moral agency requires understanding and the formation of right moral emotions. (…) The hope of controlling human moral emotions is insufficient for the formation of virtuous people. Moral agents

23 Bernard Baertschi explains that these disagreements can partly be attributed to different (i.e. cognitivist versus sentimentalist) meta-ethical positions (Baertschi, 2014).

> are not engineered but trained through the development of a vision of the good life and an understanding of human flourishing. (Jotterand, 2011, p. 8)

These comments suggest that at least some philosophical positions in the ongoing debate align with public opinion. Indeed, the reasons participants offered in their free responses reflect to large extent key themes discussed in the neuroethics literature. Furthermore, it is one thing to argue eloquently for or against the propriety of such things as mandatory biomedical moral enhancement and quite another to accept that mandate for yourself, or even more importantly for your children. Except under the auspices of a totalitarian state, the prospect of widely disseminating moral bioenhancement depends entirely upon the accession of the public. Our data demonstrate quite clearly that support for such a project is absent, even though advancing the moral skills of the populace enjoys widespread support.

Appendix B
Vignettes and questions

PHARMACOLOGICAL/ OTHER'S CHILD BULLIES	PHARMACOLOGICAL/ OWN CHILD BULLIES
Imagine that your 13-year-old child was being bullied by another student at school. The school has a program that has been shown to be effective in reducing bullying in carefully carried out studies. The program involves the following: over the course of 4 weeks, each day the bully takes a pill that increases empathy for others. The pill is based on the natural hormone oxytocin, and improves the bully's ability to understand what other people are feeling. Studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete. [104 words]	Imagine that your 13-year-old child was bullying another student at school. The school has a program that has been shown to be effective in reducing bullying in carefully carried out studies. The program involves the following: over the course of 4 weeks, each day the bully takes a pill that increases empathy for others. The pill is based on the natural hormone oxytocin, and improves the bully's ability to understand what other people are feeling. Studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete. [102 words]
NON-PHARMACOLOGICAL/ OTHER'S CHILD BULLIES	NON-PHARMACOLOGICAL/ OWN CHILD BULLIES
Imagine that your 13-year-old child was being bullied by another student at school. The school has a program that has been shown to be effective in reducing bullying in carefully carried out studies. The program involves the following: over the course of 4 weeks, each day the bully plays a video game that increases empathy for others. The video game is based on best educational practices, and improves the bully's ability to understand what other people are feeling. Studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete. [105 words]	Imagine that your 13-year-old child was bullying another student at school. The school has a program that has been shown to be effective in reducing bullying in carefully carried out studies. The program involves the following: over the course of 4 weeks, each day the bully plays a video game that increases empathy for others. The video game is based on best educational practices, and improves the bully's ability to understand what other people are feeling. Studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete. [103 words]

Q1. To what degree do you think that it would be a good idea for the bully to participate in a program like the one described above?

0 100

Bad idea Good idea

Q2. Please tell us why you answered as you did.

Q3. To what degree do you think that it would be a good idea for the bully described above to be required to participate in the program?

0 100

Bad idea Good idea

Q4. Please tell us why you answered as you did.

Q5. If there was a reliable test that identified children who are at higher risk of being bullies in the future, to what degree do you think that it would be a good idea if they would be required to participate in the program?

0 100

Bad idea Good idea

Q6. Given that the program increases empathy, to what degree do you think it would be a good idea for all children (not just bullies or potential bullies) to be required to participate in the program?

0 100

Bad idea Good idea

Q7. If this program increased empathy in everyone, to what degree would you think that society would be better off if the general population was required to participate in the program?

0 100

Much worse off Much better off

Q8. To what degree would you be willing to participate in the program yourself?

0 100

Entirely unwilling Entirely willing

Q9.

There is an alternative program that is equally effective in increasing empathy for others, but involves the following: over the course of 4 weeks, each day the bully takes a pill that increases empathy for others. The pill is based on the natural hormone oxytocin, and improves the bully's ability to understand what other people are feeling. Once again, studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete.	There is an alternative program that is equally effective in increasing empathy for others, but involves the following: over the course of 4 weeks, each day the bully plays a video game that increases empathy for others. The video game is based on best educational practices, and improves the bully's ability to understand what other people are feeling. Once again, studies have shown that the program reduces bullying by 40%, with no side effects. The reduction in bullying persists for 6 months after the program is complete.

Given that this alternative program exists, to what degree do you think that it would be a good idea for the bully described above to be required to participate in this alternative program?

[Your rating for the original program was XX.]

0 100

Bad idea Good idea

Q10. Please tell us why you answered as you did.

Q11. In the vignette described above you were asked to imagine that
Your 13-year-old child was being heavily bullied by another student at school
Your 13-year-old child was heavily bullying another student at school

Q12. Have you or any of your family members ever been so substantially bullied that it interfered with your daily activities? [optional question]
Yes / No

Q13. Have you or any of your family members ever been so much of a bully that it interfered with someone else's daily activities? [optional question]
Yes / No

Appendix C
Table 4: Code sheet

Prefer Use	**Safe:** this code captures comments that the program is safe.
	Effective: this code captures comments that the program is effective or that mention the degree of effectiveness of the program.
	Objective justifies means: this code captures comments that in spite of potential negative effects, the program is justified or worth trying given the negative effects of bullying.
	Positive Impact bully: this code captures comments that the program will likely benefit the child that is bullying (will give him more chances in life, will make him a better person, will increase his flourishing).
	Positive Impact victims: this code captures comments that the program will likely benefit the child that is bullied.
	Program is enjoyable: this code captures comments that the bully likely would enjoy the program, be motivated to participate in it, and that this would make it more likely that (s)he will stick with it.
	Good alternative to present approaches: this code captures comments that the program provides a good or viable alternative to present approaches to bullying (such as ignoring the problem, or punishment), and that the bullying behaviour needs to be addressed, that something needs to be done to stop the bullying.
	Pill not bad: this code captures comments that there is nothing (inherently) wrong with taking a pill.
Against Program because	**Safety Concerns:** this code captures comments about side effects and safety concerns, about risk of becoming dependent or addicted, and comments that express disbelief that the program has no side effects, or no side effects in the long term.
	Efficacy Disbelief: this code captures comments that explicitly express disbelief that the program will be effective in reducing bullying (for example because 'once a bully, always a bully', or because increase in empathy will not decrease bullying).
	Drugs should not be used to remedy behavioural problems in children/ concerns about medicalization of normal child behaviour: this code captures comments that behavioural problems should not be remedied by taking drugs, comments that there is nothing medically wrong with the child/ no medical condition/ no acute medical condition, general comments that drugs are bad, wrong, or should not be used, as well as concerns that the program is artificial or not natural (e.g. chemical feelings, hormones).
	Need for (more) permanent and less superficial solution: this code captures concerns that the program offers no permanent or durable solution to the bullying, and/ or that the program should not sidestep the real issues, that it should address underlying causes, not symptoms, as well as comments that express disbelief that the program will be effective after six months, after the program stops, in the long term, etc.
	Alternatives First/ last resort: this code captures comments that other ways of dealing with the bullying child need to be tried first, as well as comments that the program should be a last resort, or only be used for the most grave instances.
	Autonomy: this code captures comments that the bully should have the right to consent to the program, or that the program infringes on his autonomy.
	Identity, personality, fundamental changes to the self, diversity: this code captures comments that a person should not be asked, or forced to change his personality, as well as comments that ask who should decide what kind of changes in personality would in fact be desirable, and comments that people differ and that diversity in personality traits should be embraced instead of eliminated.
	Parents should decide: this code captures comments that parents should be consulted, and consent to the program.
	Doctor should be consulted: this code captures comments that a doctor needs to be consulted.

Ambivalent	**Ambivalent:** When it was overtly stated one or more reasons for and against the program, or that the commenter was unsure.
Appropriate reaction to bullying/ The problem underlying bullying that needs to be addressed is	**Bully needs to be taught:** this code captures comments that the bullying child needs to be taught.
	Bully needs to understand: this code captures comments that the bullying child needs to understand why bullying is wrong.
	Bully needs to empathize with victims: this code captures comments that the bully needs to empathize with victims, and feel and understand how his actions hurt other people.
	Bully needs punishment: this code captures comments that the bullying child needs punishment.
	Bully needs help: this code captures comments that the bully needs help, that he likely has problems, is a victim himself.
	Bully needs to take responsibility/ agency: this code captures comments that the bully needs to bear the consequences of his (past) behaviour, needs to take responsibility for (future) behaviour.
Program provides an appropriate reaction to bullying	**Program provides:** this code captures comments that program is an adequate reaction to bullying.
	Program does not provide: this code captures comments that the program is not an adequate reaction to bullying.
	Unclear whether program provides: this code captures comments where it is unclear whether the program provides an adequate reaction to bullying.

Appendix D

Table 5: Sample demographics

Characteristic	*% (n)*
Gender	
Male	50.9% (149)
Female	49.1% (144)
Age (groups)	
18-30	41.6% (122)
31-45	38.9% (114)
46-65	18.1% (53)
>65	1.4% (4)
Education (groups)	
Some high school	0.3% (1)
High school diploma	13.0% (38)
Some college	31.1% (91)
College degree	43.0% (126)
Some post-graduate	2.7% (8)
Post-graduate degree	9.9% (29)

Appendix E
Table 6: Overall coding results

	All	Pill	Video Game	Pill own child	Pill other's child	Video game own child	Video game other's child	Own child	Other's child
GOOD Efficacy	106	**22**	***84**	9	13	39	45	48	58
GOOD Safety	50	**17**	***33**	9	8	15	18	24	26
GOOD Good alternative	22	**5**	***17**	1	4	6	11	7	15
GOOD Objective justifies means	17	5	12	2	3	4	8	6	11
GOOD Positive impact bully	15	4	11	0	4	8	3	8	7
GOOD Other									
GOOD Program enjoyable	12	**0**	***12**	0	0	7	5	7	5
GOOD Positive impact victim	5	3	2	1	2	1	1	2	3
GOOD Pill not bad	3	3	0	1	2	0	0	1	2
BAD Drugs are bad	68	***68**	**0**	36	32	0	0	36	32
BAD Superficiality	36	***28**	**8**	17	11	1	7	18	18
BAD Alternatives first	31	***27**	**4**	***18**	**9**	1	3	19	12
BAD Efficacy disbelief	32	12	20	3	9	12	8	15	17
BAD Safety concerns	16	***16**	**0**	***11**	**5**	0	0	11	5
BAD Other									
BAD Fundamental changes to self	5	1	4	1	0	2	2	3	2
BAD Parents should consent	7	5	2	3	2	0	2	3	4
BAD Autonomy	3	2	1	1	1	0	1	1	2
BAD Doctor should be consulted	4	4	0	3	1	0	0	3	1
Ambivalent	35	17	18	***12**	**5**	9	9	21	14
Overall FOR	139	**31**	***108**	**7**	***24**	52	56	59	80
Overall AGAINST	111	***94**	**17**	45	49	7	10	52	59
Overall UNCLEAR	9	6	3	4	2	2	1	6	3
NEED FOR Teaching	48	30	18	15	15	8	10	23	25
NEED FOR Understanding	35	17	18	5	12	11	7	16	19
NEED FOR Empathy	41	***20**	**21**	9	11	13	8	22	19
NEED FOR Punishment	8	4	4	1	3	1	3	2	6
NEED FOR Help	19	***14**	**5**	8	6	2	3	10	9
NEED FOR Agency	2	1	1	0	1	1	0	1	1
YES, program provides	46	**14**	***32**	4	10	18	14	22	24
NO, program does not provide	49	***40**	**9**	20	20	3	6	23	26
UNCLEAR whether program provides	7	3	4	3	0	3	1	6	1

7
General discussion

Psychopaths and God Machines: A deeply provocative and puzzling debate

The end of human civilization is nearing, as a consequence of a toxic mix of rapid technological developments, an outdated moral psychology, and an evil minority that is capable to cause ever greater – even *existential* – harm. Our only hope in stopping climate change, preventing terrorism, solving global injustices, in short, in averting disaster, is to find ways to change no less than the moral character of humanity. In fact, there is such an abundance of evil, and the urgency is so great, that the development and implementation of state initiated, mandatory bioenhancement programs is justified.[24] Quietly awaiting people to acknowledge and change their flawed natures themselves is simply not an option.

This assessment of the apocalyptic character of (parts of) the debate on moral enhancement might seem a bit of a caricature. Still, in writing this thesis, many a time I have marvelled at the unworldly nature, and at times even silliness, of the debate. To borrow the words of Harris Wiseman:

> The scope of moral bioenhancement is great: presented by some in terms of pure fantasy, by others as hard-hitting real future prospects and as offering remedies for every last moral concern from all our petty and mostly harmless vices and to various addictions and all sorts of "undesirable behaviours" up to the ultimate fate of humanity and moral bioenhancement's apparently salvatory promise for humanity against its own inner biological evils. (Wiseman, 2017, p. 398)

My main motivation in writing this thesis has been to understand in what ways the debate on moral enhancement can be relevant for and could impact on existing or emergent human practices – and the other way around. Some authors express scepticism towards the potential real world relevance of the debate and have posed the question whether "moral enhancement is ever to gain relevance apart from merely theoretical interest" (Beck, 2015, p. 234)? In this final chapter, contrary to such scepticism, I will take the claim that a range of present and emerging practices already contain ele-

24 In Savulescu and Persson's God Machine thought experiment, such a large scale project is described as follows: "The Great Moral Project was completed in 2045. This involved construction of the most powerful, self-learning, self-developing bioquantum computer ever constructed called the God Machine. The God Machine would monitor the thoughts, beliefs, desires and intentions of every human being. It was capable of modifying these within nanoseconds, without the conscious recognition by any human subjects." (Savulescu & Persson, 2012, pp. 412-413) See the introduction of this thesis for a more extensive discussion of The God Machine thought experiment.

ments of "moral enhancement by proxy" (Wiseman, 2016, p. 219) as a starting point. I continue to understand moral enhancement as formulated in the working definition in the introduction, to include biomedical and non-biomedical interventions, and assume that in principle, both of these merit equally intense ethical scrutiny.

As technological possibilities (and perhaps public and political willingness) to influence and alter moral behaviour and moral capacities increase, the responsibility to make explicit this moral enhancement dimension in our current practices, and to think carefully about the associated ethical issues, increase as well.

Evaluating moral enhancement

As discussed before, this thesis aims at formulating conditions for ethically justifiable moral enhancement practices, and focuses to a lesser degree on analysing particular moral enhancement technologies. This represents a choice, based on the observation that at present there are very few concrete moral enhancement technologies available – or at least, that there is little consensus on the criteria to decide which technologies would in fact constitute moral enhancement technologies. The choice is also based on the observation that there is little agreement on what would be a suitable evaluative framework to analyse moral enhancement, and what kind of elements should be included in it.

Before formulating conditions for ethically justifiable moral enhancement practices, first some questions need to be addressed and some distinctions made.

Do we know or simply assume the alleged positive and negative effects of potential moral bioenhancement technologies?

Commentators in this debate have, to a lesser or larger extent, assumed moral bioenhancement to take certain forms, be focused on certain capacities, to have certain effects; and from those implicit or explicit assumptions, have taken up a position towards moral bioenhancement. For example: Harris argues that moral bio-enhancement interventions will necessarily impede on human freedom, because he assumes that moral enhancement will bypass or distort reasoning, and therefore interfere with ethical analysis rather than enhance it. Likewise, Reichlin claims that biotechnological means will have serious negative effects on moral agency and capacity of authentic moral behaviour: "biotechnological ways of producing moral progress (…) risk having serious negative effects on our moral agency, by causing a substantial loss of freedom and capacity of authentic moral behaviour, by affecting our moral identity and by

imposing a standard conception of moral personality" (Reichlin, 2017). These assumptions about the effects of emerging or fictitious moral enhancement interventions are not necessarily warranted.[25]

It has been argued, for example, that neuromodulation techniques can have disrupting, restorative, or enhancing effects on autonomy, depending on the context, the exact effects, and the person in question (Schermer, 2015). Moreoever, there are no guarantees that traditional procedures, like talk therapy, to name one example of a nonbiomedical intervention, will necessarily foster autonomy or other morally relevant capacities – it could do the exact opposite. Rather than differentiating ethically appropriate from ethically worrisome means, it is preferable to explore the conditions under which a particular intervention is most likely to undermine human freedom or human agency – and whether there are measures that would minimize these risks – and, vice versa, the conditions under which an intervention can restore or enhance them. The potential – positive or negative effects – of moral enhancement interventions should not be assumed, but be considered as largely *open questions,* which can only be answered in relation to a specified context or practice.

Distinguishing moral self-enhancement from moral other-enhancement

Can we imagine there ever to be a commercial market for moral enhancers, resembling current (semi-legal) markets for cognitive enhancers? Is it likely that there will be do-it-yourself communities experimenting with moral *self* enhancement, similar to current communities that are experimenting with neurostimulation for self-improvement purposes (Wexler, 2015, 2017)? Are people in fact interested in morally enhancing themselves, or do they think that morally enhancing *others* should be prioritized?[26] These are, to a certain extent, empirical questions with regard to which the jury is still out – as is true for many questions surrounding the debate on moral enhancement.

