

# Counterproductive criminal rehabilitation:

## Dealing with the double-edged sword of moral bioenhancement via cognitive enhancement

Elizabeth Shaw

(This paper is a pre-publication draft of: Shaw, E. Counterproductive Criminal Rehabilitation: Dealing with the Double-Edged Sword of Moral Bioenhancement via Cognitive Enhancement. *International Journal of Law and Psychiatry*. DOI: 10.1016/j.ijlp.2018.07.006)

### 1. Introduction

Given the high criminal recidivism rate, there is a pressing need to find new approaches to offender rehabilitation (Ministry of Justice, 2013). Recently, there has been increasing academic discussion of using biomedical interventions in rehabilitation programmes (see, e.g. Birks and Douglas, 2018). The discussion of biomedical criminal rehabilitation intersects with the literature on “moral bioenhancement” – the idea that biomedical interventions could be used in order to morally improve individuals' motivations, decision-making and conduct (and to decrease the likelihood of seriously immoral actions, including criminal behaviour). There is a growing consensus among advocates of moral bioenhancement that the most defensible type of moral enhancement would involve enhancing cognitive capacities and enhancing control over one's thought-processes, emotions and behaviour. This article will discuss a potential problem with attempting to use cognitive enhancement to rehabilitate offenders - a problem that, so far, has not been the subject of detailed analysis. Cognitive enhancement is a double-edged sword. The very same cognitive capacities that are required for moral reasoning, and which would be targeted by the envisaged enhancement technologies, are also the capacities that could be used for criminal behaviour. Therefore, there is a risk that cognitive enhancement might be counterproductive and increase the likelihood of reoffending. Enhancing the cognitive skills of badly motivated people may have a greater chance of making them more dangerous than of making them more moral. If the double-edged sword problem cannot be overcome, it would seriously undermine the case for moral bioenhancement. It would be a severe restriction on the usefulness of moral bioenhancement if the most promising means of achieving it could only be given to those who are least in need of moral enhancement – those who are already well motivated.

Much has been written about failed attempts at rehabilitation that have no effect on criminal behaviour and the resulting waste of time and resources. Even more problematic (although less frequently discussed) is the possibility that a rehabilitative intervention might actually make offenders worse. If an intervention increases the risk of reoffending,

arguably the state would bear some responsibility for causing the subsequent offences, rather than just failing to prevent them. The risk that treatments for criminal behaviour could be counterproductive is an issue of significant current concern in relation to existing (non-biomedical) rehabilitation programmes. In 2017 a Ministry of Justice study reported that sex offenders who participated in the Core Sex Offender Treatment Programme (Core SOTP) in England and Wales were more likely to reoffend than similar offenders who had not participated in the programme (Mews et al., 2017). Although, it is uncertain whether the results were biased by the limitations of the study, there is a real possibility that the treatment programme itself made the participants worse. It is therefore essential to anticipate ways in which attempts at rehabilitation might backfire, to develop strategies for mitigating this risk and to consider the ethical implications of any attempt to reverse the effects of the treatment.

As the types of cognitive enhancement discussed in this article (and their effect on reoffending) have not yet been extensively studied in criminal populations it is not possible to find direct empirical evidence that these interventions have criminogenic effects. Instead, I will consider the cognitive capacities which these interventions aim to enhance, and will discuss reasons for thinking that these cognitive capacities can enable criminals to carry out crimes and avoid detection. I will then suggest possible ways of mitigating the risk that could be posed by cognitive enhancement of offenders. The rest of the article will focus primarily on the following ethical and legal questions arising from this discussion. What factors should be taken into account when considering whether cognitive enhancement should be withheld from an offender who requests it and might benefit from it? What are the ethical differences between withholding and reversing cognitive enhancement and under what conditions, if any, would it be permissible to reverse cognitive enhancement? This article will draw attention to a range of (disputed) philosophical issues that are relevant to settling these questions (e.g. the acts/omissions distinction and rights to mental and bodily integrity) and will discuss the implications of certain positions about these philosophical issues for the topic of cognitive moral bioenhancement (rather than providing an in-depth defence of these positions).

The topic discussed in this article potentially has wider implications for non-biomedical interventions, e.g. talking therapy, education and training. However, this article will focus on biomedical interventions for the following reasons. Firstly, researchers have become increasingly interested in biomedical moral enhancement, because of the possibility that, in the future, biomedical techniques might alter mental capacities to a larger extent and faster than non-biomedical methods (Persson and Savulescu, 2012). The double-edged sword problem discussed in this article implies that the greater these supposed advantages, the greater the corresponding disadvantages, if the intervention turns out to be counterproductive. Secondly, the use of biomedical techniques to alter mental capacities suggests the possibility of biomedically reversing the effects of the intervention, if it turns out to be counterproductive; whereas the idea of

removing a capacity that an individual acquired through, e.g., education, gives rise to different practical and ethical worries that can only be briefly mentioned here.

## 2. Moral bioenhancement via cognitive enhancement

### 2.1. What is moral bioenhancement?

Unless otherwise indicated, this article adopts the following broad definition of the term “moral bioenhancement”: the use of biomedical interventions to increase conformity with moral norms (this article will focus specifically on norms against criminal behaviour), whether or not those interventions could also be classified as treatments. Similarly, the term cognitive “enhancement” will refer to biomedical means of improving cognition, whether or not this also involves a treatment. I have adopted this broad definition of enhancement for the following reasons. Firstly, a narrower definition of enhancement which excluded “treatments” given to individuals with disorders, would have made it necessary to enter into the controversy about whether certain conditions are “disorders” or merely involve “differences” in behaviour and personality. Specifically, this controversy relates to syndromes that are associated with criminality (and which might be targeted by moral bioenhancement), e.g. antisocial personality disorder, psychopathy and ADHD. It is not possible within the scope of this paper to engage with such debates, so this paper adopts a definition of “enhancement” that is neutral on this issue. Secondly, the traits that will be discussed in this article, such as poor planning skills, poor concentration and impulsivity, lie on continuum ranging from severe to normal. For these traits, there is no clearly non-arbitrary cut-off point between normal functioning and disordered functioning and the same biomedical interventions that are beneficial below the cut-off point (and therefore might qualify as “treatments”) may also have a similar effect above the cut-off point (Pugh, 2014). Therefore defining moral bioenhancement in a way that excluded “treatments” would have been somewhat arbitrary. Thirdly, the practical problem that will be discussed in this article could, in principle, arise regardless of whether the intervention is labelled a treatment or an enhancement – the problem that improving cognition might, in some cases, increase the chance of criminal behaviour. However, I will return to the treatment/enhancement distinction later in the article, when discussing whether the state has the right to deny cognitive enhancements to offenders who request them.