Everyday experience suggests that behaving morally, and being motivated to do so, is not out of the ordinary. People are prone to reflect normatively, behave altruistically, and sacrifice themselves for what they consider good causes. For her book *Strangers Drowning: Grappling with Impossible Idealism, Drastic Choices, and the Urge to Help*, Larissa

25 Horstkötter and colleagues caution against undue "bio-exceptionalism" and the apparent gap between the intensity of ethical scrutiny life sciences approaches (to antisocial and criminal behavior) receive as compared to psychosocial approaches (Horstkötter, Berghmans, & de Wert, 2014).

26 Persson and Savulescu argue that those most in need of moral enhancement, are least likely to pursue it: "If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them" (Ingmar Persson & Savulescu, 2008, p. 174).

MacFarquhar portrays a range of people who, in different ways, go at great lengths to do good (MacFarquhar, 2015). Being motivated to *better* oneself morally is also not uncommon. To give but one example, the so-called Effective Altruism movement describes the most effective ways of doing right for people who want to do better but do not know how (Singer, 2015). The Giving What We Can charity and the 80.000 Hours movement encourage people to donate at least 10% – and then gradually more – of their income, and to choose the most high-earning career (not the profession one has a passion for) in order to be able to give away as much as possible.[27]

Research shows that people consider moral traits to be central to their identity, and suggests that they consider them equally or even more important than other mental faculties, such as (emotional and autobiographical) memory, lower-level cognition, and perception (Strohminger & Nichols, 2014). At the same time, studies show that most people (irrationally) believe themselves to be morally superior to the average person:

> Most people strongly believe they are just, virtuous, and moral; yet regard the average person as distinctly less so. (…) virtually all individuals irrationally inflated their moral qualities, and the absolute and relative magnitude of this irrationality was greater than that in the other domains of positive self-evaluation. (…) Taken together, these findings suggest that moral superiority is a uniquely strong and prevalent form of "positive illusion." (Tappin & McKay, 2016, p. 623)

If people care about their moral qualities, yet tend to overestimate them, this raises the question whether and to what extent people indeed consider themselves to be a candidate for moral enhancement. A study by Jason Riis and colleagues confirms that people consider morally relevant traits such as empathy and kindness fundamental to their self, but are least willing to pharmacologically modify these morally relevant traits compared to other traits they considered less morally relevant, such as wakefulness and reflexes, possibly precisely because these traits are so closely related to personal identity (Riis et al., 2008).

Chapter 6 of this thesis suggests that in general people are open to morally improving themselves (in terms of increasing their empathy for others), but that they care about (and change their preferences based on) the process, or means by which this change in their moral capacities comes about (Specker et al., 2017). Participants' willingness

27 See the website of the Oxford based Centre for Effective Altruism: https://www.centreforeffectivealtruism.org/; the website of Giving What We Can: https://www.givingwhatwecan.org/; and the 80.000 Hours Organization: https://80000hours.org/.

to participate in an empathy-enhancing program themselves was greater for the non-pharmacological as compared to the pharmacological program. Taken together, these findings suggest that people are open to morally improving themselves, but that widespread willingness to participate in moral self enhancement by means of biomedical interventions is lacking.

Arguably, the distinction between moral self enhancement and moral other enhancement is relevant when ethically evaluating (potential) moral enhancement practices (see chapter 3 and 6). As we have seen, a central and recurring question in the moral enhancement debate is what the need for and objectives of developing and implementing moral enhancement interventions are. These needs and objectives are often described in terms of potential advantages to others (reduction of harm) or society (significant overall decrease in criminal, violent, or otherwise antisocial behaviour).[28] Indeed, in his first article on moral enhancement, Douglas argued that because moral enhancement benefits others, it compares favourably to other forms of enhancement:

> Unlike the most frequently mentioned varieties of enhancement, enhancements satisfying this formula for moral enhancement could not easily be criticised on the ground that their use by some would disadvantage others. On any plausible moral theory, a person's having morally better motives will tend to be to the advantage of others. (…) One could not object to moral enhancement on the ground that it would systematically impose morally gratuitous disadvantage on others. (Douglas, 2008, p. 230)

With respect to moral self enhancement, the decision to pursue moral enhancement and to determine the goal and appropriate means of doing so, is to a large degree up to the person herself, provided a number of safeguards are put in place. It is the responsibility of those offering potential 'moral enhancement' interventions, programs, and courses, to safeguard that what they offer is safe and effective, and to provide ample opportunity for informed consent.[29] Doing so yields many challenges in and of itself, comparable to the challenges faced by anyone offering counselling, training, or therapy in the medical or semi-medical domain to people who come to them with a

28 DeGrazia provides the following rationale for why moral bioenhancement needs serious consideration: "because the status quo of moral behaviour is deeply problematic and traditional means of moral enhancement may prove inadequate to achieve needed improvements—notwithstanding the phenomenon of moral progress. The status quo is deeply problematic because there is such an abundance of immoral behaviour, with devastating consequences, and serious risk of worse to come." (DeGrazia, 2014, p. 362)

29 In principle, this is equally important for biomedical and nonbiomedical interventions and programs.

request for help. It may also be comparable to the challenges faced by regulators and scientists who are trying to design sensible regulations and robust safety standards for direct-to-consumer personal genome testing (Bunnik et al., 2014) or home use of neurostimulation (Fitz & Reiner, 2013). For example, ethicists have called on scientists and journalists to balance their enthusiasm about the promises of new brain stimulation technologies with restraint, and to provide realistic information about potential harmful effects as well (Fitz & Reiner, 2013, p. 411).

Clearly, in the case of moral enhancement of *others*, there are additional and much larger responsibilities to justify the need, objectives, and procedure and the means employed, as well as the balancing of potential benefits and harms. This is true especially in the case of interventions that are initiated, incentivized, or mandated by a state or state body.[30] If we take proponents of moral enhancement at their word, at least some of their proposals would involve some sort of moral enhancement *program*. Mark Walker's proposal for a Genetic Virtue Program "to reduce evil in our world" by "engineering genetic virtue" might serve as an example (Walker, 2009, 2010).

Distinguishing between different target populations

As was discussed in chapters 3 and 4, in order to determine a suitable target population for moral enhancement programs *for others*, for example in terms of who will benefit most from a specific intervention, or of which individual or groups need to be targeted to optimize societal benefit, some form of screening will be necessary (Specker & Schermer, 2017). Moreoever, specific concerns are associated with implementing moral enhancement interventions in particularly vulnerable target gourps.

In children, there might be unique (but oftentimes unknown) effects on the developing brain (Cohen Kadosh et al., 2016)[31], challenges surrounding decision-making processes and consent (Focquaert, 2013; Maslen et al., 2014), and potential effects on a child's future (and developing) autonomy. Concerns about negative effects of screening and selection apply to educational contexts, with regard to programs aimed at lowering or preventing a particular kind of problem behaviour (e.g. the example of bullying in chapter 6). One potential risk is that children or youngsters, who are at risk and in need of help, but exhibit internalizing rather than externalizing problem behaviour,

30 Sarah Carter discusses ethical issues surrounding incentivizing programs for moral enhancement (S. Carter, 2015). In addition, considerable attention has been devoted to the distinction between voluntary and mandatory moral enhancement (Rakić, 2014b, 2014a; S. Carter, 2015; Baccarini & Malatesti, 2017; Rakić, 2017).

31 Again, in principle, this applies to potential effects on the developing brain of both biomedical and nonbiomedical interventions.

risk staying out of sight. In these contexts especially, the best interests of all children should be of concern of educators, social workers, and policy makers, and an exclusive focus on 'troublemakers' is not justified (Kaltiala-Heino & Eronen, 2015).

The group of forensic psychiatric patients is also vulnerable,[32] for example in terms of concerns over coersive measures (Nedopil, 2016) or so-called coercive offers, and the question whether people in forensic psychiatric care contexts can ever truly give their informed consent (McMillan, 2013, 2014; Ryberg, 2015; Adshead & Davies, 2016) (see chapters 3-5 for more extensive discussion of the specific challenges associated with the forensic field).

Conditions for ethically justifiable moral enhancement practices

Proponents of enhancement in previous debates have often defended and shown a strong, even libertarian, commitment to individual liberty and freedom of choice. However, in the debate on moral enhancement, a number of central commentators appear to depart from this commitment, when they advocate for mandatory, population-wide moral enhancement programs, aimed at solving societal problems like climate change, war, and social evil by changing individual biology.[33]

Those who advocate population-wide, mandatory moral enhancement programs would have to convincingly argue that the need for such large-scale programs is so great, that abandoning central and quite basic principles of liberal democracy is warranted.[34] In addition, they would have to show that their proposed solution – correcting individual moral deficiencies – in fact meets that urgent need adequately and effectively.

In addition, a range of regrettable historical as well as current examples caution against putting the full force of the state behind moral enhancement programs aiming to change a person in such a way, as to make her conform to (or refrain from) what in the eyes of those that order it, is (un)wanted or (im)moral. An often discussed example

---

32 "In the emotionally (and politically) charged context of crime, the imposition of unproven technical "fixes" on the always unpopular class of "criminals" seems quite plausible" (Greely, 2007, p. 1129).

33 Stefan Schlag argues that Persson and Savulescu misunderstand the nature of collective action problems: "The argumentative difficulties possibly arise because the authors misunderstand the basic problem of the tragedy of the commons. Social dilemmas are not caused by individual moral deficiencies but are rooted in the problematic structure of human interaction. Neuroscientific interventions into the biological basis of human behaviour are fundamentally inappropriate means to solve problems of this type" (Schlag, 2016, pp. 11-12).

34 One such a fundamental liberty is the right to privacy – which Persson and Savulescu recognize as a legal but not a moral right, and for that reason should not stand in the way of implenting moral enhancement programs (Ingmar Persson & Savulescu, 2012; but see: Bublitz, 2016).

concerns homosexuality, which has long been criminalized (and still is in large parts of the world), and which has been classified by the American Psychiatric Association as a mental disorder until 1973 and as a sexual orientation disturbance until 1987 (American Psychological Association Task Force on Appropriate Therapeutic Responses to Sexual Orientation, 2009). Various sexual orientation change efforts are still applied today.

These kinds of examples show the importance of freedom and diversity of thought, civil disobedience, and 'praise for defiance' (Arnold, 2015; Ripley, 2016). Instead of imposing a particular comprehensive moral doctrine, a fundamental recognition of the worth of value pluralism and of moral disagreement and debate should be at the forefront of any moral enhancement program. From this recognition it follows that in line with the views from moral progressivists (chapter 3), an important 'guiding concern' in implementing potential moral enhancement programs is to focus on enhancing morally relevant capacities, not on moral conformity.

Moral conformity could be described as the conformity of human behaviour with social or moral norms (Civai & Ma, 2017). Instead of focusing on changing a person in such a way as to make her conform to social or moral norms, arguably a more promising approach is to focus on what it means to 'be better at being good', and what kind of capacities are needed to become 'better at being good.' Such a capacitarian and developmental approach focuses on the set of capacities that "mature moral functioning"[35] (Narvaez, 2010) or "moral intelligence"[36] (Tanner & Christen, 2014) presuppose, and asks what kind of environments would foster these capacities.

To conclude, a focus on (present and emerging) moral enhancement practices invites reflection on a range of questions that are largely missing from the current debate on moral enhancement: What goal/ purpose does the intervention serve? What political, sociological, institutional context are we dealing with? Who will likely benefit and in what way(s)? Who will risk harm, and what kind of harm? What kind of background theory of moral development and moral growth is referred to? Is the intervention justified in terms of benefits to the individual undergoing the intervention, or in terms of benefits to relevant others/ public safety/ common good – or both? Are all those

35 In short, mature moral functioning refers to "individual capacities for habituated empathic concern and moral metacognition—moral locus of control, moral self-regulation, and moral self-reflection—comprise mature moral functioning, which also requires collective capacities for moral dialogue and moral institutions" (Narvaez, 2010, p. 163).

36 Tanner and Christen define moral intelligence as "the agent's capacity to process and manage moral problems" (Tanner & Christen, 2014, p. 120).

involved aware of the attempt to morally enhance, especially the person undergoing the intervention; has the person given her *informed consent*? Do all those involved agree with the goal and the means, especially the person undergoing the intervention? Will the person undergoing the intervention as a result of the intervention most likely be better at being good, or will he more reliably do the good thing?

The reader might find these questions slightly banal and deceptively simple. Yet, in order to focus attention on the fundamental issues concerning individual freedom and right to self-realisation that are at stake in current or emerging moral enhancement practices, these are, I think, exactly the kinds of questions that need to be asked.

Strengths and limitations

As both the debate and the implications of the science of morality are far from settled, it is also far from clear what the central moral issues are. In imagining potential domains in which elements of moral enhancement are recognizable, the main goal of this thesis has been to identify ethical issues that are not central in the debate now.

In response to an article in which we suggested a number of domains that we think are relevant for the debate on moral enhancement (chapter 3), bioethicist Bert Gordijn has formulated a range of reservations with respect to our suggested set of practices (Gordijn, 2017). His main concern is that the success of our strategy "is predicated on the correctness of suppositions about future moral bioenhancement scenarios. Unfortunately, however, these assumptions are inherently speculative. It is fundamentally problematic to determine in which contexts moral enhancements might first be implemented. It is equally challenging to identify contemporary domains that might, in their central aspects, be comparable to future moral bioenhancement practices" (Gordijn, 2017, p. 427).

Gordijn is correct that we cannot know for sure whether the domains we have identified will indeed be relevant for future moral enhancement scenarios. Our choice of practices might indeed turn out to be irrelevant for future moral enhancement scenarios. However, Gordijn does not go into the question why our particular choice of practices is wrong, or unlikely to be relevant for moral enhancement in the future. We have argued that elements of moral enhancement/improvement are present within current forensic practices, depending on how one defines moral enhancement – Gordijn has not refuted that. Moreover, our central claim and position that moral enhancement

should be analysed *contextually* and in relation to particular practices, still stands, even if our choice of practices turns out to be mistaken.

Without any doubt a range of other practices might be relevant and worth considering further. "Impartiality" moral enhancement, for example, could be offered to those who are in a position of power or fulfil an influential role, with the aim of nullifying implicit biases.[37] Douglas discusses studies of biological influences on fairness related behaviour (Douglas, 2015, p. 31). One often-cited study shows that judges' parole decisions vary dependent on extraneous variables, like for example how long ago the judge had a (food) break, letting the authors conclude that "justice is what the judge ate for breakfast" (Danziger et al., 2011). Neutralizing these extraneous influences would arguably make judicial sentencing fairer. Whether judges would in fact voluntary seek such impartiality enhancing interventions, or whether some sort of persuasion or coercion would be necessary (and justified) merits further discussion. Moreover, one would still have to determine who would be the fairer judge: the hungry of the satisfied one?

An additional potential concern is the fact that the working definition of moral enhancement as it was formulated in the introduction is very broad and remains silent on a number of arguably important distinctions. As the debate itself was the main subject of research, I did not want to exclude potential perspectives and practices in advance. However, such a broad interpretation has disadvantages as well. For example, a consequence of analyzing a diversity of practices through a moral enhancement lens might be that conceptualizing these practices as 'moral enhancement' practices could unintentionally promote their acceptance (chapter 5).

Concluding reflections and suggestions for further research

To enhance the debate, in my opinion it would be necessary to direct further research towards the following themes and questions:

Medicalization of immoral behaviour

Approaching violence (D. J. Williams & Donnelly, 2014) or other examples of arguably immoral behaviour from a health instead of a criminal justice lens, is an example of medicalization: quite literally, 'to make medical.' Medicalization is understood as the process by which previously nonmedical problems become defined and treated

37 One could think of judges, business and political leaders, parents, police and military personnel, programmers, American presidents, etc.

as medical problems, usually as diseases or disorders (Conrad, 2013, p. 196). Another example concerns the question whether addiction should be considered an illness or a moral failing (Wiseman, 2016). Although processes of medicalization often elicit a negative connotation the concept is, in principle, value neutral. There are both (sociological) studies *describing* processes of medicalization as well as critical analyses of instances of over- or under medicalization, and merely coining the term does not in itself differentiate between good and bad forms of medicalization (Conrad, 2013, pp. 1199-1200; Parens, 2013, p. 28; Horstkötter et al., 2015).

In the case of medicalization of violence, problematic, violent behaviour is defined, or redefined, as a medical problem (as a psychiatric disorder for instance). Medicalization of deviant or violent behaviour can have positive and negative effects. On the one hand, it can open opportunities for adequate care and treatment, where these were previously lacking. Instead of responding to this kind of behaviour with indignation and punishment, the objective of treatment is to address underlying illness or disorder and to limit chances of reoffending.

A possible worry is that by viewing (potential) offenders foremost from a medical angle, they are more likely to be viewed as basically incompetent, or at the mercy of forces beyond themselves:

> When we argue, say, against the medicalization of badness – e.g., against treating criminal behaviour as the symptom of a psychiatric disorder – we are arguing against the view of ourselves as objects at the mercy of forces beyond ourselves, and for the view of ourselves as subjects who can choose. (Parens, 2013, p. 29)

Another concern with respect to health approaches to violent behaviour concerns individualization: the risk that causes are (quite literally) sought within the individual, as a genetic, neurological, biological disorder, and to a far lesser extent in social, institutional, familiar, or other environmental circumstances. Further conceptual and ethical research is necessary here to consider the justification, and the benefits and risks of medicalization of 'badness'. Such research should ideally be performed by psychological and psychiatric forensic scientists and professionals, in close collaboration with ethicists, anthropologists, and sociologists.[38]

38 The perspectives of offenders themselves should also be included. Jeremy Dixon (Dixon, 2018) (Glover, 2014)

The roles and responsibilities of medical doctors. What should doctors do?