This article will focus primarily on moral-enhancement-via-cognitive-enhancement, since this specific type of moral enhancement has received growing attention in the academic literature. It will also occasionally refer to enhancing self-control, as self-control enhancement is thought to have similar advantages to cognitive enhancement (see ‘The Allure of the Cognitive Route to Moral Enhancement’, below). I do not mean to imply that there is always a sharp distinction between ‘cognitive enhancement’, ‘control enhancement’ and ‘emotional enhancement’. By ‘cognitive enhancement’, I mean interventions that are intended primarily to target cognitive capacities (e.g. alertness and concentration), but this is a matter of emphasis and it should be noted that interventions that are currently used for this purpose seem also to

have some effects on emotions (Vrecko, 2013). In this article, I aim to focus only on the double-edged sword challenge to moral bioenhancement of offenders, as this objection has received insufficient attention. In order to do this, it is necessary for the following sections to assume certain positions about the nature and possibility of moral bioenhancement.<sup>1</sup>

## 2.2. Examples of relevant cognitive capacities and potential cognitive enhancements

Cognitive enhancement might be achieved through pharmaceuticals, such as modafinil or methylphenidate. Modafinil is currently used to treat sleep disorders, such as narcolepsy. Methylphenidate is a treatment for attention deficit hyperactivity disorder (ADHD) and for narcolepsy. Both drugs can increase attention and improve working memory. Currently neither of these drugs is licensed for use as cognitive enhancers by the healthy, but they are sometimes used illicitly for this purpose, e.g. by students attempting to improve their revision, essay or exam performance. There is some evidence that giving methylphenidate to violent offenders with ADHD decreases recidivism (Ginsberg et al., 2012). It is possible that this is partly due to methylphenidate's cognition-enhancing effects, such as improving the ability to focus on relevant considerations when deciding how to act, including reasons against criminal behaviour. In addition, methylphenidate directly reduces impulsivity, which may also partly explain the effect on recidivism. As already mentioned, traits such as poor concentration and impulsivity (which characterize ADHD) exist on a spectrum. Biomedical interventions that effectively treat ADHD might in the future benefit offenders who have such traits, but who would not be diagnosed with a disorder.

A condition that is strongly associated with criminality, particularly violent offences, is psychopathy (Salekin et al., 1996). Psychopathic offenders have a high risk of recidivism (Douglas et al., 2006).<sup>2</sup> While only 1% of the population have this condition, it has been estimated that between 15% and 20% of the U.S. prison population have psychopathy (Kiehl and Hoffman, 2011). Recently, researchers have begun to challenge the traditional view that psychopathy cannot be successfully treated, but further investigations are still needed to establish whether biological treatments for psychopathic offenders are effective and safe (Choy et al., 2018). Much of the existing research discusses whether biomedical interventions might increase psychopaths' emotional responsiveness and decrease their aggressive impulsivity (for useful summaries see Glannon, 2014, Felthous, 2015, Ling and Raine, 2018). However, there is some evidence that psychopaths also suffer from cognitive problems, which might, in the future, be targeted by biomedical interventions. These cognitive problems include difficulties taking on board peripheral information that is outside their focus of attention (Koenigs and Newman, 2013); a tendency to make decisions in a way that fails to take account of the likelihood that they will regret the decision (Baskin-Sommers et al., 2016); and problems perspective-taking, in particular with automatically realizing what someone else is thinking (Drayton et al., 2018). These cognitive problems could contribute to psychopaths' immoral and criminal behaviour by interfering with their ability to avoid/extricate

themselves from conflict situations, learn from mistakes/punishment, and generally engage in practical reasoning (i.e. reasoning about how to act).<sup>3</sup> Currently, there are no proven biological interventions for these cognitive difficulties in criminal psychopaths, but as researchers gain a greater understanding of the relevant brain areas, new interventions which target these areas may become available. In addition to pharmacological treatments, cognitive enhancement might be achieved through deep brain stimulation (DBS), which involves implanting a small device into the brain that transmits electrical stimulation to specific brain regions (for discussion of DBS in relation to psychopaths' emotional deficits see Hübner and White, 2016, Choy et al., 2016).<sup>4</sup> A less controversial possibility is transcranial direct current stimulation (tDCS) – a non-invasive technique which delivers a constant low direct current to areas of the brain via electrodes placed outside the skull. There is some evidence that tDCS can affect perspective-taking in healthy humans (Sellaro et al., 2016). As the traits that characterize psychopathy also exist on a spectrum, it is possible that cognitive enhancers might reduce the criminal behaviour of offenders who have some of these traits, but who would not be diagnosed as psychopaths.

Neurofeedback can enhance the cognitive capacity to be aware of one's emotional responses/impulses (or their neural basis). Neurofeedback involves watching real-time images of one's brain functioning and thereby, over time, learning to alter one's own brain activity in response to this visual feedback. Studies have shown behavioural improvements in juvenile offenders who were able to re-train their brain-wave patterns and thereby gain more control over their impulses (Smith and Marvin, 2006).

### 2.3. The allure of the cognitive route to moral enhancement

The literature on moral enhancement includes a wide range of opposing views (for an overview, see Specker et al., 2014). However, there is a growing consensus that the most defensible type of moral enhancement would involve enhancing cognitive capacities and enhancing control over one's thought-processes, emotions and behaviour. This section will briefly indicate some of the supposed advantages of the cognitive approach to moral bioenhancement. My aim is not to defend the idea that it has these advantages, but to demonstrate how favourably this specific type of enhancement has been viewed in the literature, in order to show why the risk that such interventions could be counterproductive is of significance and deserves greater attention.

Schaefer and Savulescu (2016) advocate enhancing the cognitive capacities that are necessary for reliably making correct moral judgments, arguing that this approach avoids problems faced by alternative, more “substantive” approaches to moral enhancement. Substantive approaches aim directly to affect the outcome of moral reasoning, i.e. to instil beliefs, motivations and behaviour that the enhancers consider morally desirable.<sup>5</sup> In contrast, cognitive enhancement aims to increase individuals' ability to engage in the process of moral reasoning and of making moral judgements for themselves. For example, Schaefer and Savulescu discuss enhancing memory, the ability to acquire

factual knowledge and the ability to make logical deductions and inferences. These are general capacities that are used in both moral and non-moral reasoning. The cognitive enhancement approach thus, “allows one to avoid the controversial issue of imposing one's values on individuals while still promising moral improvement” (Schaefer and Savulescu, 2016, 9). Furthermore, this approach “avoids begging the question for or against particular views of morality. A more substantive approach... would require us to prejudge the content of ... moral judgments” (ibid). Paulo and Publitz (2017) have also argued that counteracting “failures of rationality by improving moral deliberation would be less worrisome than other forms” of moral bioenhancement, which might interfere with the freedom of opinion that is necessary in a democracy. They argue that cognitive enhancement, rather than undermining democratic rights, might be regarded as “creating and safeguarding necessary cognitive competencies for the democratic Citizen” (ibid, pp16–17). Earp et al. (2017) have also advocated cognitive enhancement as a promising route to moral enhancement, because it is a good means of enhancing “higher-order capacities to modulate one's moral responses in a flexible, reason-sensitive, and context-dependent way”. They claim that this approach is much less vulnerable to objections that have been raised against certain “emotional” moral enhancements. Attempting to bring about moral enhancement simply by increasing or decreasing the level of certain emotions is likely to misfire, as the same emotion is typically morally appropriate in some contexts, but not in others. For instance, decreasing someone's level of aggression could reduce the likelihood of unjustified hostility, but could also prevent the individual from using aggression appropriately, e.g. to defend herself or others from harm. Increasing a person's compassion could promote altruistic acts, but could also leave the individual paralysed by “compassion fatigue” if he encountered too many people in extreme distress. In contrast the value of general intelligence or self-control is less context-dependent. Ahlskog (2017) has identified another way in which enhancing compassion (or emotional empathy) is likely to misfire. He argues that, given a limited quantity of mental energy, increasing empathy for strangers is likely to decrease empathy for kin and vice versa. However, he argues that cognitive enhancement could increase the total available mental energy and lead to morally better conduct towards both kin and strangers.<sup>6</sup>

John Harris is one of the most well-known opponents of moral enhancement by emotion modulation, partly because such interventions could restrict people's freedom, by making it emotionally difficult or impossible to perform certain actions (Harris, 2008). However, he does favour cognitive enhancement, because it does not diminish freedom in this way. Furthermore, Earp et al. (2017) suggest that enhancing cognition and control could positively increase freedom. Relatedly, a number of writers have opposed the idea of the enhancer attempting to control the recipient's thoughts and behaviour. Such attempts could be viewed as objectification, treating the individual merely as a means, or failing to respect the individual's rational agency (see e.g., Reichlin, 2017, Hauskeller, 2017, Shaw, 2014). Since cognitive

enhancement does not enable the enhancer to control the outcome of the recipient's reasoning process, it seems less vulnerable to these objections.

To summarise: an increasing number of theorists claim that cognitive enhancement is the most promising route to moral enhancement. This does not mean, of course, that providing such interventions to offenders would be uncontroversial. Many theorists object to biomedical enhancement in general. Such objections include the idea that biomedical enhancement is “unnatural”, motivated by arrogance, would place unenhanced individuals at a competitive disadvantage and would undermine social solidarity (Earp et al., 2017).<sup>7</sup> In the specific context of attempting to morally enhance offenders via cognitive enhancement, it might also be objected that this would be unfair (if cognitive enhancement is not available to the public generally) and (to the extent that cognitive enhancement confers an advantage on offenders) it would appear to reward bad behaviour. It is beyond the scope of this article to consider the aforementioned objections. Instead, this article will focus on the following argument: that attempts to morally enhance offenders via cognitive enhancement pose an unacceptably high risk of being counterproductive and increasing the likelihood of reoffending, because enhancing the cognitive skills of badly motivated people has a greater chance of making them more dangerous than of making them more moral.<sup>8</sup> If successful, this last argument would seriously undermine the case for moral bioenhancement. It would be a severe restriction on the usefulness of moral bioenhancement if the most promising means of achieving it could only be given to those who are least in need of moral enhancement – those who are already well motivated.

### 3. Cognitive enhancement is a double-edged sword

#### 3.1. Ways in which cognitive enhancement of offenders might backfire

Cognitive enhancement might have counterproductive side effects that are not part of or consequent on the cognitive enhancement itself. For instance, aggression is a potential side effect of modafinil in some individuals and methylphenidate can cause irritability (Levy et al., 2014). Cognitive enhancers that are developed in the future might also cause counterproductive side effects, such as sexually compulsive behaviour.<sup>9</sup> However, this article will focus on a distinctive risk arises from the nature of cognitive enhancement itself. Cognitive enhancement is a double-edged sword, particularly when given to offenders. Exactly the same capacities that can enable an individual to behave more morally can also allow someone to become a more effective criminal, depending on how the individual chooses to exercise their capacities. This could happen in a number of ways – a) by increasing the offender's ability to carry out certain crimes, b) by increasing the offender's ability to evade detection and c) by undermining punishment's deterrent effect.

Most (if not all) forms of cognitive enhancement result in the acquisition of 'transferrable skills'. For instance, recall that Schaefer and Savulescu (2016) propose increasing memory and factual knowledge and the ability to make logical deductions and inferences. These improved rational abilities are transferrable - they can be used in order to engage in better moral reasoning, or to plan crimes or evade detection. Cognitive enhancements might enable certain offenders to commit types of crime that they had been unable to commit before, e.g. crimes involving advanced computer expertise (a type of crime that is likely to become increasingly common, due to the growing importance of computers in society) and complex financial frauds. Enhancing perspective-taking abilities might lead to compassion for potential victims, or alternatively, could increase the ability of criminals to deceive and manipulate their victims.<sup>10</sup> Strengthening willpower might be used in order to resist inclinations to break the law; or might enable offenders to become more effective at resisting inclinations to be law-abiding, e.g. arising from fear of punishment or the urgings of conscience.

Schaefer and Savulescu (2016) have persuasively argued that decreasing cognitive biases can increase moral reasoning ability. However, decreasing certain cognitive biases could also undermine punishment's ability to act as a special deterrent. For example, the availability bias is the tendency to estimate the probability of an event based on its availability in one's memory, even when this is not a reliable guide to the likelihood of the event occurring (Over, 2004). The availability of a memory is influenced by factors such as how recent the event was and whether it affected one personally. If a recently released offender is prone to the availability bias, this offender may be reluctant to reoffend because he overestimates the likelihood of being caught and punished again, due to the ease with which he can bring to mind his experience of previously being caught and punished. If a criminal rehabilitation programme were to decrease the offender's susceptibility to the availability bias, this might be counterproductive, resulting in the removal of a form of irrationality that is useful to the criminal justice system.<sup>11</sup>

It might be argued that, even if cognitive enhancement weakened punishment's deterrent effect, this would not matter, as offenders' new appreciation of the moral reasons for obeying the law would replace the weakened prudential incentives. However, the idea that enhancing offenders' rationality will lead them to reflect on whether they morally ought to be committing crimes in the first place seems overly optimistic. Many (perhaps most) criminals will be in the habit of thinking about how to commit crimes, and will not be used to reflecting on whether their actions are morally permissible. Once they are released back into the community, they are likely to find themselves in the same situations that they were in when they formed the habits that led to their initial offending and may also be associating with a criminal peer group. Therefore, even if cognitive enhancement facilitates moral reasoning when the individual is participating in a rehabilitation programme in prison, it could backfire when offenders return to their previous environment.

Furthermore, the motivational force of a moral reason that the offender grasps as a result of the cognitive enhancement may be weaker than the motivational force of a self-interested reason that occurs to the offender as a result of the cognitive enhancement. For instance, imagine that as a result of the cognitive enhancement, a new idea occurs to the offender about how to commit a profitable crime with a low risk of detection, and, as a result, the offender's temptation to break the law increases.

### 3.2. Ways of mitigating the counterproductive effects of cognitive enhancement

The risk that cognitive enhancement could make some offenders worse, does not necessarily mean that it should be ruled out altogether as a form of rehabilitation. There are ways in which this risk could be mitigated.

#### 3.2.1. Combining cognitive enhancement with moral dialogue/education/talking therapy

It is necessary to engage with the offender about issues such as the wrongfulness of his crime, the factors that led to it and ways to avoid reoffending (Shaw, 2014). As Earp et al. (2017, 9) put it, “the drug or technology in question [should be] used as an aid or adjunctive intervention to well-established “traditional” forms of moral learning or education (rather than used, as it were, in a vacuum)”.<sup>12</sup> Cognitive enhancement might make the process of moral dialogue or talking therapy more effective by enabling the offender to concentrate on the dialogue/therapy; and the insights gained during this process might remain with the offender once he has been released. However, it is important to bear in mind talking therapy must be carefully designed and conducted, as this technique also can be counterproductive. For instance, group therapy, by bringing the offender together with others in a similar position, might “normalize” the offender's behaviour and give him the opportunity to learn new criminal techniques (Mews et al., 2017). Talking therapies aimed at enhancing empathy have been counterproductive for some offenders, possibly by enabling them to become more manipulative (D'Silva et al., 2004).

#### 3.2.2. Tailoring the intervention to the individual offender

Cognitive enhancement might be more suitable as a form of rehabilitation for certain types of offending behaviour than for others. Consider the following hypothetical examples. 1) The highly egocentric, amoral offender. Such offenders are rarely motivated by moral norms and lack any alternative ideology, but regularly commit crimes for purely selfish reasons. This group might include certain career criminals and psychopaths. 2) The ideologically committed offender. These offenders, e.g. terrorists and extremists, act in accordance with misguided moral norms. 3) The negligent offender. These offenders may recognise and be motivated by moral norms, but they do not give sufficient weight/thought to moral considerations. 4) The impaired/inexperienced moral reasoner. These individuals have the potential to be motivated by moral norms, but they have difficulties engaging in moral reasoning, due to impaired cognitive abilities and/or lack of adequate moral education/role models in early life. 5) Akratic individuals: These individuals wish to act in accordance with moral norms, but fail to do so due to weakness of will. For example, they may be easily influenced

by others, or may suffer from addiction. It seems plausible that offenders in the first two categories would be most likely to misuse cognitive enhancement (unless sufficient supplementary measures were taken to mitigate this risk). It may be easier to rehabilitate offenders in the other three categories through measures including cognitive enhancement and/or moral education.