The debate on moral enhancement has proceeded without much attention for the kinds of institutional contexts in which moral enhancement would be implemented – or is already implemented – and the implications of those contexts in terms of professional responsibilities of those who would be responsible for the implementation.

On the one hand, perhaps biomedically trained professionals ('clinical technicians') could administer and monitor the technical aspects of certain interventions. At the same time however, focussing on the technical aspects exclusively, might come at the expense of attention for more comprehensive care aspects of the interaction. Perhaps a new sub discipline would emerge, in which hybrid doctors would move in between the domains of medicine (neuroscience, psychiatry), public health (prevention of health-related risks, safety, monitoring), and public policy (safety, crime prevention). Such a professional would be well-established in a diversity of domains, and would ideally be able to do justice to the specific responsibilities flowing from these diverse domains. There is a risk, however, that these diverse domains yield dual, multiple roles, and diffuse responsibilities (chapter 3 and 5). The merging of care and safety responsibilities warrants more fundamental reflection on the roles and responsibilities of medical professionals, as well as on the limits of those roles. This is true for forensic professionals in particular. Yet, forensic medical practice has thus far not received as much attention from medical ethics as would be warranted by the moral complexities of this practice.

Changing the individual, and/or changing the environment? Parallels with other debates

In order to stimulate people to act morally, or in accordance with social norms, a diversity of strategies can be employed. The debate on moral enhancement focuses on strategies that are aimed at changing the person herself (Klincewicz, 2016). Glenn Cohen for example tentatively considers the option of changing the biological makeup of human beings in such a way as to reduce serious human rights violations. He argues that "instead of merely crafting laws and setting up structures that get human beings such as they are to respect human rights, that the human rights approach should also consider embracing attempts to remake human beings (and more specifically human brains) into the kinds of things that are more respectful of human rights law" (Cohen, 2015, p. 1). In various other debates a variety of strategies that focus on changing the environment are discussed.[39]

39 These strategies could be used together, in concert, as is often done, for example in public health, where interventions are targeted at both the individual and the environment.

There have been proposals to change the so-called 'choice architecture' of the environment in order to 'nudg' people towards pro-social instead of egoistic choices (Capraro et al., 2017)[40], for instance by designing robots to serve as "moral nudgers" in order to foster "socially just" tendencies in humans (Borenstein & Arkin, 2016). Ismaili M'hamdi and colleagues have coined the concept of "other-regarding" nudges, in other to describe cases where "the principal but not necessarily sole beneficiary of the nudge is not the nudgee," but someone else (Ismaili M'hamdi et al., 2017, p. 702). Another example is the field of robot ethics, debating attempts to design "morally competent robots" (Malle, 2016).[41] Comparable to discussions about to who responsibility should be assigned to when a self-driving car ends up in an accident – the driver/passenger or the driver/car – one can ask who is responsible when a morally enhanced person goes astray. Who is responsible when the God Machine derails?

Another example concerns speculations about (and first experiments with) using so-called wearables to provide persons with insight as to biological processes underlying their activities and behaviour. The Empatica wristband for example supposedly can provide biofeedback about actual aggression and stress levels. This information can be used for self-regulation purposes, but could also be used for monitoring and surveillance, for example in secure prison contexts (de Kogel & Cornet, 2016; van Hintum, 2018). Such proposals should be reflected on in relation to existing efforts to monitor behaviour. To give one example, the Chinese government is experimenting with a social credit scheme through which citizens can earn (and lose) points and accompanying privileges based on online and offline behaviour (Creemers, 2017).[42] Data sources that are included in the current experiments include online behaviour, financial information, and behaviour in traffic and in public transport gathered by face-recognition

40 Nudges has been defined as "approaches that steer people in particular directions" (Sunstein, 2017, p. 4) while preserving freedom of choice.

41 In thinking about what a morally competent robot would look like, this field of robot ethics proves an interesting testing ground for thinking about the elements that are needed for "full" human moral agency.

42 "an ambitious proposed social credit scheme is intended to create a range of benefits and sanctions for online and offline behaviour. According to State Council plans, it is intended that social credit information will be connected with individuals' identity card numbers, creating unique and traceable files that can be used to facilitate citizens' access to financial and government services. At the same time, the plan called for the introduction of real-name identity-based appraisal and scoring of individual online comportment, as well as of blacklists for those perpetrating various kinds of fraud, deception and 'harm to others' lawful rights and interests'. In other words, it is not unlikely that undesired behaviour online may affect citizens' ability to gain a livelihood, find schools for their children or take out insurance." (Creemers, 2017, p. 97) See this webpage for a game that simulates the workings of the social credit scheme: https://app.nos.nl/social-credit-score/index.html.

cameras (Nieuwsuur, 2018). It is not unthinkable that future social credit schemes would include biomedical data as well.[43]

The common denominator in these discussions is reflection on a broadening range of technologies that are or can potentially be used to make people behave in accordance with social norms. In light of these parallel debates, one important direction for future research would be a more comprehensive ethical analysis of the diversity or "omnipresence"[44] of ways in which societal actors can stimulate, incentivize, nudge, or force persons towards certain behaviour, and the accompanying ethical concerns. All these emerging practices require moral scrutiny, and many of the relevant considerations will be common to all these debates. The 'moral enhancement debate' can function as a useful resource here.

Challenges and the relevance of the new sciences of morality

I find the findings of the new sciences of morality fascinating, in particular the ways this new and rich landscape of empirical studies on animal and human morality can challenge and enrich both philosophical and folk understandings of morality. These disciplines presuppose, paint, promote, and also undermine certain ideas of human nature. The debate on moral bioenhancement in a way represents an extreme example of this challenge: Is human morality to be understood on individual/biological level, and should neuro/bio-insight be taken to imply that our self-understanding needs an 'upgrade', and if so, in what ways?

It is notable that images of man and of human nature as they are usually being painted by proponents (e.g. transhumanist) of human enhancement are generally rather positive. They tend to focus on fundamental human capacities for imagination, self-creation, and empowerment. Proponents of moral enhancement on the other hand tend to depart from this positive depiction of human nature, and stress the ways human moral nature is fundamentally defective and for that reason needs a (preferably quick) 'fix'. The truth no doubt lies somewhere in the middle.[45]

---

43 See (Singh & Sinnott-Armstrong, 2014).

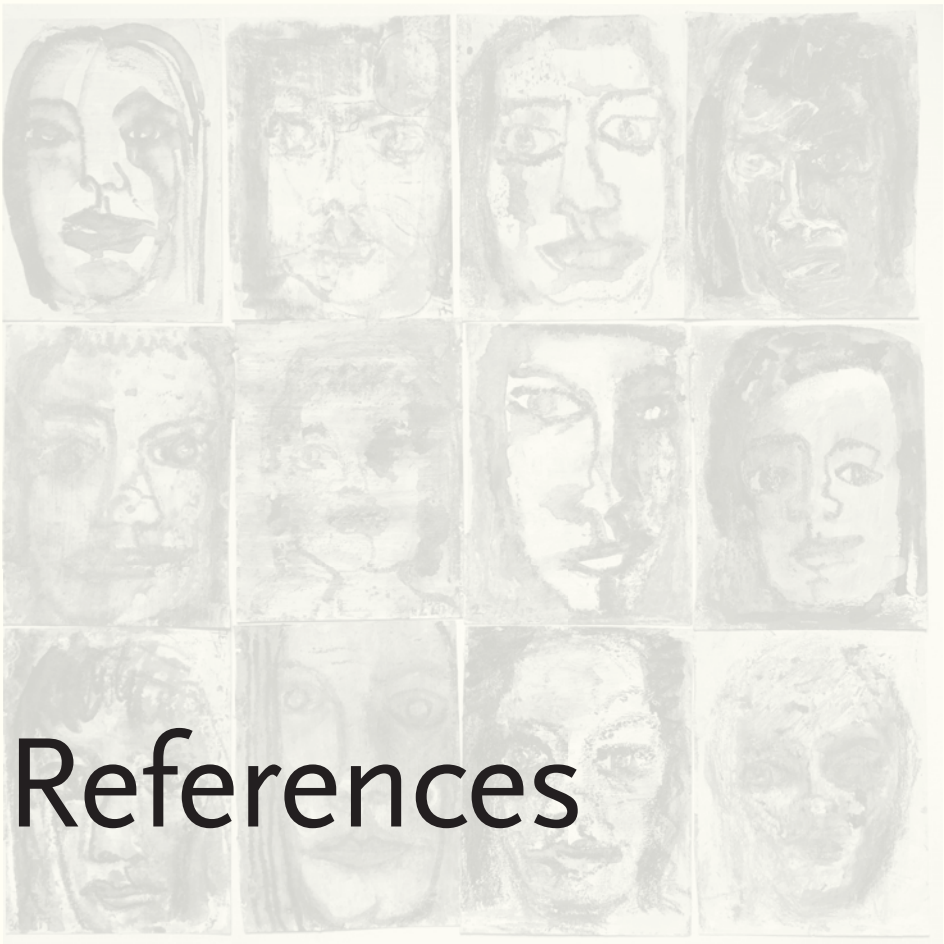
44 Such a comprehensive analysis is comparable to the argument that the ethical evaluation of health checks should go beyond weighing the harms and benefits of individual tests, but should take into account the potential beneficial and harmful effects of what has been termed an "omnipresence of health checks" (Stol et al., 2017), that is, the combined effects resulting from the multitude of tests on offer.

45 Are human beings inherently and thoroughly selfish, with morality only being an after-thought or thin, cultural layer of "veneer" around a corrupt core; or are human capacities for morality antique and part of our genetic inheritance (de Waal, 2009, pp. 7-12)?

We should do justice to human capacities for giving and exchanging reasons for actions (Sie, 2014), holding and giving responsibility, while at the same time "biting the bullet" by acknowledging the diversity of challenges (Levy, 2006). Despite a range of challenges, based on fascinating studies explaining the varieties of ways human moral behaviour, moral thought, and moral emotions are influenced by situationalist, environmental, genetic, and biological influences, we should be careful to not rush towards – debunking – conclusions.[46] Humans arguably have evolved into the kind of beings to whom 'moralizing' comes natural; we tend to evaluate ourselves and others in moral terms. As Harris writes; "We have certainly evolved to have a vigorous sense of justice and right, that is, with a virtuous sense of morality" (Harris, 2011, p. 104). To reconcile these different ways of thinking about human morality is a challenge for science and philosophy: for the explanatory models and their underlying presuppositions. But it is a challenge also for folk conceptions of the moral self/ moral identity. Along with providing new input or challenges to various conceptual and scientific *puzzles*, studies on human morality arguably can impact the way we view ourselves as moral agents as well.

To the extent that 'the new sciences of morality' indeed open up 'the black box' of human moral psychology, the question is how we want to use this new, evolving body of knowledge. If we truly are coming closer to understanding the many influences on human moral decision making and behaviour, the question is how this knowledge is interpreted and translated. What in my view is paradoxical is that in many discussions we quickly move in the direction of applications where the insight is not distributed more widely/ democratically, but is apparently limited to an unidentified 'we', a small group (scientists, moral philosophers, bioethicists), who discuss how that new knowledge can be used to change 'people'. Many of the applications that are discussed (nudging, using robots, environmental intervention, virtual reality, etc.) seem to be conceived, designed from 'above', with little consideration of the possible ways in which people themselves could (or would like to) benefit from this knowledge. Future research should focus on ways in which the new sciences of morality could indeed empower people to become better at being good.

46 Why is it that the realization that we typically/ often 'bypass reasons' often perceived as disturbing? Because, philosopher Regina Rini argues; "Doxastic embarrassment results from my awareness of a gap between the considerations that seem correct to me in my conscious thought, and the factors that actually drive my automated moral beliefs. What psychological research exposes is a form of disunity in my functioning as a moral agent. My conscious, reflective self is not appropriately unified with my automated, effective self" (Rini, 2016, p. 1449).



References

A

Abend, G. (2013). What the science of morality doesn't say about morality. *Philosophy of the Social Sciences, 43*(2), 157-200.

Adshead, G., & Davies, T. (2016). Wise restraints: Ethical issues in the coercion of forensic patients. In B. Völlm & N. Nedopil (Eds.), *The Use of Coercive Measures in Forensic Psychiatric Care: Legal, Ethical and Practical Challenges* (pp. 69-86). Cham: Springer International Publishing.

Aeon. (2015). If we one day invent a pill for morality, would it be immoral to withhold it from someone? Retrieved April 9, 2018, from https://aeon.co/conversations/if-we-one-day-invent-a-pill-for-morality-would-it-be-immoral-to-withhold-it-from-someone-171.

Agar, N. (2010). Enhancing genetic virtue? *Politics and the Life Sciences, 29*(1), 73-75.

Agar, N. (2014). A question about defining moral bioenhancement. *Journal of Medical Ethics, 40*(6), 369-370.

Agar, N. (2015a). Moral bioenhancement and the utilitarian catastrophe. *Cambridge Quarterly of Healthcare Ethics, 24*(1), 37-47.

Agar, N. (2015b). Moral bioenhancement is dangerous. *Journal of Medical Ethics, 41*(4), 343-345.

Aggarwal, N. K. (2009). Neuroimaging, culture, and forensic psychiatry. *The Journal of the American Academy of Psychiatry and the Law,* 37(2), 239-244.

Aharoni, E., Funk, C., Sinnott-Armstrong, W., & Gazzaniga, M. (2008). Can neurological evidence help courts assess criminal responsibility? Lessons from law and Neuroscience. *Annals of the New York Academy of Sciences, 1124*(1), 145-160.

Aharoni, E., Vincent, G. M., Harenski, C. L., Calhoun, V. D., Sinnott-Armstrong, W., Gazzaniga, M. S., & Kiehl, K. A. (2013). Neuroprediction of future rearrest. *Proceedings of the National Academy of Sciences USA, 110*(15), 6223-6228.

American Psychological Association Task Force on Appropriate Therapeutic Responses to Sexual Orientation. (2009). Report of the American Psychological Association Task Force on Appropriate Therapeutic Responses to Sexual Orientation. Retrieved April 14, 2018, from http://www.apa.org/pi/lgbc/publications/therapeutic-resp.html.

Appelbaum, P. S. (1990). The parable of the forensic psychiatrist: Ethics and the problem of doing harm. *International Journal of Law and Psychiatry, 13*(4), 249-259.

Appelbaum, P. S. (1997). A theory of ethics for forensic psychiatry. *Journal of the American Academy of Psychiatry and the Law, 25*(3), 233-247.

Appiah, A. (2008). *Experiments in Ethics*. Cambridge, MA: Harvard University Press.

Arnhart, L. (2010). Can virtue be genetically engineered? *Politics and the Life Sciences, 29*(1), 79-81.

Arnold, C. (2015, December 18). In praise of defiance. *AEON*. Retrieved December 1, 2017, from https://aeon.co/essays/we-should-listen-to-the-defiant-not-diagnose-and-medicate-them.

Asherson, P., & Cormand, B. (2016). The genetics of aggression: Where are we now? *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 171*(5), 559-561.

Aspinwall, L. G., Brown, T. R., & Tabery, J. (2012). The double-edged sword: Does biomechanism increase or decrease judges' sentencing of psychopaths? *Science, 337*(6096), 846-849.

Atkinson, R., & Flint, J. (2004). Snowball sampling. In M. S. Lewis-Beck, A. Bryman, & T. Futing Liao (Eds.), *The SAGE Encyclopedia of Social Science Reserach Methods* (Vol. 3, pp. 1043-1044). Thousand Oaks: Sage Publications, Inc.

B

Baccarini, E., & Malatesti, L. (2017). The moral bioenhancement of psychopaths. *Journal of Medical Ethics, 43*(10), 697-701.

Baertschi, B. (2014). Neuromodulation in the service of moral enhancement. *Brain Topography, 27*(1), 63-71.

Baron-Cohen, S. (2011). *Zero Degrees of Empathy: A New Theory of Human Cruelty. London*: Penguin UK.

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena *The Social Neuroscience of Empathy* (pp. 3-15). The MIT Press.

Baum, M. L. (2013). The monoamine oxidase A (MAOA) genetic predisposition to impulsive violence: Is it relevant to criminal trials? *Neuroethics, 6*(2), 287-306.

Baylis, F. (2013). "I am who I am": On the perceived threats to personal identity from deep brain stimulation. *Neuroethics, 6*(3), 513-526.

Bazeley, P., & Jackson, K. (2013). *Qualitative Data Analysis with NVivo. London*: Sage Publications Limited.

Beauchamp, T. L. (2015). Are we unfit for the future? *Journal of Medical Ethics, 41*(4), 346-348.

Beaver, K. M., Schwartz, J. A., & Gajos, J. M. (2015). A review of the genetic and gene–environment interplay contributors to antisocial phenotypes. In J. Morizot & L. Kazemian (Eds.), *The Development of Criminal and Antisocial Behavior: Theory, Research and Practical Applications* (pp. 109-122). Cham: Springer International Publishing.

Beck, B. (2015). Conceptual and practical problems of moral enhancement. *Bioethics, 29*(4), 233-240.

Berdik, C. (2017, February 2). Can virtual reality teach empathy? *Slate*. Retrieved February 16, 2017, from http://www.slate.com/articles/technology/future_tense/2017/02/can_virtual_reality_in_the_classroom_help_teach_empathy.html.

Bergström, L. S., & Lynöe, N. (2008). Enhancing concentration, mood and memory in healthy individuals: an empirical study of attitudes among general practitioners and the general population. *Scandinavian Journal of Public Health, 36*(5), 532-537.

Berkvens, J. (2013, June 19). Biologische oorzaken criminaliteit: Affaire Buikhuisen. [Biological causes of crime: The affair Buikhuisen], *EenVandaag*. Nederland.

Berryessa, C. M., Chandler, J. A., & Reiner, P. B. (2016). Public attitudes toward legally coerced biological treatments of criminals. *Journal of Law and the Biosciences, 3*(3), 447-467.

Blackford, R. (2010). Genetically engineered people. *Politics and the Life Sciences, 29*(1), 82-84.

Blank, R. H. (2007). Policy implications of the new neuroscience. *Cambridge Quarterly of Healthcare Ethics, 16*(02), 169-180.

Bloemink, S. (2013, February 2). In de amygdala-hoek. Het verzet tegen de breincultuur groeit. [In the amygdala corner. The resistance to brain culture is growing]. *De Groene Amsterdammer*.

Boodman, S. G. (2015, March 15). How to teach doctors empathy. *The Atlantic*. Retrieved March 20, 2018, from http://www.theatlantic.com/health/archive/2015/03/how-to-teach-doctors-empathy/387784/.

Borenstein, J., & Arkin, R. (2016). Robotic nudges: The ethics of engineering a more socially just human being. *Science and Engineering Ethics, 22*(1), 31-46.

Borry, P., Schotsmans, P., & Dierickx, K. (2005). The birth of the empirical turn in bioethics. *Bioethics, 19*(1), 49-71.

Bostrom, N. (2013). Existential Risk Prevention as Global Priority. *Global Policy, 4*(1), 15-31.

Brey, P. A. E. (2012). Anticipatory ethics for emerging technologies. *NanoEthics, 6*(1), 1-13.

Bronstein, J. (2010). Objecting to the Genetic Virtue Program. *Politics and the Life Sciences, 29*(1), 85-87.

Brooks, T. (2012). Moral Frankensteins. *AJOB Neuroscience, 3*(4), 28-30.

Bruneau, E. (2015). Putting neuroscience to work for peace. In E. Halperin & K. Sharvit (Eds.), *The Social Psychology of Intractable Conflicts* (Vol. 27, pp. 143-155). Cham: Springer.

Bruni, T. (2011). The ambivalence of moral psychology. *AJOB Neuroscience, 2*(4), 13-15.

Bublitz, J. C. (2015). Moral enhancement and mental freedom. *Journal of Applied Philosophy, 33*, 88-106.

Bublitz, J. C. (2016). Saving the world through sacrificing liberties? A critique of some normative arguments in Unfit for the Future. *Neuroethics,* 10.1007/s12152-016-9265-8.

Bublitz, J. C., & Merkel, R. (2014). Crimes against minds: On mental manipulations, harms and a human right to mental self-determination. *Criminal Law and Philosophy, 8*(1), 51-77.

Buchanan, A., & Grounds, A. (2011). Forensic psychiatry and public protection. *The British Journal of Psychiatry, 198*(6), 420-423.

Bunnik, E. M., Janssens, A. C. J. W., & Schermer, M. H. N. (2014). Informed consent in direct-to-consumer personal genome testing: The outline of a model between specific and generic consent. *Bioethics, 28*(7), 343-351.

Burstin, K., Doughtie, E. B., & Raphaeli, A. (1980). Contrastive Vignette Technique: An indirect methodology designed to address reactive social attitude measurement. *Journal of Applied Social Psychology, 10*(2), 147-165.

C

Cabrera, L. Y., Fitz, N. S., & Reiner, P. B. (2014). Reasons for comfort and discomfort with pharmacological enhancement of cognitive, affective, and social domains. *Neuroethics, 8(2)*, 93-106.

Cabrera, L. Y., & Reiner, P. B. (2016). A novel sequential mixed-method technique for contrastive analysis of unscripted qualitative data: Contrastive quantitized content analysis. *Sociological Methods and Research,* 10.1177/0049124116661575.

Campbell, E., & Ross, L. (2004). Attitudes of healthcare professionals and parents regarding genetic testing for violent traits in childhood. *Journal of Medical Ethics, 30*(6), 580-586.

Capraro, V., Jagfeld, G., Klein, R., Mul, M., & van de Pol, I. (2017). What's the right thing to do? Increasing pro-sociality with simple moral nudges.

Carter, J. A., & Gordon, E. C. (2015). On cognitive and moral enhancement: A reply to Savulescu and Persson. *Bioethics, 29*(3), 153-161.

Carter, S. (2015). Putting a price on empathy: Against incentivising moral enhancement. *Journal of Medical Ethics, 41*(10), 825-829.

Carter, S. (2016). Could moral enhancement interventions be medically indicated? *Health Care Analysis, 25*(4), 338-353.

Casal, P. (2015). On not taking men as they are: Reflections on moral bioenhancement. *Journal of Medical Ethics, 41*(4), 340-342.

Chan, S., & Harris, J. (2011). Moral enhancement and pro-social behaviour. *Journal of Medical Ethics, 37*(3), 130-131.

Chandler, J. A. (2016). The use of neuroscientific evidence in Canadian criminal proceedings. *Journal of Law and the Biosciences, 2*(3), 550-579.

Chhangur, R. R., Weeland, J., Matthys, W., & Overbeek, G. (2015). Gene by environment research to prevent externalizing problem behavior: Ethical questions raised from a public healthcare perspective. *Public Health Ethics, 8*(3), 295-304.

Christen, M., & Narvaez, D. (2012). Moral development in early childhood is key for moral enhancement. *AJOB Neuroscience, 3*(4), 25-26.

Churchland, P. S. (2011). *Braintrust: What neuroscience tells us about morality. Princeton*: Princeton University Press.

Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations An integrative review. *Perspectives on Psychological Science, 9*(3), 245-274.

Civai, C., & Ma, I. (2017). The enhancement of social norm compliance: Prospects and caveats. *Journal of Cognitive Enhancement, 1*(1), 26-30.

Clausen, J. (2010). Ethical brain stimulation–neuroethics of deep brain stimulation in research and clinical practice. *European Journal of Neuroscience, 32*(7), 1152-1162.

Cohen, I. G. (2015). This is your brain on human rights: Moral enhancement and human rights. *The Law & Ethics of Human Rights, 9*(1), 1-41.

Cohen Kadosh, K., Lisk, S., & Lau, J. Y. F. (2016). The ethics of (neuro) feeding back to the developing brain. *AJOB Neuroscience, 7*(2), 132-133.

Conrad, P. (2013). Medicalization: Changing contours, characteristics, and contexts. In W. Cockerham (Ed.), *Medical Sociology on the Move* (pp. 195-214). Dordrecht: Springer.

Cornet, L. J. M., de Kogel, C. H., Nijman, H. L. I., Raine, A., & van der Laan, P. H. (2013). Neurobiological factors as predictors of cognitive–behavioral therapy outcome in individuals with antisocial behavior: A review of the literature. *International Journal of Offender Therapy and Comparative Criminology, 58*(11), 1279-1296.

Creemers, R. (2017). Cyber China: Upgrading propaganda, public opinion work and social management for the twenty-first century. *Journal of Contemporary China, 26*(103), 85-100.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). London: Sage Publications, Inc.

Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology, 111*(4), 636-653.

Crockett, M. J. (2014a). Moral bioenhancement: A neuroscientific perspective. *Journal of Medical Ethics, 40*(6), 370-371.

Crockett, M. J. (2014b, June 3). Morality pills: Reality or science fiction? *The Guardian*. Retrieved April 12, 2018, from https://www.theguardian.com/science/headquarters/2014/jun/03/morality-pills-reality-or-science-fiction.

Crockett, M. J. (2014c). Pharmaceutical effects on moral behavior: A neuroscientific perspective. *Philosophy, Psychiatry, & Psychology, 21*(2), 131-134.

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010a). Reply to Harris and Chan: Moral judgment is more than rational deliberation. *Proceedings of the National Academy of Sciences, 107*(50), E184-E184.

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010b). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences, 107*(40), 17433-17438.

Curtis, B. L. (2012). Moral enhancement as rehabilitation? *AJOB Neuroscience, 3*(4), 23-24.

D

Damasio, A. R. (2000). *Descartes' error: Emotion, reason, and the human brain*. New York: Quill.

Danaher, J. (2013, July 14). Moral enhancement and superficiality: compassion pills. *IEET Blog*. Retrieved December 5, 2016, from http://ieet.org/index.php/IEET/more/danaher20130714.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences, 108*(17), 6889-6892.

Darragh, M., Buniak, L., & Giordano, J. (2015). A four-part working bibliography of neuroethics: part 2-neuroscientific studies of morality and ethics. *Philosophy, Ethics, and Humanities in Medicine, 10*(2).

Davies, R., Ives, J., & Dunn, M. (2015). A systematic review of empirical bioethics methodologies. *BMC medical ethics, 16*(1), 15.

Day, A., & Casey, S. (2009). Values in forensic and correctional psychology. *Aggression and Violent Behavior, 14*(4), 232-238.

de Haan, W. J. M. (2009). "Wat nooit had gemogen". Perverse effecten van een essay over chronische delinquentie. ["What should have never happened". Perverse effects of an essay on chronic delinquency]. *Delikt en Delinkwent, 40*(6), 527-538.

de Kogel, C. H., & Cornet, L. J. M. (2016). Toepassingsmogelijkheden van Quantified Self-data. Enkele voorbeelden uit de forensisch psychiatrische praktijk. [Potential applications of Quantified Self-Data. Some examples from forensic psychiatric practice]. *Justitiele Verkenningen, 42*(1), 79-94.

de Kogel, C. H., & Westgeest, E. J. M. C. (2015). Neuroscientific and behavioral genetic information in criminal cases in the Netherlands. *Journal of Law and the Biosciences, 2*(3), 580-605.

de Mey, T. (2006). Imagination's grip on science. *Metaphilosophy, 37*(2), 222-239.

de Ridder, D., Langguth, B., Plazier, M., & Menovsky, T. (2009). Moral Dysfunction: Theoretical Model and Potential Neurosurgical Treatments. In J. Verplaetse, J. Schrijver, S. Vanneste, & J. Braeckman (Eds.), *The Moral Brain: Essays on the Evolutionary and Neuroscientific Aspects of Morality* (pp. 155-183). Dordrecht: Springer Netherlands.

de Vrieze, J. (2017, February 22). Eigen volk... eh, tweede? Het morele brein. [Own people.. eh, second? The moral brain]. *De Groene Amsterdammer*.

de Waal, F. (1996). *Good Natured. The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

de Waal, F. (2009). *Primates and Philosophers: How Morality Evolved*. Princeton, NJ: Princeton University Press.

de Waal, F., & Sherblom, S. A. (2018). Bottom-up morality: The basis of human morality in our primate nature. *Journal of Moral Education, 47*(2), 248-258.

Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences, 18*(7), 337-339.

Decety, J., & Wheatley, T. (Eds.). (2015). *The Moral Brain. A Multidisciplinary Perspective*. Cambridge, MA: MIT Press.

Decety, J., & Yoder, K. J. (2017). The emerging social neuroscience of justice motivation. *Trends in Cognitive Sciences, 21*(1), 6-14.

DeGrazia, D. (2005). Enhancement technologies and human identity. *Journal of Medicine and Philosophy, 30*(3), 261-283.

DeGrazia, D. (2014). Moral enhancement, freedom, and what we (should) value in moral behaviour. *Journal of Medical Ethics, 40*(6), 361-368.

Dehue, T. (2011, November 2). Ik maak drukte want ik ben een druktemaker. De medicalisering van 'ongewenst' gedrag. [I make noice because I am a noicy fellow. The medicalisation of 'unwanted' behaviour]. *De Groene Amsterdammer*.

Dehue, T. (2014). *Betere Mensen [Better People]*. Amsterdam: Atlas Contact.

Dennett, D. C. (2013, April 24). Are criminals born or made? *Prospect Magazine*. Retrieved July 27, 2017, from http://www.prospectmagazine.co.uk/arts-and-books/the-anatomy-of-violence-adrian-raine-daniel-dennett-review.

Dijkstra, A. M., & Schuijff, M. (2015). Public opinions about human enhancement can enhance the expert-only debate: A review study. *Public Understanding of Science, 25*(5), 588-602.

Dixon, J. (2018). Narratives of illness and offending: mentally disordered offenders' views on their offending. *Sociology of Health & Illness,* 10.1111/1467-9566.12740.

Dooley, J. J., Pyzalski, J., & Cross, D. (2009). Cyberbullying versus face-to-face bullying. A theoretical and conceptual review. *Zeitschrift Fur Psychologie-Journal of Psychology, 217*(4), 182-188.

Douglas, T. (2008). Moral enhancement. *Journal of Applied Philosophy, 25*(3), 228-245.

Douglas, T. (2013). Moral enhancement via direct emotion modulation: A reply to John Harris. *Bioethics, 27*(3), 160-168.

Douglas, T. (2014a). Criminal rehabilitation through medical intervention: Moral liability and the right to bodily integrity. *The Journal of Ethics, 18*(2), 101-122.

Douglas, T. (2014b). Enhancing moral conformity and enhancing moral worth. *Neuroethics, 7*(1), 75-91.

Douglas, T. (2014c). The morality of moral neuroenhancement. In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics* (pp. 1227-1249). Dordrecht: Springer Netherlands.

Douglas, T. (2015). The harms of enhancement and the conclusive reasons view. *Cambridge Quarterly of Healthcare Ethics, 24*(01), 23-36.

Dressing, H., Salize, H. J., & Gordon, H. (2007). Legal frameworks and key concepts regulating diversion and treatment of mentally disordered offenders in European Union member states. *European Psychiatry, 22*(7), 427-432.

Dubljević, V. (2017, May 30). "Moral enhancement" is science fiction, not science fact. *Scientific American*.

Dubljević, V., & Racine, E. (2017). Moral enhancement meets normative and empirical reality: Assessing the practical feasibility of moral enhancement neurotechnologies. *Bioethics, 31*(5), 338-348.

E

Edworthy, R., Sampson, S., & Völlm, B. (2016). Inpatient forensic-psychiatric care: Legal frameworks and service provision in three European countries. *International Journal of Law and Psychiatry, 47*, 18-27.

Eerkens, M. (2017, December 8). Hoe maken we onze kinderen empathischer? [How can we make our children more empathetic?]. *De Correspondent*.

Ehni, H.-J., & Aurenque, D. (2012). On moral enhancement from a Habermasian perspective. *Cambridge Quarterly of Healthcare Ethics, 21*(2), 223-234.

Eichelberger, R., & Barnes, J. C. (2015). Biosocial criminology. In W. G. Jennings (Ed.), *The Encyclopedia of Crime and Punishment* (10.1002/9781118519639.wbecpx185). John Wiley & Sons, Inc.

F

Fallon, J. (2013). *The Psychopath Inside: A Neuroscientist's Personal Journey Into the Dark Side of the Brain. New York*: Penguin.

Farahany, N. A. (2016). Neuroscience and behavioral genetics in US criminal law: an empirical analysis. *Journal of Law and the Biosciences, 2*(3), 485-509.

Faust, H. S. (2008). Should we select for genetic moral enhancement? A thought experiment using the MoralKinder (MK+) haplotype. *Theoretical Medicine and Bioethics, 29*(6), 397-416.

Feinberg, J. (1998). The child's right to an open future. In P. H. Hirst & P. White (Eds.), *Philosophy of Education: Society and Education* (Vol. 3, pp. 250-270). London and New York: Taylor & Francis.

Fenton, E. (2010). The perils of failing to enhance: A response to Persson and Savulescu. *Journal of Medical Ethics, 36*(3), 148-151.

Ferguson, C. J. (2010). Genetic contributions to antisocial personality and behavior: A meta-analytic review from an evolutionary perspective. *The Journal of Social Psychology, 150*(2), 160-180.

Fitz, N. S., & Reiner, P. B. (2013). The challenge of crafting policy for do-it-yourself brain stimulation. *Journal of Medical Ethics, 41*(5), 410-412.

Focquaert, F. (2013). Deep brain stimulation in children: Parental authority versus shared decision-making. *Neuroethics, 6*(3), 447-455.

Focquaert, F. (2014). Mandatory neurotechnological treatment: ethical issues. *Theoretical Medicine and Bioethics, 35*(1), 59-72.

Focquaert, F. (2017, March 1). On the impossibility of justifying the moral responsibility system. *Syndicate*. Retrieved May 30, 2017, from https://syndicate.network/symposia/philosophy/the-stubborn-system-of-moral-responsibility/.

Focquaert, F., & Raine, A. (2012). Ethics of community-based sanctions. In S. Barton-Bellessa (Ed.), *Encyclopedia of Community Corrections* (pp. 144-148). London: SAGE Publications.

Focquaert, F., & Schermer, M. H. N. (2015). Moral enhancement: Do means matter morally? *Neuroethics, 8*(2), 139-151.