### 3.2.3. Taking measures to support prisoners post-release

Enabling offenders to remain in regular contact with a “mentor”, e.g. a probation officer, or therapist, might encourage the offender to avoid crime, after his sentence has been completed and could allow any counterproductive effects of the intervention to be detected early. Providing offenders with vocational training and providing incentives for employers to employ ex-offenders could reduce the risk that offenders will simply return to the situation that led to their original offending, but this time with an enhanced ability to offend. Such measures would require considerable public resources. However, they might generate economic savings in the long run, if they led to crime reduction.

### 3.2.4. Combining cognitive enhancement with emotional enhancement

It might be argued that combining cognitive enhancement with emotion modulation, e.g. decreasing aggression and/or increasing affective empathy,<sup>13</sup> might reduce the risk that offenders will misuse their cognitive capacities for immoral purposes. However, as noted above, advocates of cognitive enhancement have expressed considerable reservations about trying to biomedically increase or reduce the levels of certain emotions. Earp et al. (2017, 9) have suggested a more sophisticated approach – enhancing empathic control – an ability, firstly, to identify when it is morally desirable to feel or be motivated by empathy and, secondly, the ability to feel and act accordingly. However, enhancing empathic control would not provide a guarantee against the misuse of enhanced cognitive capacities, since, on this approach, offenders would retain the ability to suppress their feelings of empathy.<sup>14</sup>

Given that there is no failsafe way of preventing cognitive enhancement of offenders from being counterproductive, it is important to consider when (if ever) it would be permissible to a) withhold cognitive enhancement and b) to reverse the effects of cognitive enhancement in circumstances where it carries a significant risk to public safety. These issues will be considered in the remainder of this article.

## 4. Withholding cognitive, moral enhancement

The risk that cognitive enhancement will increase the likelihood of criminal behaviour raises the following question: if there is reason to believe, in a particular case, that this risk may exist, is it permissible to withhold an intervention that the individual requests and that otherwise appears to be in his best interests? Theorists are beginning to discuss the ethical pre-conditions for offering moral bioenhancement to offenders (e.g. Bomann-Larsen, 2013, Shaw, 2014, Ryberg, 2015). There is also debate concerning whether moral bioenhancement should ever be imposed on a mandatory basis

(Douglas, 2014; Focquaert, 2014; Shaw, 2014). However, so far, a thorough investigation of the ethical pre-conditions for withholding moral bioenhancement from offenders has not been undertaken.<sup>15</sup>

As far as the law is concerned, just because a patient strongly desires a particular treatment does not mean that the patient is legally entitled to it. As a general rule, the courts are extremely reluctant to order doctors to provide treatments that the doctors do not wish to provide.<sup>16</sup> Although doctors are subject to a duty of care to their patients,<sup>17</sup> the law in the UK is not entirely clear about the circumstances under which this duty requires them to provide treatment. There are at least three situations where doctors are not required to provide treatment: a) a competent patient refuses treatment,<sup>18</sup> b) the treatment is not in the best interests of the patient<sup>19</sup> or c) treatment is reasonably withheld from some patients, due to scarce resources.<sup>20</sup> It is unclear whether doctors have even more leeway when deciding whether or not to provide enhancements, particularly as there is no legal definition of what an “enhancement” is.

Nielsen (2013), 415) argues that people have a moral entitlement to health care, “due to the mere fact that they are persons—that is, being a person entails the right to be able to live a life of normal human functioning—and thus these entitlements are inalienable... Furthermore, you cannot even lose your right to basic health care to some degree compared to others, since you do not become less of a person by being a bad person.”<sup>21</sup> However, it might be objected that, if there is an inalienable, personhood-based right to receive health care, this right only extends to treatments and therefore does not cover cognitive, moral enhancement of offenders. On this view, it might be argued that enhancements should be withheld from offenders, if those enhancements carry a risk to others, but that it is harder to justify withholding treatments. This line of argument assumes that there is a meaningful, ethically significant distinction between treatment and enhancement - an idea that has been contested (see, e.g. Levy, 2007, Erler, 2017). So far, in this article, I have been using the term “moral enhancement” and “cognitive enhancement” to include interventions that improve moral or cognitive functioning, whether or not they might also be considered “treatments”. At this point, however, it is important to consider the possible relevance of the treatment/enhancement distinction.

Theorists who believe that a meaningful treatment/enhancement distinction can be drawn have not reached a consensus about how to do so. A treatment might be regarded as an intervention that is necessary to maintain or restore health, but given that there are various different conceptions of “health” and “disorder”, an intervention to improve cognitive functioning might constitute a “treatment” according to one conception, and an “enhancement” according to another (Carter, 2017). Different definitions of health and disorder might be suitable for different purposes (Reimer, 2008). For example, biostatistical accounts of health and disease, which emphasise statistically subnormal biological functioning (e.g. Boorse, 1977; Daniels, 2000) might be particularly relevant to medical researchers investigating whether certain conditions have a common aetiology and whether these conditions might be responsive to the same forms of treatment. However, normative conceptions of illness (e.g. Neilson's account which includes deviation from normal social

functioning) might be more appropriate to ethicists considering whether individuals have a right to treatment.<sup>22</sup> I will consider two situations in which withholding a cognitive intervention seems particularly ethically problematic, regardless of whether the intervention is regarded as an ‘enhancement’ rather than a ‘treatment’ according to biostatistical definitions of ‘treatment’.

Firstly, imagine that someone suffers a brain injury resulting in a reduction in their cognitive capacities compared their own prior level of functioning. It seems ethically questionable to deny them an intervention to mitigate the effects of the injury (if such an intervention were available) merely because the individual's level of cognitive functioning post-injury is statistically average. Denying treatment still seems problematic even if there were a risk of the individual misusing their restored cognitive capacities. This example does involve mitigating the effects of an injury and so might be considered a treatment on certain biological conceptions of treatment. However, it is not obvious that “normal” age-related cognitive decline should be dealt with any differently (if an intervention to reverse this decline were available) merely because this intervention might be classified as an enhancement. The intuition that people have a right to be restored to an earlier level of cognitive functioning might be based on the idea that one's cognitive capacities are central to one's identity as a rational agent. The state has an obligation to respect (and arguably to foster) the rational agency of its citizens. Although the state's obligation to respect rational agency surely does not ground a right to state-funded cognitive enhancement under all circumstances (e.g. transforming an offender who had always been of average intelligence into a genius, if such a transformation were possible); the obligation to respect and foster rational agency, combined with a concern for the individual's identity might ground a right to restoration of diminished cognitive capacities.