Fozdar, M. A. (2016). The Relevance of Modern Neuroscience to Forensic Psychiatry Practice. *Journal of the American Academy of Psychiatry and the Law, 44*(2), 145-150.

Frentz, S.-L., Gavaghan, C., & Schwarz, K. (2015, December 3). My genes made me do it: The problem of genetic evidence and diminished culpability. *The Conversation*.

Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education, 37*(3), 333-356.

Frith, L. (2012). Symbiotic empirical ethics: A practical methodology. *Bioethics, 26*(4), 198-206.

Fröding, B. E. E. (2011). Cognitive enhancement, virtue ethics and the good life. *Neuroethics, 4*(3), 223-234.

G

Gaudet, L. M., Kerkmans, J. P., Anderson, N. E., & Kiehl, K. A. (2016). Can neuroscience help predict future antisocial behavior? *Fordham Law Review, 85*(2), 503-531.

Ginther, M. (2016). Neuroscience or neurospeculation? Peer commentary on four articles examining the prevalence of neuroscience in criminal cases around the world. *Journal of Law and the Biosciences, 3*(2), 324-329.

Giordano, J., Kulkarni, A., & Farwell, J. (2014). Deliver us from evil? The temptation, realities, and neuroethico-legal issues of employing assessment neurotechnologies in public safety initiatives. *Theoretical Medicine and Bioethics, 35*(1), 73-89.

Glannon, W. (2014). The Limitations and Potential of Neuroimaging in the Criminal Law. *The Journal of Ethics, 18*(2), 153-170.

Glaser, B. (2009). Treaters or punishers? The ethical role of mental health clinicians in sex offender programs. *Aggression and Violent Behavior, 14*(4), 248-255.

Glaser, B. (2010). Sex offender programmes: New technology coping with old ethics. *Journal of Sexual Aggression, 16*(3), 261-274.

Glenn, A. L., Focquaert, F., & Raine, A. (2015). Prediction of antisocial behavior. In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics* (pp. 1689-1701). Dordrecht: Springer Netherlands.

Glenn, A. L., Johnson, A. K., & Raine, A. (2013). Antisocial Personality Disorder: A current review. *Current Psychiatry Reports, 15*(12), 427.

Glenn, A. L., & Raine, A. (2014). Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour. *Nature Reviews Neuroscience, 15*(1), 54-63.

Glover, J. (2014). *Alien Landscapes?: Interpreting Disordered Minds*. Cambridge, MA: Harvard University Press.

Goddard, D., McDonald, M., McCarthy-Miller, B., Benz, P., Gates, T., O'Donnell, T., . . . Schur, M. (Writers). (2016). The Good Place. M. Schur, D. Miner, M. Sackett, & D. Goddard (Producers). Los Angeles, CA: NBC.

Goldberg, D. S. (2011). Against reductionism in law & neuroscience. *Houston Journal of Health Law & Policy, 11*(2), 321-346.

Gordijn, B. (2017). Commentary: Reservations about the lessons drawn from moral education, public health ethics, and forensic psychiatry. *Cambridge Quarterly of Healthcare Ethics, 26*(3), 427-430.

Greely, H. (2007). Neuroscience and criminal justice: Not responsibility but treatment. *University of Kansas Law Review, 56*, 1103-1138.

Greely, H., Sahakian, B., Harris, J., Kessler, R. C., Gazzaniga, M., Campbell, P., & Farah, M. J. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature, 456*(7223), 702-705.

Greene, J. (2001). From neural 'is' to moral 'ought': What are the moral implications of neuroscientific moral psychology. *Nature, 411*, 953-956.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*(12), 517-523.

Gunson, D., & McLachlan, H. (2013). Risk, Russian-roulette and lotteries: Persson and Savulescu on moral enhancement. *Medicine, Health Care and Philosophy, 16*(4), 877-884.

Gurnani, T., Ivanov, I., & Newcorn, J. H. (2016). Pharmacotherapy of aggression in child and adolescent psychiatric disorders. *Journal of Child and Adolescent Psychopharmacology, 26*(1), 65-73.

H

Hagerty, B. B. (2017, June). When your child is a psychopath. *The Atlantic*.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814.

Hallgren, I. (2012). Seeing agents when we need to, attributing experience when we feel like it. *The Review of Philosophy and Psychology, 3*(3), 369-382.

Harris, J. (2011). Moral enhancement and freedom. *Bioethics, 25*(2), 102-111.

Harris, J. (2012). What it's like to be good. *Cambridge Quarterly of Healthcare Ethics, 21*(3), 293-305.

Harris, J. (2013a). 'Ethics is for bad guys!' Putting the 'moral' into moral enhancement. *Bioethics, 27*(3), 169-173.

Harris, J. (2013a). In search of blue skies: Science, ethics, and advances in technology. *Medical Law Review, 21*(1), 131-145.

Harris, J. (2013b). Moral progress and moral enhancement. *Bioethics, 27*(5), 285-290.

Harris, J. (2013b). What it is to be good. *European Review, 21*(S1), S114-S122.

Harris, J. (2014). Taking liberties with free fall. *Journal of Medical Ethics, 40*(6), 371-374.

Harris, J. (2016a). *How to Be Good: The Possibility of Moral Enhancement*. Oxford: Oxford University Press.

Harris, J. (2016b). Moral blindness – The gift of the God Machine. *Neuroethics, 9*(3), 269-273.

Harris, J., & Chan, S. (2010). Moral behavior is not what it seems. *Proceedings of the National Academy of Sciences of the United States of America, 107*(50), E183-E183.

Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. Ecco/HarperCollins Publishers.

Hauskeller, M. (2014). *Better humans?: Understanding the enhancement project*. New York, NY: Routledge.

Hauskeller, M. (2015). Being good enough to prevent the worst. *Journal of Medical Ethics, 41*(4), 289-290.

Hauskeller, M. (2016). The art of misunderstanding critics. The case of Ingmar Persson and Julian Savulescu's defense of moral bioenhancement. *Cambride Quarterly of Healthcare Ethics, 25*(1), 153-161.

Hauskeller, M. (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.

Hofmann, B. (2018). The gene-editing of super-ego. *Medicine, Health Care and Philosophy,* 10.1007/s11019-018-9836-z.

Hope, T. (1999). Empirical medical ethics. *Journal of Medical Ethics, 25*(3), 219-220.

Horstkötter, D. (2015). Forensic screening and prevention in children and adolescents: Public health ethical aspects. *Public Health Ethics, 8*(3), 266-269.

Horstkötter, D., Berghmans, R., De Ruiter, C., Krumeich, A., & de Wert, G. (2012). "We are also normal humans, you know?" Views and attitudes of juvenile delinquents on antisocial behavior, neurobiology and prevention. *International Journal of Law and Psychiatry, 35*(4), 289-297.

Horstkötter, D., Berghmans, R., & de Wert, G. (2012). Moral enhancement for antisocial behavior? An uneasy relationship. *AJOB Neuroscience, 3*(4), 26-28.

Horstkötter, D., Berghmans, R., & de Wert, G. (2014). Early prevention of antisocial behavior (ASB): A comparative ethical analysis of psychosocial and biomedical approaches. *BioSocieties, 9*(1), 60-83.

Horstkötter, D., Berghmans, R., Feron, F., & De Wert, G. (2014). "One Can Always Say No." Enriching the bioethical debate on antisocial behaviour, neurobiology and prevention: Views of juvenile delinquents. *Bioethics, 28*(5), 225-234.

Horstkötter, D., & de Wert, G. (2013). The prevention of psychopathy: What we owe to young people. *AJOB Neuroscience, 4*(2), 19-20.

Horstkötter, D., Dondorp, W., & de Wert, G. (2015). Medicalization, demedicalization and beyond: antisocial behaviour and the case of the Dutch youth law. *Public Health Ethics, 8*(3), 295-304.

Hubbeling, D. (2009). Pharmacology and human morality. *The British Journal of Psychiatry, 194*(2), 187-188.

Hübner, D., & White, L. (2016). Neurosurgery for psychopaths? An ethical analysis. *AJOB Neuroscience, 7*(3), 140-149.

Hughes, J. (2011). After happiness, cyborg virtue. *Free Inquiry, 32*(1), 34-37.

Hughes, J. (2013). Using neurotechnologies to develop virtues: A Buddhist approach to cognitive enhancement. *Accountability in Research-Policies and Quality Assurance, 20*(1), 27-41.

I

Ismaili M'hamdi, H., Hilhorst, M., Steegers, E. A. P., & de Beaufort, I. (2017). Nudge me, help my baby: On other-regarding nudges. *Journal of Medical Ethics, 43*(10), 702-706.

Ives, J. (2014). A method of Reflexive Balancing in a pragmatic, interdisciplinary and reflexive bioethics. *Bioethics, 28*(6), 302-312.

J

Jebari, K. (2014). What to enhance: Behaviour, emotion or disposition? *Neuroethics, 7*(3), 253-261.

Jefferson, W., Douglas, T., Kahane, G., & Savulescu, J. (2014). Enhancement and Civic Virtue. *Social Theory and Practice, 40*(3), 499-527.

Jones, D. G. (2013). The importance of realism in assessing technological possibilities: The role of Christian thinking. *Christian Perspectives on Science and Technology, 9*.

Jörg, P., Heino, S., & Hans, W. (2012). Dual loyalty in prison health care. *American Journal of Public Health, 102*(3), 475-480.

Jotterand, F. (2011). "Virtue engineering" and moral agency: Will post-humans still need the virtues? *AJOB Neuroscience, 2*(4), 3-9.

Jotterand, F. (2014). Questioning the moral enhancement project. *American Journal of Bioethics, 14*(4), 1-3.

Joyce, R. (2013). Unfit for the future: The need for moral enhancement. *Analysis, 73*(3), 587-589.

Juengst, E. (1998). What does 'enhancement' mean? In E. Parens (Ed.), *Enhancing Human Traits* (pp. 29-47). Washington: Georgetown University Press.

K

Kabasenche, W. P. (2012). Moral enhancement worth having: Thinking holistically. *AJOB Neuroscience, 3*(4), 18-20.

Kabasenche, W. P. (2013). Engineering for virtue? Toward holistic moral enhancement. In J. Basl & R. L. Sandler (Eds.), *Designer Biology: The Ethics of Intensively Engineering Biological and Ecological Systems* (pp. 69-86). Plymouth: Lexington Books.

Kahane, G. (2011, July 5). Would we swallow a 'morality' pill? *The Globe and Mail*. Retrieved April 12, 2018, from https://www.theguardian.com/science/headquarters/2014/jun/03/morality-pills-reality-or-science-fiction.

Kahane, G., & Savulescu, J. (2015). Normal human variation: Refocussing the enhancement debate. *Bioethics, 29*(2), 133-143.

Kahn, J. (2012, May 11). Can you call a 9-year old a psychopath? *The New York Times Magazine*.

Kaltiala-Heino, R., & Eronen, M. (2015). Ethical issues in child and adolescent forensic psychiatry: A review. *The Journal of Forensic Psychiatry & Psychology, 26*(6), 759-780.

Kass, L. (2003). *Beyond therapy: biotechnology and the pursuit of happiness* (H. Perennial Ed.): President's Council on Bioethics.

Katz, L. D. (2000). *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives* (Vol. 1). Imprint Academic.

Khan, O., Ferriter, M., Huband, N., Powney, M. J., Dennis, J. A., & Duggan, C. (2015). Pharmacological interventions for those who have sexually offended or are at risk of offending. *Cochrane Database of Systematic Reviews,* 10.1002/14651858.CD007989.pub2(2).

Kim, M., & Sankey, D. (2009). Towards a Dynamic Systems Approach to moral development and moral education: a response to the JME Special Issue, September 2008. *Journal of Moral Education, 38*(3), 283-298.

Kitcher, P. (2007). Biology and ethics. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory. Oxford*: Oxford University Press.

Kitcher, P. (2011). *The Ethical Project*. Cambridge, MA: Harvard University Press.

Klincewicz, M. (2016). Artificial intelligence as a means to moral enhancement *Studies in Logic, Grammar and Rhetoric* (Vol. 48, pp. 171).

Koelewijn, J. (2017, July 3). Het antisociale brein bestaat niet [The antisocial brain does not exist]. *NRC Handelsblad*. Retrieved July 7, 2017, from https://www.nrc.nl/nieuws/2017/07/03/het-antisociale-brein-bestaat-niet-11520765-a1565320.

Krznaric, R. (2015, June 29). Can you teach people to have empathy? *BBC News Magazine*. Retrieved August 12, 2015, from http://www.bbc.com/news/magazine-33287727#.

L

Långström, N., Enebrink, P., Laurén, E.-M., Lindblom, J., Werkö, S., & Hanson, R. K. (2013). Preventing sexual abusers of children from reoffending: systematic review of medical and psychological interventions. *British Medical Journal, 347*.

Lapsley, D., & Carlo, G. (2014). Moral development at the crossroads: New trends and possible futures. *Developmental Psychology, 50*(1), 1-7.

Lechner, S. (2014). Why moral bioenhancement is a bad idea and why egalitarianism would make it worse. *American Journal of Bioethics, 14*(4), 31-32.

Lee, B. X. (2015). Causes and cures II: The biology of violence. *Aggression and Violent Behavior, 25, Part B*, 204-209.

Lem, S. (2014 [1965]). Altruizine, or A True Account of How Bonhomius the Hermetic Hermit Tried to Bring About Universal Happiness, and What Came of It (M. Kandel, Trans.) *The Cyberiad. Fables for the Cybernetic Age*. London, UK: Penguin.

Lev, O. (2012). Enhancing the capacity for moral agency. *AJOB Neuroscience, 3*(4), 20-22.

Levitt, M., & Manson, N. (2007). My genes made me do it? The implications of behavioural genetics for responsibility and blame. *Health Care Analysis, 15*(1), 33-40.

Levitt, M., & Pieri, E. (2009). "It could just be an additional test couldn't it?" Genetic testing for susceptibility to aggression and violence. *New Genetics and Society, 28*(2), 189-200.

Levy, N. (2006). Cognitive scientific challenges to morality. *Philosophical Psychology, 19*(5), 567-587.

Levy, N. (2007). Rethinking Neuroethics in the Light of the Extended Mind Thesis. *The American Journal of Bioethics, 7*(9), 3-11.

Levy, N. (2015a, April 10). Can our DNA turn us into criminals? *The Telegraph*. Retrieved April 13, 2015, from http://www.telegraph.co.uk/news/science/11526272/Can-our-DNA-turn-us-into-criminals.html.

Levy, N. (2015b, July 16). Common drugs can affect our minds and morals - but should we be worried about it? *The Conversation*. Retrieved April 21, 2017, from http://theconversation.com/common-drugs-can-affect-our-minds-and-morals-but-should-we-be-worried-about-it-44660.

Levy, N., Douglas, T., Kahane, G., Terbeck, S., Cowen, P. J., Hewstone, M., & Savulecu, J. (2014a). Are you morally modified?: The moral effects of widely used pharmaceuticals. *Philosophy, Psychiatry, & Psychology, 21*(2), 111-125.

Levy, N., Douglas, T., Kahane, G., Terbeck, S., Cowen, P. J., Hewstone, M., & Savulecu, J. (2014b). Disease, normality, and current pharmacological moral modification. *Philosophy, Psychiatry, & Psychology, 21*(2), 135-137.

Liégeois, A., & Eneman, M. (2008). Ethics of deliberation, consent and coercion in psychiatry. *Journal of Medical Ethics, 34*(2), 73-76.

Liu, J. (2011). Early health risk factors for violence: Conceptualization, evidence, and implications. *Aggression and Violent Behavior, 16*(1), 63-73.

Lotz, M. (2006). Feinberg, Mills, and the child's right to an open future. *Journal of Social Philosophy, 37*(4), 537-551.

M

Macer, D. (2014). A public ethos of enhancement across Asia. *American Journal of Bioethics, 14*(4), 45-47.

MacFarquhar, L. (2015). *Strangers Drowning: Grappling with Impossible Idealism, Drastic Choices, and the Urge to Help*. Penguin Press.

Malle, B. F. (2016). Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology, 18*(4), 243-256.

Mandelbaum, R. F. (2017, May 19). Can a pill make you more moral? *Gizmodo*. Retrieved May 19, 2017, from https://www.gizmodo.com.au/2017/05/can-a-pill-make-you-more-moral/.

Marsh, A. A. (2018). The neuroscience of empathy. *Current Opinion in Behavioral Sciences, 19*, 110-115.

Marshall, F. (2014). Would moral bioenhancement lead to an inegalitarian society? *American Journal of Bioethics, 14*(4), 29-30.

Maslen, H., Earp, B. D., Cohen Kadosh, R., & Savulescu, J. (2014). Brain stimulation for 'enhancement'in children: An ethical analysis. *Frontiers in Human Neuroscience, 8*, 953.

Mcleod, J. D., Pescosolido, B. A., Takeuchi, D. T., & White, T. F. (2004). Public attitudes toward the use of psychiatric medications for children. *Journal of Health and Social Behavior, 45*(1), 53-67.

McMillan, J. (2013). The kindest cut? Surgical castration, sex offenders and coercive offers. *Journal of Medical Ethics, 40*(9), 583-590.

McMillan, J. (2014). Surgical castration, coercive offers and coercive effects: it is still not about consent. *Journal of Medical Ethics, 40*(9), 596-596.

McSwiggan, S., Elger, B., & Appelbaum, P. S. (2017). The forensic use of behavioral genetics in criminal proceedings: Case of the MAOA-L genotype. *International Journal of Law and Psychiatry, 50*, 17-23.

Mendez, M. F. (2009). The neurobiology of moral behavior: review and neuropsychiatric implications. *CNS spectrums, 14*(11), 608.

Meynen, G. (2014). Neurolaw: Neuroscience, ethics, and law. Review essay. *Ethical Theory and Moral Practice, 17*(4), 819-829.

Miczek, K. A., de Almeida, R. M. M., Kravitz, E. A., Rissman, E. F., de Boer, S. F., & Raine, A. (2007). Neurobiology of escalated aggression and violence. *Journal of Neuroscience, 27*(44), 11803-11806.

Mills, C. (2003). The child's right to an open future? *Journal of Social Philosophy, 34*(4), 499-509.

Millum, J. (2014). The foundation of the child's right to an open future. *Journal of Social Philosophy, 45*(4), 522-538.

Molewijk, B., Stiggelbout, A. M., Otten, W., Dupuis, H. M., & Kievit, J. (2004). Empirical data and moral theory. A plea for integrated empirical ethics. *Medicine, Health Care and Philosophy, 7*(1), 55-69.

Monks, C. P., Smith, P. K., Naylor, P., Barter, C., Ireland, J. L., & Coyne, I. (2009). Bullying in different contexts: Commonalities, differences and the role of theory. *Aggression and Violent Behavior, 14*(2), 146-156.

Morioka, M. (2014). Some remarks on moral bioenhancement. In A. Akabayashi (Ed.), *The Future of Bioethics: International Dialogues* (pp. 120-125). Oxford: Oxford University Press.

Morse, S. J. (2016). Actions speak louder than images: the use of neuroscientific evidence in criminal cases. *Journal of Law and the Biosciences, 3*(2), 336-342.

Morse, S. J. (2017). Neuroethics: Neurolaw *Oxford Handbooks Online* (February 2017 ed., Vol. 17, pp. Available at SSRN: https://ssrn.com/abstract=2919011): Oxford University Press.

Munthe, C., & Radovic, S. (2015). The return of Lombroso? Ethical aspects of (visions of) preventive forensic screening. *Public Health Ethics, 8*(3), 270-283.

N

Narvaez, D. (2006). Integrative ethical education. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 703-733). Mahwah, NJ: Erlbaum.

Narvaez, D. (2010). Moral complexity the fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science, 5*(2), 163-181.

Nedopil, N. (2016). Special considerations in forensic psychiatry. In B. Völlm & N. Nedopil (Eds.), *The Use of Coercive Measures in Forensic Psychiatric Care: Legal, Ethical and Practical Challenges* (pp. 135-149). Cham: Springer International Publishing.

Nichols, S. (2005). Innateness and moral psychology. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind: Structure and Contents* (Vol. 1, pp. 353-370). Oxford: Oxford University Press.

Nieuwsuur. (2018, April 15). Big Brother 2.0: In China bepaalt je 'sociale score' je leven. [Big Brother 2.0: In China, your 'social score' determines your life]. Retrieved April 15, 2018, from https://nos.nl/nieuwsuur/artikel/2227440-big-brother-2-0-in-china-bepaalt-je-sociale-score-je-leven.html.

Nordmann, A. (2007). If and then: A critique of speculative nanoethics. *Nano Ethics, 1*(1), 31-46.

Nuffield Council on Bioethics. (2013). Novel neurotechnologies: Intervening in the brain. Retrieved June 6, 2014, from https://nuffieldbioethics.org/wp-content/uploads/2013/06/Novel_neurotechnologies_report_PDF_web_0.pdf.

O

Ogloff, J. R. P., Roesch, R., & Eaves, D. (2000). International perspective on forensic mental health systems. *International Journal of Law and Psychiatry, 23*(5–6), 429-431.

Oosterhuis, H. (2014). Treatment as punishment: Forensic psychiatry in The Netherlands (1870–2005). *International Journal of Law and Psychiatry, 37*(1), 37-49.

P

Pacholczyk, A. (2011). Moral enhancement: What is it and do we want it? *Law, Innovation and Technology, 3*(2), 251-277.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*(5), 411-419.

Parens, E. (2013). On good and bad forms of medicalization. *Bioethics, 27*(1), 28-35.

Parens, E. (Ed.) (1998). *Enhancing human traits: Ethical and social implications*. Washington, D.C.: Georgetown University Press.

Partridge, B., Lucke, J., & Hall, W. (2014). "If you're healthy you don't need drugs": Public attitudes towards "brain doping" in the classroom and "legalised doping" in sport. *Performance Enhancement & Health, 3*(1), 20-25.

Pearce, S., & Pickard, H. (2009). The moral content of psychiatric treatment. *British Journal of Psychiatry, 195*(4), 281-282.

Pellissier, H. (2012). IEET Readers Endorse Wide Availability of Morality Pills. Retrieved April 12, 2018, from https://ieet.org/index.php/IEET/more/poll20120323.

Persson, I., & Savulescu, J. (2008). The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity. *Journal of Applied Philosophy, 25*(3), 162-177.

Persson, I., & Savulescu, J. (2010). Moral transhumanism. *Journal of Medicine and Philosophy, 35*(6), 656-669.

Persson, I., & Savulescu, J. (2011). The turn for ultimate harm: A reply to Fenton. *Journal of Medical Ethics, 37*(7), 441-444.

Persson, I., & Savulescu, J. (2012). *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press.

Persson, I., & Savulescu, J. (2013). Getting moral enhancement right: The desirability of moral bioenhancement. *Bioethics, 27*(3), 124-131.

Persson, I., & Savulescu, J. (2014). Against fetishism about egalitarianism and in defense of cautious moral bioenhancement. *American Journal of Bioethics, 14*(4), 39-42.

Persson, I., & Savulescu, J. (2014a). Reply to commentators on Unfit for the Future. *Journal of Medical Ethics*, 10.1136/medethics-2013-101796.

Persson, I., & Savulescu, J. (2014b). Should moral bioenhancement be compulsory? Reply to Vojin Rakić. *Journal of Medical Ethics, 40*(4), 251-252.

Persson, I., & Savulescu, J. (2015a). The art of misunderstanding moral bioenhancement. *Cambridge Quarterly of Healthcare Ethics, 24*(1), 48-57.

Persson, I., & Savulescu, J. (2015b). Summary of Unfit for the Future. *Journal of Medical Ethics, 41*(4), 338-339.

Persson, I., & Savulescu, J. (2016). Enharrisment: A reply to John Harris about moral enhancement. *Neuroethics, 9*(3), 275-277.

Persson, I., & Savulescu, J. (2017). The duty to be morally enhanced. *Topoi,* 10.1007/s11245-017-9475-7.

Petersen, T. S., & Kragh, K. (2017). Should violent offenders be forced to undergo neurotechnological treatment? A critical discussion of the 'freedom of thought' objection. *Journal of Medical Ethics, 43*(1), 30-34.

Petrila, J., & de Ruiter, C. (2011). The competing faces of mental health law: Recovery and access versus the expanding use of preventive confinement. *Amsterdam Law Forum, 3*(1), 59-67.

Pickard, H. (2011). What is personality disorder? *Philosophy, Psychiatry, & Psychology, 18*(3), 181-184.

Pieri, E., & Levitt, M. (2008). Risky individuals and the politics of genetic research into aggressiveness and violence. *Bioethics, 22*(9), 509-518.

Polak, N. (2015, March 16). Gaat Virtual Reality ons betere mensen maken? [Will Virtual Reality make us better people?]. *De Correspondent*.

Poldrack, R. A. (2017, 12-07-2017). The risks of reading the brain. *Nature, 541,* 156.

Poldrack, R. A., Monahan, J., Imrey, P. B., Reyna, V., Raichle, M. E., Faigman, D., & Buckholtz, J. W. (2018). Predicting Violent Behavior: What Can Neuroscience Add? *Trends in Cognitive Sciences, 22*(2), 111-123.

Pugh, J. (2017). Moral bio-enhancement, freedom, value and the Parity Principle. *Topoi,* 10.1007/s11245-017-9482-8.

Pustilnik, A. C. (2009). Violence on the brain: a critique of neuroscience in criminal law. *Wake Forest Law Review, 44*(1), 183-238.

R

Raine, A. (2013). *The Anatomy of Violence: The biological Roots of Crime*. New York: Random House LLC.

Rakić, V. (2012). From cognitive to moral enhancement: A possible reconciliation of religious outlooks and the biotechnological creation of a better human. *Journal for the Study of Religions and Ideologies, 11*(31), 113-128.

Rakić, V. (2014a). Voluntary moral bioenhancement is a solution to Sparrow's concerns. *American Journal of Bioethics, 14*(4), 37-38.

Rakić, V. (2014b). Voluntary moral enhancement and the survival-at-any-cost bias. *Journal of Medical Ethics, 40*(4), 246-250.

Rakić, V. (2017). Compulsory administration of oxytocin does not result in genuine moral enhancement. *Medicine, Health Care and Philosophy, 20*(3), 291-297.

Rakić, V., & Hughes, J. (2015). Reflections on moral enhancement: Can we? Should we? *Cambridge Quarterly of Healthcare Ethics, 24*(01), 3-6.

Ram-Tiktin, E. (2014). The possible effects of moral bioenhancement on political privileges and fair equality of opportunity. *American Journal of Bioethics, 14*(4), 43-44.

Rampton, V. (2017, March 20). Are you creeped out by the idea of a "Moral Enhancement" pill? *Slate*. Retrieved March 21, 2017, from http://www.slate.com/articles/technology/future_tense/2017/03/why_we_are_so_alarmed_by_the_idea_of_a_moral_enhancement_pill.html.

Raus, K., Focquaert, F., Schermer, M. H. N., Specker, J., & Sterckx, S. (2014). On defining moral enhancement: A clarificatory taxonomy. *Neuroethics, 7*(3), 263-273.

Reed, D. C., & Stoermer, R. M. (2008). Towards an integrated model of moral functioning: an overview of the Special Issue. *Journal of Moral Education, 37*(3), 417-428.

Reichlin, M. (2017). The moral agency argument against moral bioenhancement. *Topoi,* 10.1007/s11245-017-9471-y.

Reimer, M. (2010). Moral aspects of psychiatric diagnosis: the Cluster B Personality Disorders. *Neuroethics, 3*(2), 173-184.

Reiner, P. B. (2013, November 5). Experimental neuroethics. Retrieved November 26, 2015, from http://www.theneuroethicsblog.com/2013/11/experimental-neuroethics_5.html.

Reiner, P. B. (2017, August 30). What can neuroethicists learn from public attitudes about moral bioenhancement? Retrieved October 28, 2017, from http://www.theneuroethicsblog.com/2017/08/what-can-neuroethicists-learn-from.html.

Riis, J., Simmons, J. P., & Goodwin, G. P. (2008). Preferences for enhancement pharmaceuticals: The reluctance to enhance fundamental traits. *Journal of Consumer Research, 35*(3), 495-508.

Rini, R. A. (2016). Why moral psychology is disturbing. *Philosophical Studies, 174*(6), 1439-1458.

Ripley, A. (2016, September 12). Can teenage defiance be manipulated for good? *The New York Times*. Retrieved December 12, 2017, from https://www.nytimes.com/2016/09/13/upshot/can-teenage-defiance-be-manipulated-for-good.html?_r=1.

Robertson, M. D., & Walter, G. (2008). Many faces of the dual-role dilemma in psychiatric ethics. *Australian and New Zealand Journal of Psychiatry, 42*(3), 228-235.

Robichaud, P. (2014). Moral capacity enhancement does not entail moral worth enhancement. *American Journal of Bioethics, 14*(4), 33-34.

Rocque, M., Welsh, B. C., & Raine, A. (2012). Biosocial criminology and modern crime prevention. *Journal of Criminal Justice, 40*(4), 306-312.

Rose, N. S. (2000). The biology of culpability: Pathological identity and crime control in a biological culture. *Theoretical criminology, 4*(1), 5-34.

Rose, N. S. (2010). 'Screen and intervene': Governing risky brains. *History of the Human Sciences, 23*(1), 79-105.

Rose, N. S. (2013). The human sciences in a biological age. *Theory, Culture & Society, 30*(1), 3-34.

Rose, N. S., & Abi-Rached, J. M. (2013). *Neuro: The New Brain Sciences and the Management of the Mind*. Princeton, NJ: Princeton University Press.

Roskies, A. L., Schweitzer, N. J., & Saks, M. J. (2013). Neuroimages in court: less biasing than feared. *Trends in Cognitive Sciences, 17*(3), 99-101.

Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior, 43*, 304-307.

Roychowdhury, A., & Adshead, G. (2014). Violence risk assessment as a medical intervention: Ethical tensions. *The Psychiatrist, 38*(2), 75-82.

Ryberg, J. (2015). Is coercive treatment of offenders morally acceptable? On the deficiency of the debate. *Criminal Law and Philosophy, 9*(4), 619-631.

S

Sadler, J. Z. (2013). Vice and mental disorders. In K. W. M. Fulford, M. Davies, R. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. I. M. Thornton (Eds.), *The Oxford Handbook of Philosophy and Psychiatry* (pp. 451-479). Oxford: Oxford University Press.

Sadler, J. Z. (2014). Conduct disorder as a vice-laden diagnostic concept. In C. D. Perring & L. A. Wells (Eds.), *Diagnostic dilemmas in child and adolescent psychiatry: Philosophical perspectives* (pp. 166-181). Oxford: Oxford University Press.

Salize, H. J., Dressing, H., & Kief, C. (2007). *Mentally disordered persons in European prison systems - Needs, programmes and outcome (EUPRIS). Final report*. Germany: Central Institute of Mental Health.

Sampson, S., Edworthy, R., Völlm, B., & Bulten, E. (2016). Long-term forensic mental health services: An exploratory comparison of 18 European countries. *International Journal of Forensic Mental Health, 15*(4), 333-351.

Savulescu, J., Douglas, T., & Persson, I. (2014). Autonomy and the ethics of biological behaviour modification. In A. Akabayashi (Ed.), *The Future of Bioethics: International Dialogues* (pp. 91-112). Oxford: Oxford University Press.

Savulescu, J., & Maslen, H. (2015). Moral enhancement and artificial intelligence: Moral AI? In J. Romportl, E. Zackova, & J. Kelemen (Eds.), *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide* (pp. 79-95). Cham: Springer International Publishing.

Savulescu, J., & Persson, I. (2012). Moral enhancement, freedom, and the God Machine. *Monist, 95*(3), 399-421.

Schaefer, G. O. (2011). What Is the goal of moral engineering? *AJOB Neuroscience, 2*(4), 10-11.

Schaefer, G. O. (2015). Direct vs. indirect moral enhancement. *Kennedy Institute of Ethics Journal, 25*(3), 261-289.

Schechtman, M. (2009). Philosophical reflections on narrative and deep brain stimulation. *Journal of Clinical Ethics, 21*(2), 133-139.

Schermer, M. H. N. (2006). Voorspellende en preventieve criminologie. Parallellen met de geneeskunde. [Predictive and preventive criminology. Parallels with medicine]. *Justitiele Verkenningen, 32*(8), 103-116.

Schermer, M. H. N. (2014). Ethics of pharmacological mood enhancement. In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics* (pp. 1177-1190): Springer Netherlands.

Schermer, M. H. N. (2015). Reducing, restoring, or enhancing autonomy with neuromodulation techniques. In W. Glannon (Ed.), *Free Will and the Brain. Neuroscientific, Philosophical, and Legal Perspectives* (pp. 205-227). Cambridge: Cambridge University Press.

Schermer, M. H. N. (2016). Cognitive enhancement in The Netherlands. Practices, public opinion, and ethics. In F. Jotterand & V. Dubljević (Eds.), *Cognitive Enhancement: Ethical and Policy Implications in International Perspectives* (pp. 181-195). Oxford: Oxford University Press.

Schirmann, F. (2013a). Badness, madness and the brain–the late 19th-century controversy on immoral persons and their malfunctioning brains. *History of the Human Sciences, 26*(2), 33-50.

Schirmann, F. (2013b). Invoking the brain in studying morality: A theoretical and historical perspective on the neuroscience of morality. *Theory & Psychology, 23*(2), 289-304.

Schlag, S. (2016). The tragedy of biomedical moral enhancement. *Neuroethics,* 10.1007/s12152-016-9284-5.

Schleim, S. (2014). Critical neuroscience—or critical science? A perspective on the perceived normative significance of neuroscience. *Frontiers in Human Neuroscience, 8*, 336.

Schuijff, M., & Munnichs, G. (2012). *Goed, beter, betwist: publieksonderzoek naar mensverbetering. [Good, better, disputed: public attitudes towards human enhancement]*. Den Haag: Rathenau Instituut.

Schuitema, J., Ten Dam, G., & Veugelers, W. (2008). Teaching strategies for moral education: A review. *Journal of Curriculum Studies, 40*(1), 69-89.

Sedgwick, O., Young, S., Das, M., & Kumari, V. (2016). Objective predictors of outcome in forensic mental health services-a systematic review. *CNS spectrums, 21*(6), 430-444.

Selgelid, M. J. (2014). Freedom and moral enhancement. *Journal of Medical Ethics, 40*(4), 215-216.