A second situation in which denying cognitive enhancement becomes problematic is when the individual's capacities are below normal compared with a relevant reference group. Offenders with serious learning disabilities plausibly have a right to available treatments, even if there is a risk that they might abuse their new cognitive skills. Furthermore, arguably certain individuals whose cognitive problems are still within the statistically normal range might also be covered by a personhood-based right to cognitive treatment/enhancement. Such a person's cognitive limitations (which she cannot help) might seriously interfere with her “ability to live a life of normal human functioning”, which plausibly includes social functioning. How could individuals whose cognitive abilities are within the statistically normal range be unable to live a life of normal social functioning? Here is one possibility: cognitive/self-control problems within the normal range, combined with other factors (e.g. emotional problems and an adverse environment) could result in problems with social functioning. To illustrate this point consider the following testimony from a woman who received a number of criminal convictions:

Personal Account

“I realised that I was different from other kids when I was at primary school. I remember having both the desire to do really bad things and then acting them out, like poking my mum in the eye with a pencil or ripping up the book she was reading. I really struggled at school with reading [partly]...because of my impulsiveness ...and used to steal money from my parents to pay other children to read the books I was supposed to so that I was able to tell the teacher the story. I thought I was evil inside and...I thought my whole life would be bad and nobody seemed to take my concerns seriously... I gravitated towards similar kids and started experimenting with soft drugs and alcohol at around 11 years old...[By the age of 13] I stopped attending school because I found it too difficult and... I hung around town shoplifting and drinking. I got cautioned by the police several times. I often got into physical fights both in and out of school and started carrying a knife. I never really remember being satisfied with what I was doing. I tried to run away from home on several occasions... My father had a terrific temper and we often got into verbal and physical fights ... When I finished school I left home and drifted through a number of manual jobs, not ever being able to complete the tasks required of me...I made quick and silly decisions; for example, I often stole cars and drove while drunk or drug-impaired. I got involved with credit card fraud and...I spent a brief time in prison on drugs-related charges too...[I continued] to indulge in high-risk behaviour, which led to a serious motorbike accident that has left me disabled.”<sup>23</sup>

It is plausible that this individual's ability to engage in normal social functioning was seriously impaired. She struggled to succeed academically, to hold down a job and to stay on the right side of the law. She found much of her behaviour “unsatisfying” and found that her early life was shaped by “bad desires” that she did not endorse. She also sustained serious physical injuries due to her poor decision-making. In addition to the passage quoted, she described how for many years she seemed unable to alter this pattern of behaviour. This led to or exacerbated her low self-esteem, resulting in depression and suicide attempts. It seems that a range of different factors contributed to this individual's impaired social functioning. She experienced cognitive and self-control problems, had negative emotions and an adverse environment (including a delinquent peer group, the availability of drugs and alcohol and conflict with a violent parent). Some of her problems received medical diagnoses, including ADHD and dyslexia.<sup>24</sup> However, it is possible to imagine someone whose problems with social functioning are similar to those described in this personal account, but whose cognitive difficulties and impulsivity fall just below the threshold for a diagnosis of ADHD or dyslexia. Call this hypothetical offender “Alison”.

Imagine that efforts have been made to address Alison's emotional problems and adverse environment and yet these efforts are insufficient, by themselves, to enable her to lead a life of normal social functioning. There are certain pressures and obstacles in her life that would be impossible/unethical to remove, e.g. it is impossible to turn back the clock and improve her upbringing/education and it would be unethical to forcibly separate her from her family and long-term friends, even though they might be a bad influence on her, if she wishes to continue these relationships. It is

conceivable that she would cope much better with these pressures and have a much better chance of leading a fulfilling, law-abiding life, if she had access to biomedical interventions to increase her cognitive capacities, e.g. concentration, reading and planning abilities, and self-control (together with other measures, such as counseling, support from a social worker etc). If that were the case, there would seem to be good grounds for providing her with these biomedical interventions, even if they were classed as form of cognitive “enhancement”, on a biostatistical conception of enhancement.

Arguably, Alison should still be provided with cognitive enhancement, even if there was a risk that she would misuse her new skills in order to commit crimes more effectively or evade detection. Cognitive enhancement seems necessary in order to give her an adequate opportunity to lead a fulfilling life and to place her on a more level playing field with other members of the community. Furthermore, it should be noted that although there is a risk she might choose to use her enhanced capacities to become a more serious offender, the intervention still improves her ability to be law-abiding. If the state fails to provide her with this intervention, and, as a result, her ability to obey the law is seriously impaired, the state's moral authority to punish her for future offending is arguably weakened, as, arguably, without the intervention, she is at a serious disadvantage compared with others in terms of her ability to obey the law.

To conclude this section, when deciding whether it is ethically permissible to withhold cognitive enhancement from an offender who requests it, it is particularly relevant to consider the following factors: a) Whether the cognitive enhancement is necessary to restore the recipient to a level of functioning that he/she had enjoyed at a prior point in his/her life; b) Whether the cognitive enhancement is necessary in order to enable the recipient to function on a roughly equal footing with other members of society.

## 5. Reversing failed attempts at cognitive, moral enhancement

### 5.1. Differences between withholding and reversing interventions

It cannot be assumed that the ethical pre-conditions for withholding a cognitive enhancement intervention are the same as the pre-conditions for reversing it, but these differences have not yet been sufficiently explored in the literature. This section will consider a number of possible ethical differences between reversing and withholding cognitive enhancement.

### 5.2. Interference with bodily integrity

Academics typically interpret the right to bodily integrity as protecting against all forms of active interference with one's body without one's consent (see, e.g. Douglas, 2014).<sup>25</sup> In contrast, failing to provide a medical intervention may not be considered a violation of bodily integrity at all, especially if the intervention would only treat a minor health condition or would constitute an enhancement. Intuitively, failure to provide treatment might only be considered a violation of

bodily integrity in more extreme cases - e.g. when the failure results in death or severe disability - or alternatively such failure might instead be considered an infringement of other rights, such as the right against degrading treatment.<sup>26</sup> (This way of delimiting the scope of the right to bodily integrity may be partly based on the acts/omissions distinction, discussed below.)

Actively interfering with the offender's bodily integrity might be the only way of reversing the effects of certain biomedical interventions. It is possible to imagine that the original intervention might have caused a long-term change in the offender's mental capacities and reversing the effects of the intervention might require another type of medication as an "antidote". In cases where the biomedical intervention took the form of an implant, the only way to reverse its effects might be an operation to remove the implant. If the device could be switched off without an operation (e.g. DBS), doing this without the recipient's consent could still arguably infringe the right to bodily integrity (Gasson and Koops, 2013). If it turned out that cognitive enhancement made a particular offender more dangerous, would it be permissible, for instance, to compel him to take the antidote or even undergo an operation to reverse the procedure? Currently, the majority of medical ethicists oppose the compulsory treatment of capacitous individuals (cf Douglas, 2014). The right to bodily integrity – understood as a prohibition against active interference with one's body - is also given strong legal protection, much stronger than any positive right to receive interventions. As a general rule, it is only legally permitted to give medical interventions with the patient's consent. However, legislation in England and Wales permits the following exception: capacitous individuals can be given medication without their consent if they pose a danger to themselves or others as a result of a mental illness.<sup>27</sup> The legislative definition of mental illness is very broad - covering 'any disorder or disability of the mind'.<sup>28</sup> Yet, reversing the effects of cognitive enhancement would not fall within this exception, as the danger does not arise from any form of mental illness, but rather stems from the offender's decision to misuse enhanced cognitive skills. Should the exception be broadened? Douglas (2014) has suggested that it might be permissible to give compulsory medical interventions to dangerous, sane individuals, if they have been convicted of a crime. He relies on a comparison with incarceration. Most theorists assume that offenders, by committing certain crimes, have rendered themselves "morally liable" to incarceration. So, why, Douglas asks, have they not also rendered themselves morally liable to receive compulsory medical interventions aimed at reducing their dangerousness? Although Douglas focuses on the example of injecting an offender with an aggression-lowering drug, his reasoning could be applied in the context of administering a drug that reverses a counterproductive cognitive enhancement. If a cognitively enhanced offender misuses his enhanced capacities in order to reoffend, he could be liable to re-incarceration as punishment for that offence. So why is he not also liable to being forced to take an "antidote" to reverse the enhancement that he has misused?