Sharma, S., & Sharma, G. (2006). Exploring evolving concepts and challenges in forensic psychiatry. *World Psychiatry, 5*(2), 97-98.

Shaw, D. (2014). Neuroenhancing public health. *Journal of Medical Ethics, 40*(6), 389-391.

Shaw, E. (2014). The use of brain interventions in offender rehabilitation programs: Should it be mandatory, voluntary, or prohibited? In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics* (pp. 1381-1398). Dordrecht: Springer Netherlands.

Shaw, E. (2016). Psychopathy, moral understanding and criminal responsibility. *European Journal of Current Legal Issues, 22*(2).

Shook, J. R. (2012). Neuroethics and the possible types of moral enhancement. *AJOB Neuroscience, 3*(4), 3-14.

Shook, J. R., & Giordano, J. J. (2016a). Moral enhancement? Acknowledging limitations of neurotechnology and morality. *AJOB Neuroscience, 7*(2), 118-120.

Shook, J. R., & Giordano, J. J. (2016b). Neuroethics beyond normal. *Cambridge Quarterly of Healthcare Ethics, 25*(1), 121-140.

Shook, J. R., & Giordano, J. J. (2017). Moral bioenhancement for social welfare: Are civic institutions ready? *Frontiers in Sociology, 2*(21).

Sie, M. (2014). Self-knowledge and the minimal conditions of responsibility: A traffic-participation view on human (moral) agency. *The Journal of Value Inquiry, 48*(2), 271-291.

Simkulet, W. (2012). On moral enhancement. *AJOB Neuroscience, 3*(4), 17-18.

Singer, P. (2015). *The Most Good You Can Do: How Effective Altruism Is Changing Ideas About Living Ethically*. New Haven and London: Yale University Press.

Singer, P., & Sagan, A. (2012, January 28). Are we ready for a "Morality Pill"? *The New York Times*. Retrieved April 12, 2018, from https://opinionator.blogs.nytimes.com/2012/01/28/are-we-ready-for-a-morality-pill/.

Singh, I., & Rose, N. S. (2009). Biomarkers in psychiatry. *Nature, 460*(7252), 202-207.

Singh, I., & Sinnott-Armstrong, W. P. (2014). Introduction: Deviance, classification, and bioprediction. In I. Singh, W. P. Sinnott-Armstrong, & J. Savulescu (Eds.), *Bioprediction, Biomarkers, and Bad Behavior: Scientific, Legal, and Ethical Challenges* (pp. 1-11). Oxford Scholarship Online.

Singh, I., Sinnott-Armstrong, W. P., & Savulecu, J. (2013). *Bioprediction, Biomarkers, and Bad Behavior: Scientific, Legal, and Ethical Challenges*. Oxford Scholarship Online.

Sio, F. S. d., Maslen, H., & Faulmüller, N. (2012). The necessity of objective standards for moral enhancement. *AJOB Neuroscience, 3*(4), 15-16.

Slaby, J. (2013). The new science of morality: A bibliographic review. *The Hedgehog Review, 15*(1), 46-54.

Smith, P. K., Cowie, H., Olafsson, R. F., & Liefooghe, A. P. D. (2002). Definitions of bullying: A comparison of terms used, and age and gender differences, in a Fourteen–Country international comparison. *Child Development, 73*(4), 1119-1133.

Sparrow, R. (2014a). Better living through chemistry? A reply to Savulescu and Persson on 'moral enhancement'. *Journal of Applied Philosophy, 31*(1), 23-32.

Sparrow, R. (2014b). Egalitarianism and moral bioenhancement. *American Journal of Bioethics, 14*(4), 20-28.

Specker, J., Focquaert, F., Raus, K., Sterckx, S., & Schermer, M. H. N. (2014). The ethical desirability of moral bioenhancement: A review of reasons. *BMC medical ethics, 15*(1), 67.

Specker, J., Focquaert, F., Sterckx, S., & Schermer, M. H. N. (2018). Forensic practitioners' expectations and moral views regarding neurobiological interventions in offenders with mental disorders. *BioSocieties, 13*(1), 304-321.

Specker, J., & Schermer, M. H. N. (2017). Imagining moral bioenhancement practices. Drawing inspiration from moral education, public health ethics, and forensic psychiatry. *Cambridge Quarterly of Healthcare Ethics, 26*(3), 415-426.

Specker, J., Schermer, M. H. N., & Reiner, P. B. (2017). Public attitudes towards moral enhancement: Evidence that means matter morally. *Neuroethics, 10*(3), 405-417.

Spence, S. A. (2008). Can pharmacology help enhance human morality? *The Britisch Journal of Psychiatry, 193*(3), 179-180.

Sprinkle, R. H. (2010). Moral suasion, installed. *Politics and the Life Sciences, 29*(1), 88-89.

Starr, S. B. (2014). Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Review, 66*(4), 803-872.

Stol, Y. H., Schermer, M. H. N., & Asscher, E. C. A. (2017). Omnipresent health checks may result in over-responsibilization. *Public Health Ethics, 10*(1), 35-48.

Stone, A. A. (1984). The ethical boundaries of forensic psychiatry: A view from the ivory tower. *Journal of the American Academy of Psychiatry and the Law, 12*(3), 209-219.

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*(1), 159-171.

Strong, K. A., Lipworth, W., & Kerridge, I. (2010). The strengths and limitations of empirical bioethics. *Journal of Law and Medicine, 18*(2), 316-319.

Sugarman, J. (2004). The future of empirical research in bioethics. *The Journal of Law, Medicine & Ethics, 32*(2), 226-231.

Summers, J. A., & Caplan, P. J. (1987). Laypeople's attitudes toward drug treatment for behavioral control depend on which disorder and which drug. *Clinical Pediatrics, 26*(5), 258-262.

Sunstein, C. R. (2017). Nudges that fail. *Behavioural Public Policy, 1*(1), 4-25.

T

Tanner, C., & Christen, M. (2014). Moral intelligence–A framework for understanding moral competences *Empirically Informed Ethics: Morality between facts and norms* (pp. 119-136). Springer.

Tappin, B. M., & McKay, R. T. (2016). The illusion of moral superiority. *Social Psychological and Personality Science, 8*(6), 623-631.

Temel, Y., Hescham, S., Melse, M., & Visser-Vandewalle, V. (2016). Deep brain stimulation: Emerging indications. In C. Hamani, P. Holtzheimer, A. M. Lozano, & H. Mayberg (Eds.), *Neuromodulation in Psychiatry* (pp. 309-324). John Wiley & Sons.

Tennison, M. N. (2012). Moral transhumanism: The next step. *Journal of Medicine and Philosophy, 37*(4), 405-416.

Thibaut, F., Barra, F. D. L., Gordon, H., Cosyns, P., & Bradford, J. M. W. (2010). The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the biological treatment of paraphilias. *The World Journal of Biological Psychiatry, 11*(4), 604-655.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349-357.

Tonkens, R. (2013). Feeling good about the end: Adderall and moral enhancement. *AJOB Neuroscience, 4*(1), 15-16.

Tracy, D. G., Michael, G. V., & Robert, A. P. (2010). Behavioral genetics in antisocial spectrum disorders and psychopathy: A review of the recent literature. *Behavioral Sciences and the Law, 28*(2), 148-173.

Tremonti, A. M. (2017, March 28). Is it ethical to swallow a morality pill? Retrieved March 30, 2017, from http://www.cbc.ca/radio/thecurrent/the-current-for-march-28-2017-1.4042750/is-it-ethical-to-swallow-a-morality-pill-1.4042754.

Triviño, J. L. P. (2013). On the need of moral enhancement. A critical comment of "Unfit for the future" of I. Persson and J. Savulescu. *Dilemata*(12), 261-269.

Ttofi, M. M., Farrington, D. P., & Lösel, F. (2012). School bullying as a predictor of violence later in life: A systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior, 17*(5), 405-418.

Turiel, E. (2008). Foreword. *Journal of Moral Education, 37*(3), 279-288.

Turner, D., Petermann, J., Harrison, K., Krueger, R., & Briken, P. (2017). Pharmacological treatment of patients with paraphilic disorders and risk of sexual offending: An international perspective. *The World Journal of Biological Psychiatry,* 10.1080/15622975.2017.1395069.

V

van de Werff, T., Slatman, J., & Swierstra, T. (2016). Can we "remedy" neurohype, and should we? Using neurohype for ethical deliberation. *AJOB Neuroscience, 7*(2), 97-99.

van der Gronde, T., Kempes, M., van El, C., Rinne, T., & Pieters, T. (2014). Neurobiological Correlates in Forensic Assessment: A Systematic Review. *PLoS ONE, 9*(10), e110672.

van Goozen, S. H. M., & Fairchild, G. (2008). How can the study of biological processes help design new interventions for children with severe antisocial behavior? *Development and Psychopathology, 20*(3), 941-973.

van Hintum, M. (2018, March 4). Op afstand de bloeddruk van een TBS'er volgen; is dat een goed idee? [Monitoring the blood pressure of an offender from a distance; is that a good idea?]. *Trouw*. Retrieved March 8, 2018, from https://www.trouw.nl/home/op-afstand-de-bloeddruk-van-een-tbs-er-volgen-is-dat-een-goed-idee-~a158cb73/.

van Marle, H. J. (2002). The Dutch Entrustment Act (TBS): Its principles and innovations. *International Journal of Forensic Mental Health, 1*(1), 83-92.

van Thiel, G. J. M. W., & van Delden, J. J. M. (2010). Reflective equilibrium as a normative empirical model. *Ethical Perspectives-Katholieke Universiteit, 17*(2), 183-202.

Velinov, V. T., & Marinov, P. M. (2006). Forensic psychiatric practice: Worldwide similarities and differences. *World Psychiatry, 5*(2), 98-99.

Viding, E., Blair, R. J. R., Moffitt, T. E., & Plomin, R. (2005). Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry, 46*(6), 592-597.

W

Walker, M. (2009). Enhancing genetic virtue: A project for twenty-first century humanity? *Politics and the Life Sciences, 28*(2), 27-47.

Walker, M. (2010). In defense of the Genetic Virtue Program. *Politics and the Life Sciences, 29*(1), 90-96.

Ward, T. (2013). Addressing the dual relationship problem in forensic and correctional practice. *Aggression and Violent Behavior, 18*(1), 92-100.

Ward, T., Gannon, T. A., & Fortune, C.-A. (2015). Restorative justice–informed moral acquaintance: Resolving the dual role problem in correctional and forensic practice. *Criminal Justice and Behavior, 42*(1), 45-57.

Wasserman, D. (2014a). Ethical and policy issues in genetic prediction of violence: Implications for clinicians. *Current Genetic Medicine Reports, 2*(4), 216-222.

Wasserman, D. (2014b). When bad people do good things: Will moral enhancement make the world a better place? *Journal of Medical Ethics, 40*(6), 374-375.

Wexler, A. (2015). The practices of do-it-yourself brain stimulation: implications for ethical considerations and regulatory proposals. *Journal of Medical Ethics, 42*(4), 211-215.

Wexler, A. (2017). The social context of "do-it-yourself" brain stimulation: Neurohackers, biohackers and lifehackers. *Frontiers in Human Neuroscience, 11*, 224.

White, K. L. A., Jordens, C. F. C., & Kerridge, I. (2014). Contextualising professional ethics: The impact of the prison context on the practices and norms of health care practitioners. *Journal of Bioethical Inquiry, 11*(3), 333-345.

Wikström, P.-O. H. (2010). Explaining crime as moral actions. In S. Hitlin & S. Vaisey (Eds.), *Handbook of the Sociology of Morality* (pp. 211-239). Springer.

Williams, D. J., & Donnelly, P. D. (2014). Is violence a disease? Situating violence prevention in public health policy and practice. *Public Health, 128*(11), 960-967.

Williams, E. F., & Steffel, M. (2014). Double standards in the use of enhancing products by self and others. *Journal of Consumer Research, 41*(2), 506-525.

Wilson, A. T. (2014). Egalitarianism and successful moral bioenhancement. *American Journal of Bioethics, 14*(4), 35-36.

Wiseman, H. (2014a). Moral enhancement-"hard" and "soft" forms. *American Journal of Bioethics, 14*(4), 48-49.

Wiseman, H. (2014b). SSRIs as moral enhancement interventions: A practical dead end. *AJOB Neuroscience, 5*(3), 21-30.

Wiseman, H. (2016). *The Myth of the Moral Brain: The Limits of Moral Enhancement*. Cambridge, MA: MIT Press.

Wiseman, H. (2017). Would we even know moral bioenhancement if we saw it? *Cambridge Quarterly of Healthcare Ethics, 26*(3), 398-410.

Witzel, J., Walter, M., Bogerts, B., & Northoff, G. (2008). Neurophilosophical perspectives of neuroimaging in forensic psychiatry—giving way to a paradigm shift? *Behavioral Sciences and the Law, 26*(1), 113-130.

Wolke, D., & Lereya, S. T. (2015). Long-term effects of bullying. *Archives of Disease in Childhood, 100*(9), 879-885.

Wolpe, P. R. (2013). Rethinking the implications of discovering biomarkers for biologically based criminality. In I. Singh, W. P. Sinnott-Armstrong, & J. Savulescu (Eds.), *Bioprediction, biomarkers, and bad behavior: Scientific, legal, and ethical challenges* (pp. 118-130). Oxford: Oxford University Press.

Wren, T. (2014). Philosophical moorings. In L. Nucci, D. Narvaez, & T. Krettenauer (Eds.), *Handbook of Moral and Character Education* (2 ed., pp. 11-29). New York: Taylor & Francis.

Y

Yamasue, H., Yee, J. R., Hurlemann, R., Rilling, J. K., Chen, F. S., Meyer-Lindenberg, A., & Tost, H. (2012). Integrative approaches utilizing oxytocin to enhance prosocial behavior: from animal and human social behavior to autistic social dysfunction. *The Journal of Neuroscience, 32*(41), 14109-14117a.

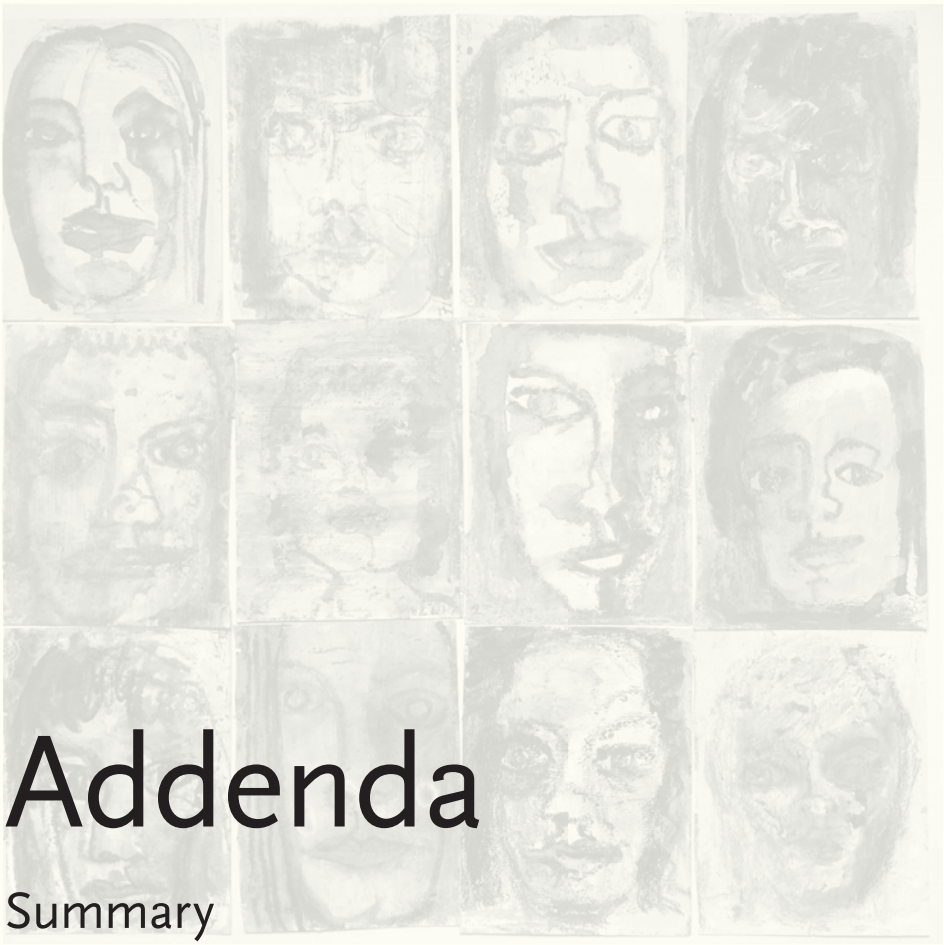
Yoder, K. J., & Decety, J. (2017). The neuroscience of morality and social decision-making. *Psychology, Crime & Law, 24*(3), 179-295.

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience, 7*(1), 1-10.

Z

Zarpentine, C. (2013). 'The thorny and arduous path of moral progress': Moral psychology and moral enhancement. *Neuroethics, 6*(1), 141-153.

Zik, J. B., & Roberts, D. L. (2015). The many faces of oxytocin: Implications for psychiatry. *Psychiatry Research, 226*(1), 31-37.



Addenda

Summary

Since 2008, the so-called 'moral enhancement debate' asks whether we should actively pursue the development of moral enhancement technologies, and whether it would be permissible – or even obligatory – to put them to use, provided that these interventions would be effective and safe?

Whereas 'traditional methods' of moral betterment (such as upbringing, socialization and education) are arguably as old as humanity itself, the debate on moral bioenhancement focuses on the desirability of methods based on novel biomedical insights and the use of biomedical methods. The debate follows a significant rise in fundamental research on the (neuro)biological and genetic underpinnings of morality. Potential interventions that are being discussed range from various types of psychopharmaceuticals, neurostimulation, and genetic selection and engineering.

Until now, the theoretical debate on moral (bio)enhancement has a strong speculative character and mostly precedes and runs ahead of realistic scientific possibilities. Moreover, to date, the debate risks lacking focus and real world impact, as different commentators fail to agree on how to understand and define moral enhancement.

This under-examination of potential *practices of moral enhancement* is significant because without specifying intended users, contexts of implementation, and the goals and objectives of developing and applying potential biomedical possibilities for optimizing morally relevant capacities, it is not clear who should be concerned about this debate. At the same time, existing or emerging practices that already contain elements of 'moral enhancement' but as yet lack ethical attention and ethical scrutiny, risk staying out of sight.

In order to address this problem, this thesis focuses on (present and emerging) moral enhancement *practices*, in order to identify ethical issues that are not necessarily part of the current debate on moral enhancement. Rather than distinguishing in a blanket fashion desirable from undesirable moral enhancement *technologies*, this thesis intends to formulate conditions and ethical requirements for ethically justifiable moral enhancement practices.

On the basis of a systematic clustering of ethical arguments that emerge in the moral enhancement debate (based on a systematic search of the literature), it is concluded, first, that there is little discussion on what distinguishes moral bioenhancement from treatment of pathological deficiencies in morality (chapter 2). Furthermore, remark-

ably little attention has been paid so far to the safety, risks and side-effects of moral enhancement, including the risk of identity changes. Finally, many authors overestimate the scientific as well as the practical feasibility of the interventions they discuss, rendering the debate too speculative.

Subsequently, a number of contexts or domains in which (future) moral bioenhancement interventions possibly or most likely will be implemented are investigated (chapter 3). By looking closely at similar or related existing practices and their relevant ethical frameworks, we identify ethical considerations that are relevant for evaluating potential moral bioenhancement interventions. Domains that are examined are, first, debates on the proper scope of moral education; second, proposals for identifying early risk factors for antisocial behaviour, and third, the difficult balancing of the best interests of the individual and third party concerns in (forensic) psychiatry.

In a series of interviews we asked forensic practitioners about the prospects, threats, and limitations of integrating neurobiological and behavioural genetic interventions in forensic psychiatric practices (chapter 4). Potential applications that were included were potential biomedical possibilities to lower aggression, the possible usage of neuroimaging in assessing legal responsibility, and the potential use of biomarkers in assessing risk for future violent and antisocial behaviour.

In addition, these interviews explored to what extent forensic practitioners consider moral development and moral growth to be a part of their current professional practices and to what extent they think that stimulating moral development is a legitimate objective in the context of forensic psychiatric treatment (chapter 5). Finally, we asked how forensic practitioners balance public safety and risk management concerns with the interests and wellbeing of the individual patient.

To gain insight into the reasons that the public may have for endorsing or eschewing pharmacological moral enhancement for themselves or for others, we used empirical tools to explore public attitudes towards these issues (chapter 6). We found that people were significantly more troubled by pharmacological as opposed to non-pharmacological moral enhancement interventions. The results indicate that members of the public for the greater part oppose pharmacological moral bioenhancement, yet are open to non-biomedical means to attain moral enhancement.

The final chapter starts from the observation that a range of present and emerging practices already contain elements of moral enhancement. As technological possibilities (and perhaps public and political willingness) to influence and alter moral

behaviour and moral capacities increase, the responsibility to make explicit this moral enhancement dimension in our current practices, and to think carefully about the associated ethical issues, increase as well.

When ethically evaluating (potential) moral enhancement practices, the distinction between moral self-enhancement and moral other-enhancement is relevant. With respect to moral self-enhancement, the decision to pursue moral enhancement and to determine the goal and appropriate means of doing so, is to a large degree up to the person herself, provided a number of safeguards are put in place. Clearly, in the case of moral enhancement of *others*, there are additional and much larger responsibilities to justify the need, objectives, and procedure and the means employed, as well as the balancing of potential benefits and harms, especially in particularly vulnerable target gourps.

Samenvatting

Sinds 2008 staat binnen het zogenaamde 'moral enhancement debat' de vraag centraal of we de ontwikkeling van technologieën ten behoeve van morele mensverbetering actief moeten nastreven en of het toegestaan of zelfs verplicht zou zijn om ze in te zetten, op voorwaarde dat deze interventies effectief en veilig zijn.

Terwijl 'traditionele methoden' van morele mensverbetering zoals opvoeding, socialisatie en onderwijs, waarschijnlijk net zo oud zijn als de mensheid zelf, richt het debat over morele mensverbetering zich op de wenselijkheid van interventies die zijn gebaseerd op nieuwe biomedische inzichten en op het gebruik van biomedische methoden. Het debat volgt op een aanzienlijke toename van fundamenteel onderzoek naar (neuro-)biologische en genetische verklaringen van moraliteit. Potentiële interventies die worden besproken variëren van verschillende soorten psychofarmaca, neurostimulatie tot genetische selectie en manipulatie.

Tot nu toe heeft het theoretische debat over morele mensverbetering een sterk speculatief karakter en loopt het ver vooruit op realistische wetenschappelijke mogelijkheden. Bovendien dreigt het debat focus en daadwerkelijke impact te ontberen, omdat verschillende commentatoren het niet eens te worden over hoe morele mensverbetering begrepen en gedefinieerd dient te worden.

Het gebrek aan onderzoek naar potentiële praktijken van morele mensverbetering is problematisch, omdat zonder de context, doelgroep en doelstellingen van het toepassen van potentiële biomedische mogelijkheden voor morele mensverbetering te expliciteren, niet duidelijk is wie zich van dit debat iets zou moeten aantrekken, terwijl bestaande of opkomende praktijken die al elementen van morele mensverbetering in zich dragen juist teveel buiten beeld dreigen te blijven.

Dit proefschrift richt zich daarom op (huidige en toekomstige) praktijken van morele mensverbetering, met het oog op het identificeren van ethische kwesties die niet noodzakelijk deel uitmaken van het huidige debat over morele mensverbetering. In plaats van onderscheid te maken tussen wenselijke en ongewenste *technologieën* voor morele mensverbetering, beoogt dit proefschrift voorwaarden te formuleren voor ethisch verantwoorde *praktijken* van morele mensverbetering.

Op basis van een systematische clustering van ethische argumenten die naar voren komen in het debat over morele mensverbetering, wordt in de eerste plaats geconcludeerd dat er weinig discussie is over het onderscheid tussen morele mensverbetering

en de behandeling van psychische stoornissen of tekorten (hoofdstuk 2). Bovendien is tot nu toe opvallend weinig aandacht besteed aan de veiligheid, risico's en bijwerkingen van morele mensverbetering, waaronder het risico op identiteitsveranderingen. Ten slotte overschatten veel auteurs zowel de wetenschappelijke als de praktische haalbaarheid van de interventies die zij bespreken, waardoor het debat te zeer een speculatief karakter heeft.

Vervolgens worden een aantal contexten of domeinen onderzocht waarin (toekomstige) morele mensverbeteringsinterventies mogelijk ingang zullen krijgen (hoofdstuk 3). Door vergelijkbare of verwante bestaande praktijken en hun relevante ethische kaders nauw te bestuderen, kunnen ethische overwegingen worden geïdentificeerd die relevant zijn voor het beoordelen van potentiële morele mensverbeteringsinterventies. Domeinen die werden onderzocht zijn, ten eerste debatten over wat geschikte doelstellingen van morele opvoeding zijn; ten tweede, voorstellen voor het identificeren van risicofactoren voor antisociaal gedrag, en ten derde, de moeilijke afweging in de (forensische) psychiatrie tussen de belangen van het individu en die van derden.

In een reeks interviews vroegen we forensische artsen naar de vooruizichten, bedreigingen en beperkingen van integratie van neurobiologische en gedragsgenetische interventies in forensisch psychiatrische praktijken (hoofdstuk 4). Toepassingen die werden besproken betroffen potentiële biomedische interventies om agressie te verlagen, het mogelijke gebruik van beeldvormend onderzoek bij het beoordelen van verantwoordelijkheid en het mogelijke gebruik van 'biomarkers' bij het bepalen van risico op toekomstig gewelddadig en antisociaal gedrag.

Daarnaast hebben we door middel van deze interviews onderzocht in hoeverre forensische artsen morele ontwikkeling en morele groei als onderdeel van hun huidige professionele praktijken beschouwen en in welke mate zij denken dat het stimuleren van morele ontwikkeling een legitiem doel is in de context van de forensisch psychiatrische behandeling (hoofdstuk 5). Daarnaast vroegen we hoe forensisch specialisten overwegingen rondom openbare veiligheid en risicobeheer afwegen tegen de belangen en het welzijn van de individuele patiënt.

Door middel van empirisch onderzoek hebben we getracht inzicht te krijgen in de redenen van het publiek om het gebruik van farmacologische middelen ten behoeve van morele mensverbetering voor zichzelf of voor anderen te steunen of af te wijzen (hoofdstuk 6). We zagen dat mensen significant meer problemen hadden met farmacologische in vergelijking met niet-farmacologische morele mensverbeteringsinterventies. De resultaten suggereren dat het publiek zich grotendeels verzet tegen

farmacologische middelen, maar open staat voor niet-biomedische middelen om morele mensverbetering te bereiken.

Het laatste hoofdstuk vertrekt vanuit de observatie dat zowel een aantal bestaande als opkomende praktijken al elementen van morele mensverbetering bevatten. Naarmate technologische mogelijkheden om moreel gedrag en morele capaciteiten te beïnvloeden en te veranderen toenemen, en tegelijk wellicht ook de publieke en politieke bereidheid om die mogelijkheden daadwerkelijk te gebruiken, neemt ook het belang van het verhelderen van deze elementen van morele mensverbetering toe, alsmede de noodzaak, om zorgvuldig na te denken over de daarmee samenhangende ethische kwesties.

Voor deze ethische analyse is het onderscheid tussen morele zelf-verbetering en morele verbetering van anderen relevant. De beslissing om morele zelf-verbetering na te streven is primair aan de persoon zelf, mits aan een aantal voorwaarden is voldaan. In het geval van morele verbetering van anderen bestaat er een veel uitgebreidere verantwoordelijkheid om de noodzaak, doelstellingen, procedure, en de gekozen interventie te rechtvaardigen, zeker in het geval van kwetsbare groepen.

Curriculum Vitae

Jona Specker is born on February 4th, 1984 in Amsterdam, the Netherlands. After living in Peru and Indonesia for a number of years, she grew up in Breda. After completing secondary school in Breda at the Stedelijk Gymnasium, she started at the newly founded University College in Maastricht. Here her studies comprised a combination of philosophy, psychology, and history. During this period, she spent a semester abroad at the University of California, Los Angeles (UCLA), studying philosophy and the history of documentary film. After obtaining a bachelor's degree (summa cum laude), Jona continued by pursuing a masters in Philosophy at the Erasmus University Rotterdam, where she wrote a thesis on evolutionary accounts of morality, supervised by Prof. Maureen Sie.

Next, she worked for three years at the Dutch Scientific Council for Government Policy (WRR) in The Hague. Here she was a member of the research team *Confidence in Citizens*, led by Prof. Pieter Winsemius, writing a report on new forms of civic engagement.

After her first daughter Line was born, Jona started a PhD project on the ethical desirability of moral enhancement at the department of Medical Ethics and Philosophy of Medicine at the Erasmus Medical Center Rotterdam under supervision of Prof. Maartje Schermer. She obtained a basic teaching qualification for higher education (BKO), and was involved in teaching medical ethics and argumentation and academic writing courses to medical students, clinical technology students, and general practitioners in training. During this period, she spent a trimester at the National Core for Neuroethics at the University of British Columbia (UBC), under supervision of Prof. Peter Reiner.

Jona lives in Katendrecht, Rotterdam, together with her husband, Theun, and two daughters, Line and Jade.

List of Publications

Raus, K., Focquaert, F., Schermer, M.H.N., Specker, J., & Sterckx, S. (2014). On defining moral enhancement: A clarificatory taxonomy. *Neuroethics*, *7*(3), 263-273.

Specker, J., Focquaert, F., Raus, K., Sterckx, S., & Schermer, M.H.N. (2014). The ethical desirability of moral bioenhancement: A review of reasons. *BMC Medical Ethics, 15*(1), 67.

Specker, J., & Schermer, M.H.N. (2017). Imagining moral bioenhancement practices. Drawing inspiration from moral education, public health ethics, and forensic psychiatry. *Cambridge Quarterly of Healthcare Ethics, 26*(3), 415-426.

Specker, J., Schermer, M.H.N., & Reiner, P.B. (2017). Public attitudes towards moral enhancement. Evidence that means matter morally. *Neuroethics, 10*(3), 405-417.

Specker, J. (2017) "Morality pills": het moral enhancement-debat belicht vanuit de praktijk. *NVBe Podium*, *24*(4), 9-12.

Specker, J., Focquaert, F., Sterckx, S., & Schermer, M.H.N. (2018). Forensic practitioners' expectations and moral views regarding neurobiological interventions in offenders with mental disorders. *BioSocieties, 13*(1), 304-321.

Specker, J., Focquaert, F., Sterckx, S., & Schermer, M.H.N. (2018). Forensic practitioners' views on stimulating moral development and moral growth in forensic psychiatric care. *Neuroethics*, 10.1007/s12152-018-9363-x.

PhD Portfolio

Name PhD student: Jona Specker	**PhD period: 2013 – 2018**
Erasmus MC Department: Medical Ethics and Philosophy	**Promotor(s): M.H.N. (Maartje) Schermer, I.D. (Inez) de Beaufort**
Research School: OZSW	**Supervisor: M.H.N. (Maartje) Schermer**

1. PhD training

	Year	Workload (ECTS)
PhD research and academic courses		
- Systematic Literature Retrieval in PubMed and other databases, Endnote, Medical Library Erasmus MC	2013	1.0
- "Integrity in research", Erasmus MC	2013	2.0
- OZSW Winter School "Ethical Theory & Applied Ethics"	2014	5.0
- OZSW Summer School "Emotions and Moral Agency"	2015	5.0
- RISBO training "Onderwijs en de kunst van het theater"	2015	1.0
- BKO: Basic Teaching Qualification for Higher Education, certificate obtained June 2016	2015 – 2016	5.0
National conferences		
- OZSW conference, Rotterdam	2013	1.0
- OZSW conference, Nijmegen. Oral presentation:*"The debate on the ethical desirability of moral bioenhancement"*	2014	2.0
- Integrity in Research congress, Erasmus MC Rotterdam. Oral presentation: *"Authorships"*	2015	2.0
- "Moral progress – concept, measurement and application" VUMC Amsterdam	2015	1.0
International conferences		
- Schuld onder de schedel, Ghent	2013	1.0
- ESPMH Congress, Ghent. Oral presentation: *"Public health approaches to violent crime"*	2015	2.0
- CBS/ Hastings Center Conference, Enhancing Understanding of Enhancement, Belgrade. Oral presentation: *"Evaluating moral enhancement"*	2015	2.0
- Justice Without Retribution III, Ghent. Oral presentation: *"Forensic Practitioners' Expectations and Moral Views Regarding Neurobiological Interventions in Offenders"*	2017	2.0

Reading groups, seminars, and workshops		
- Reading group, Department of Medical Ethics and Philosophy, Erasmus MC	2013 – 2018	3.0
- Reading group "Ethics and health", OZSW	2014	1.0
- Expertmeeting "Biology, medicine and the future of ageing", OZSW	2015	1.0
- Reading Group, Reiner Lab, University of British Columbia	2015 – 2017	2.0
- Seminars, National Core for Neuroethics, University of British Columbia	2016	1.0
- Seminar "Institutions for moral progress", Utrecht University	2015	0.5
- Workshop "Academic freedom", The Royal Netherlands Academy of Arts and Sciences (KNAW)	2016	0.5
- Seminar "Brein onder druk", Ministry of Justice and Security	2017	0.5
Other		
- Visiting International Research Student, National Core for Neuroethics, University of British Columbia	2016	

2. Teaching activities

	Year	Workload (ECTS)
- Diversity of ethical dilemmas, BA1 Medicine	2013 – 2017	2.5
- Academic writing and argumentation, BA2 Medicine	2013 – 2018	2.0
- Essay writing, BA2 Medicine (lecture, preparatory classes, grading, overall coordination)	2016 – 2018	4.0
- Bachelor Essay, BA3 Medicine (grading)	2013 – 2018	2.0
- Medical ethics, BA Clinical technology	2015 – 2018	2.0
- Lifestyle and responsibility, MA Medicine	2016	1.0
- Medical ethics, general practitioners in training	2017 – 2018	3.0
- Supervision of thesis, student Biomedical sciences Leiden University	2015	0.5
- Supervision of theses, minor Ethics of health care	2015 – 2017	2.5
- Lecture "human enhancement and disability", minor Ethics of health care	2016	0.5
- Lecture "enhancement technology", Erasmus University College	2015 – 2016	1.5