I have argued elsewhere, in reply to Douglas, that forcibly injecting an offender with a drug is relevantly different from incarceration, because violating bodily integrity expresses disrespect in a way that restricting freedom of movement through incarceration (under humane conditions) does not (Shaw, 2016). I consider a disrespectful action to involve “behaviour, whose meaning appropriately understood by members of the cultural community in which the behaviour occurs, represents [the victim's] value as less than the value she should be accorded” (Hampton, 1991, 1670). Historically, there has been a trend towards increasing the protection afforded to bodily integrity and this trend arguably reflects an increasing acceptance of the principle that all persons, including offenders, are entitled to a basic level of respect. For instance, although corporal punishment of offenders was once a widespread practice, it has been replaced in many jurisdictions by measures such as incarceration. Douglas has responded that a practice such as mild, non-public corporal punishment of criminals expresses disrespect “primarily because it involves intentional infliction of pain”, rather than because of the interference with bodily integrity per se (Douglas, 2016, 5). However, in reply, it seems implausible that the intentional infliction of pain is the main reason why this practice is disrespectful, as it still seems disrespectful (and therefore an inappropriate form of punishment in a society that upholds human rights) even if the pain inflicted is very mild and much less distressing to the offender than incarceration would be. Furthermore, non-consensual physical interference is typically disrespectful, even when no pain at all is inflicted. For instance, consider the examples of forcibly shaving a prisoner's head,<sup>29</sup> or the sexual assault of an unconscious woman who never discovers what happened (see Gardner and Shute, 2000). The violation of the right to bodily integrity can explain the intuition that these examples involve disrespectful actions.

In the context of reversing cognitive enhancement, the concept of respect is particularly important. As discussed above, many theorists favour cognitive enhancement over other forms of rehabilitation because cognitive enhancement is considered more respectful of the offender's status as an autonomous, rational agent and an equal member of the moral community. The disrespect expressed by forcibly violating bodily integrity in order to reverse cognitive enhancement would undermine one of the main reasons why many theorists have favoured cognitive enhancement as form of rehabilitation in the first place.

### 5.3. Interference with mental integrity

The scope of the right to mental integrity has received much less attention than the right to bodily integrity (see, e.g., Bublitz and Merkel, 2014, Craig, 2016).<sup>30</sup> However, intuitively, forcibly reversing cognitive enhancement seems like a greater infringement of mental integrity (and more disrespectful of rational agency) than withholding cognitive enhancement. Diminishing an offender's cognitive capacities against his will would seem to express contempt for his rational agency. In contrast, if the state simply withholds cognitive enhancement from an offender (because the enhancement might increase his risk of reoffending), this would not send out the same disrespectful message.

#### 5.4. The acts/omissions distinction

As noted above, some methods of reversing biomedical interventions clearly involve actions that interfere with bodily/mental integrity and these interventions seem particularly hard to justify (at least on a compulsory basis). In contrast, withholding an intervention is an omission and might therefore be easier to justify. However, sometimes it is unclear whether reversing an intervention is an action or an omission. For instance, simply discontinuing certain interventions may be sufficient to reverse their effects. For example, the original biomedical intervention might consist in a course of pills or injections, whose effects would cease soon after the pills or injections were stopped. An implanted device might be fitted with a timer, causing it to turn off automatically after a fixed period of time. For example, Muller et al. (2014) discuss the possibility of fitting a DBS device with a timer, causing it to switch off automatically after certain intervals. The doctors would only switch it on again if the device did not have side effects that increased the risk of harm to others (DBS can sometimes cause compulsive harmful behaviour).

The question of whether discontinuing treatment counts as an act or as an omission is most frequently discussed in the context of life-sustaining treatment. For instance, turning off a life-support machine might be represented as the act of pressing the switch or as omitting to provide further treatment. Although omitting to save life is often legally permissible in circumstances in which actively killing is not, and the law treats switching off a life support machine as an omission, the courts have not yet articulated a clear, philosophically coherent account of the significance of the acts/omissions distinction in the context of withdrawing treatment. Lord Neuberger in *R (on the application of Nicklinson) v Ministry of Justice* referred to an “understandable discomfort with the notion that switching a machine off is actually an omission”.<sup>31</sup> Lord Mustill, in *Bland*, feared that the law's distinction between acts and omissions had become “morally and intellectually misshapen”, but nevertheless believed that “we must take it as it stands” (at least until law reform takes place).<sup>32</sup> This acknowledgement of the problems with the law in this area highlights the importance of carefully considering the ethical differences between withdrawing and withholding interventions, in order to provide the groundwork for new policy guidelines and possible legal reform.

Some ethicists believe that a philosophically important distinction can be drawn between withholding treatment and withdrawing it. Miller and Truog (2011) defend this distinction by arguing that withdrawing treatment “initiates” a sequence of events, whereas withholding treatment merely involves “permitting” a pre-existing sequence of events or state of affairs to continue “without intervention to stop it”. In the context of cognitive enhancement, it could be said that withholding cognitive enhancement permits a certain state of affairs to continue, i.e. limited cognitive functioning; whereas withdrawing cognitive enhancement could initiate a sequence of events, i.e. it could initiate cognitive decline from the enhanced condition back to the unenhanced level of functioning. If this line of reasoning is correct, then the active nature per se of withdrawing cognitive enhancement may be ethically significant. It seems implausible that this

consideration could render withdrawals of cognitive enhancements impermissible in all circumstances, but it may mean that they are harder to justify, i.e. weightier reasons would be required for withdrawing an enhancement, compared to withholding it.

### 5.5. Causing a loss versus failure to benefit

Even if the act/omissions distinction per se is not ethically significant in this context, it may track an ethically significant psychological phenomenon, in cases where the patient desires cognitive enhancement. In other words, the psychological suffering experienced by someone from whom cognitive enhancement has been withdrawn may be more severe than the distress experienced by someone from whom the enhancement is withheld, all other things being equal. There is evidence that people generally find losses more aversive than failures to receive a benefit. Even if loss aversion is irrational, it is nevertheless a psychological reality (Tversky and Kahneman, 1991).<sup>33</sup>

Alternatively, there may be a rational basis for loss aversion in the context of biomedical interventions. If someone begins to receive a successful intervention and experiences an improvement in their condition, being forced to go through the experience of losing the ground that they had gained might understandably be even worse than never receiving the intervention in the first place. The additional distress resulting from a discontinuation of treatment can be illustrated in the following case. In *D v UK*, an individual was imprisoned for smuggling cocaine from St Kitts into the UK, received treatment in the UK for AIDS and challenged the authorities' attempts to return him to St Kitts, where he would cease to receive adequate treatment and would face a real risk of dying in distressing conditions.<sup>34</sup> The European Court of Human Rights held that, although the situation he would face in St Kitts would not in itself constitute inhuman treatment, forcing him to withdraw from a beneficial course of treatment and face those conditions would be inhuman and would breach article 3 ECHR.<sup>35</sup> (This decision was not based on the narrow question of whether he was fit to travel.)

It is important to distinguish between two possibilities: either the additional psychological suffering attaching to losses may be the thing that matters, or, alternatively, the causing a loss/failure to benefit distinction may have ethical significance independently of the suffering involved. Different practical implications flow from each possibility.

If psychological suffering were the only important consideration, then the relative permissibility of causing a loss (by reversing cognitive enhancement) compared with failing to benefit (by withholding cognitive enhancement), would largely depend on the size of the loss or benefit. Even though undergoing a loss is typically experienced as somewhat more aversive than failing to receive a benefit of equal size, one would expect that failing to receive a sufficiently sizeable benefit would cause more suffering than experiencing a smaller loss.

In contrast, if the causing a loss/failure to benefit distinction has ethical significance independently of the suffering involved, then the relative size of the loss or benefit may not be so important. Instead, the key issues, on this approach, seem to be whether the individual has a right to the intervention and how robust that right is. When a person receives an enhancement, over time, their enhanced capacities become the new status quo. Intuitively, this strengthens the person's right to continue to have those enhanced capacities. Whereas, any right someone may have to start receiving an enhancement in the first place seems to be a more qualified right or a right that is more easily overridden by countervailing considerations. This principle that causing a loss (involving a disruption to the status quo) is harder to justify than failing to confer a benefit can explain our intuitions in other situations as well. For example, the fact that X has already formed family relationships in one country is an important consideration when deciding whether it would breach his right to private and family life to deport him to another country; whereas X may not have a right to move to a country in order to start a family in that country. The argument that causing a loss (through disrupting the status quo) is harder to justify than failing to confer a benefit is similar to Miller and Truog's argument in defence of the acts/omissions distinction, that it is harder to justify initiating a harmful causal sequence (i.e. harmfully disrupting the status quo) than it is to justify permitting a pre-existing harmful causal sequence to continue. However, it may be possible to agree that the causing a loss/failure to benefit distinction has moral significance, without necessarily claiming that this distinction directly corresponds to the acts/omissions distinction. For instance, imagine that a patient receives an implant that significantly boosts his cognitive capacities. Over time, the patient's enhanced cognitive capacities become the new status quo. The device has been fitted with a timer that causes it to switch off automatically after five years. The doctor who implanted the device refuses to switch it on again and, as a result, the patient's cognitive capacities decline. One might be reluctant to say that the failure to switch on the device was an "act", but willing to say that this omission caused a loss. In contrast, omitting ever to implant the device in the first place would merely have been a failure to confer a benefit.

To conclude this section: Although suffering should be given considerable weight in the context of withholding/reversing enhancements, it does not seem to be the only thing that matters. It is plausible that the causing a loss/failure to benefit distinction is of independent importance, as this distinction can best explain our intuitive responses to certain cases. It is debatable whether this distinction rests on a more fundamental distinction between acts and omissions.

## 5.6. The effect on narrative identity

The withdrawal of cognitive enhancements specifically, may have distinctively distressing and ethically unacceptable consequences, because cognitive abilities are intimately connected with one's identity. Even if withdrawing cognitive enhancement would not threaten personal identity (i.e. destroy the person), it might disrupt narrative identity. One's

narrative identity includes the psychological characteristics (such as beliefs, capacities, values and desires) that are “most central and salient in a given person's self-conception” (Degrazia, 2005, 266). It is crucial that changes to one's narrative identity “can be incorporated in one's life story in a coherent way, without compromising one's sense of self” (Focquaert and Schermer, 2015, 8). Forcibly reversing cognitive enhancement against the person's will could threaten narrative identity just as much as forcibly imposing a cognitive intervention (although the latter issue has received much more attention). In contrast, withholding a cognitive intervention that someone requests will typically not adversely affect narrative identity to the same extent (unless the recipient would otherwise face cognitive decline, or their cognitive limitations are so extreme that they cannot establish a coherent narrative identity in the first place).

The impact of cognitive enhancement on identity is powerfully explored in the novel, *Flowers for Algernon*. In this novel, the central character, Charlie, who has learning disabilities, is given an experimental intervention that gradually increases his intellectual capacities, until his intelligence is considerably above average. As his cognitive skills increase, this has a marked impact on the development of his personality (a bit like the complex interaction between intellectual and personality development that occurs as a child grows into an adult). Charlie strongly identifies with his new cognitive capacities and personality. However, he has a great fear of losing his new abilities (which eventually happens, as the effects of the intervention are not permanent). At one point, he is looking in the mirror and seems to glimpse a reflection of his pre-intervention-self. Terrified, he says, “I'm not going to give up my intelligence without a struggle. I can't go back down into that cave....I'm not going to give up – no matter what they all think...I'm going to keep what they've given me...”(Keyes, 1989, 177).

It might be argued that the intervention should not have been given to the individual in the first place if it were likely to have such a profound effect on his personal identity. However, this is not obvious, particularly, if the offender requested the intervention in the first place and if it took effect gradually so that the offender had the chance to stop the intervention if he did not like the way it was changing him. It might be worried that he could not give informed consent to an intervention that dramatically increased his cognitive capacities, as he could not possibly imagine before receiving the intervention what it would be like to have these intellectual capacities. This does not seem to be a sufficiently strong reason for ruling out such interventions in the first place. In other contexts it is ethically acceptable to give someone an intervention, even if they could not fully imagine its effects beforehand, e.g. enabling someone who had been blind from birth to see. Provided that all reasonable steps are taken to explain the nature of the change, to the extent that this is possible, and the recipient is not being coerced, the magnitude of the cognitive improvement does not seem, by itself, to make the intervention impermissible. However, the possibility that bringing about a high degree of cognitive enhancement might be counterproductive, e.g. by enabling an offender to manipulate victims and evade law enforcement more effectively, does provide strong reasons to hesitate before providing such interventions, because once the

intervention has been provided, there may be no way of going back, without causing the recipient serious harm. Reversing the procedure could have potentially devastating effects on the recipient's narrative identity and the only other way of protecting the public might be incarceration. Even if the recipient gave prior consent for the enhancement to be reversed if it turned out to be counterproductive, it would seem ethically dubious to enforce this pre-commitment, as bringing about cognitive decline (in a way that could disrupt narrative identity) without the recipient's current consent would arguably fail to show respect for his rational agency .36

## 6. Preliminary conclusions and future directions

What conclusions can be drawn from above-mentioned considerations about the permissibility of reversing cognitive enhancement? Firstly, in cases where reversing the intervention requires carrying out another intervention (e.g. giving an antidote drug or performing an operation), reversing the intervention without the recipients' consent would involve actively violating their bodily and mental integrity. Because of the disrespect that would be expressed by forcibly performing such a procedure, it should never be permitted on mandatory basis. Secondly, it is less clear whether bodily or mental integrity would be actively violated if the enhancement could be reversed by simply withdrawing an intervention (e.g. stopping a course of pills or injections, or programming an implant to switch off after a certain period of time and refusing to switch it on again). However, there are still strong reasons to doubt whether it can be justified to withdraw the intervention on a mandatory basis (or on the basis of a binding pre-commitment), in order to prevent possible harm to third parties, given that withdrawal would result in cognitive decline from the enhanced state (which is a loss, rather than just a failure to benefit) and given the potential impact on narrative identity. After all, if someone committed crimes by misusing his naturally high intelligence, or by misusing capacities that he had acquired through education it would be abhorrent to damage his brain to make him lose those skills. It is not clear that “artificially” enhanced abilities should be treated any differently, just because they might be destroyed in a way that could (arguably) be characterised as an omission. Further exploration of the analogy between biomedically enhanced cognitive capacities and cognitive skills acquired in other ways would be a fruitful avenue for future research. If mandatory reversal of cognitive enhancement is considered impermissible, this raises the question of how society could be adequately protected if offenders were cognitively enhanced and misused their new capacities to commit crimes. The enhanced offender could be punished for subsequent offences (just as non-enhanced offenders would be). A further question for future research is whether, or under what circumstances, it would be ethical to offer offenders the choice between reversing cognitive enhancement and facing restrictions on their liberty.

## 7. Conclusion

In this paper, I have drawn attention to the problem that cognitive enhancement, when used as part of criminal rehabilitation, is a double-edged sword. The very same cognitive capacities that are required for moral reasoning (and

which would be targeted by the enhancement technologies favoured by leading researchers on biomedical moral enhancement) are also the capacities that could be used for criminal behaviour. Therefore, there is a risk that cognitive enhancement of offenders might be counterproductive and increase the likelihood of recidivism. As, arguably, the state would bear some responsibility for subsequent offences caused by such counterproductive attempts at rehabilitation, it is essential to anticipate ways in which attempts at rehabilitation might backfire, to develop strategies for mitigating this risk and to consider the ethical implications of any attempt to withhold such risky interventions, or to reverse their effects. This article has discussed a range of ethical considerations, which are relevant to the cognitive enhancement of offenders (and which potentially have wider implications for withholding and reversing biomedical interventions in general and for counterproductive non-biomedical interventions, such as behavioural training programmes). The topic discussed in this article has not previously been the subject of detailed analysis, but ought to receive more attention from medical ethicists and legal scholars in the future.

## References

- R. Ahlskog, Moral Enhancement Should Target Self-Interest and Cognitive Capacity. *Neuroethics*, 10 (3) (2017), pp. 363-373.
- A. Baskin-Sommers, A.M. Stuppy-Sullivan, J.W. Buckholtz. Psychopathic Individuals Exhibit But Do Not Avoid Regret During Counterfactual Decision Making. *Proceedings of the National Academy of Sciences*, 113 (50) (2016), pp. 14438-14443
- D. Birks, T. Douglas (Eds.), *Treatment for Crime*, Oxford University Press (2018)
- L. Bomann-Larsen. Voluntary Rehabilitation? On Neurotechnological Behavioural Treatment, Valid Consent and (In) Appropriate offers. *Neuroethics*, 6 (1) (2013), pp. 65-77
- C. Boorse. Health as a theoretical concept. *Philosophy of Science*, 44 (4) (1977), pp. 542-573
- J.C. Bublitz, R. Merkel. Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination. *Criminal Law and Philosophy*, 8 (1) (2014), pp. 51-77
- D. Degrazia. Enhancement Technologies and Human Identity. *Journal of Medicine and Philosophy*, 30 (3) (2005), p. 261.
- A. Caplan. Denying autonomy in order to create it. The paradox of forcing treatment on addicts. *Addiction*, 103 (12) (2008), p. 191

- S. Carter. Could Moral Enhancement Interventions be Medically Indicated? *Health Care Analysis*, 25 (4) (2017), pp. 338-353
- A. Carter, P. Ambermoon, W.D. Hall. Drug-Induced Impulse Control Disorders: A Prospectus for Neuroethical Analysis. *Neuroethics*, 4 (2) (2011), pp. 91-102
- O. Choy, C.M. Berryessa, A. Raine. The Ethics of Biological Interventions on Psychopathic Prisoners  
*AJOB Neuroscience*, 7 (3) (2016), pp. 154-156
- O. Choy, F. Focquaert, A. Raine. Benign Biological Interventions to Reduce Offending  
*Neuroethics* (2018), p. 1, 10.1007/s12152-018-9360-0
- J.N. Craig. Incarceration, Direct Brain Intervention, and the Right to Mental Integrity – A Reply to Thomas Douglas.  
*Neuroethics*, 9 (2) (2016), pp. 107-118
- M. Crockett, L. Clark, M. Hauser, T. Robbins. Serotonin Selectively Influences Moral Judgment and Behavior Through Effects on Harm Aversion  
*Proceedings of the National Academy of Sciences*, 107 (40) (2010), p. 17433
- N. Daniels. Normal Functioning and the Treatment-Enhancement Distinction  
*Cambridge Quarterly of Healthcare Ethics*, 9 (3) (2000), pp. 309-322
- S. Darwall. *Welfare and rational care*. Princeton University Press, Princeton (2002)
- T. Douglas. Criminal Rehabilitation Through Medical Intervention: Moral Liability and the Right to Bodily Integrity.  
*The Journal of Ethics*, 18 (2) (2014), pp. 101-122  
CrossRefView Record in Scopus
- T. Douglas. Nonconsensual Neurocorrectives and Bodily Integrity: a Reply to Shaw and Barn  
*Neuroethics* (2016), pp. 1-12
- K. Douglas, G. Vincent, J. Edens. Psychopathy and aggression. *Handbook of Psychopathy* (2006), pp. 533-554
- L.A. Drayton, L.R. Santos, A. Baskin-Sommers. Psychopaths Fail to Automatically Take the Perspective of Others.  
*Proceedings of the National Academy of Sciences*, 115 (13) (2018), pp. 3302-3307
- K. D'Silva, et al. Does Treatment Really Make Psychopaths Worse? A Review of the Evidence

Journal of Personality Disorders, 18 (2) (2004), p. 163

B.D. Earp, T. Douglas, J. Savulescu. Moral neuroenhancement. The routledge handbook of neuroethics, Routledge (2017)

A. Erler. The Limits of the Treatment-Enhancement Distinction as a Guide to Public Policy. Bioethics, 31 (8) (2017), pp. 608-615

A.R. Felthous. The Appropriateness of Treating Psychopathic Disorders. CNS Spectrums, 20 (3) (2015), pp. 182-189

F. Focquaert. Mandatory Neurotechnological Treatment: Ethical Issues. Theoretical Medicine and Bioethics, 35 (1) (2014), pp. 59-72

F. Focquaert, M. Schermer. Moral Enhancement: Do Means Matter Morally? Neuroethics, 8 (2) (2015), p. 139

J. Gardner, S. Shute. The wrongness of rape. J. Horder (Ed.), Oxford Essays in Jurisprudence: Fourth Series, Oxford University Press, Oxford (2000), pp. 193-217

M.N. Gasson, B.J. Koops. Attacking Human Implants: A New Generation of Cybercrime. Law, Innovation and Technology, 5 (2) (2013), pp. 248-277

Y. Ginsberg, T. Hirvikoski, M. Grann, N. Lindfors. Long-Term Functional Outcome in Adult Prison Inmates with ADHD Receiving OROS-Methylphenidate. European Archives of Psychiatry and Clinical Neuroscience (2012), pp. 705-724

W. Glannon. Intervening in the Psychopath's Brain. Theoretical Medicine and Bioethics, 35 (1) (2014), pp. 43-57

J. Hampton. Correction harms versus righting wrongs: the goal of retribution. UCLA L. Rev., 39 (1991), p. 1659

R. Hare. Psychopathy Checklist Revised. (2nd edn), Multi-health Systems, Toronto (2003)

J. Harris. Moral Enhancement and Freedom. Bioethics, 25 (2) (2008), pp. 102-111

M. Hauskeller. Is it desirable to be able to do the undesirable? moral bioenhancement and the little alex problem. Cambridge Quarterly of Healthcare Ethics, 26 (3) (2017), pp. 365-376

M.L. Hoffman. Empathy and Moral Development: Implications for Caring and Justice  
Cambridge University Press (2001)

HD. Hübner, L. White. Neurosurgery for Psychopaths? An Ethical Analysis. AJOB Neuroscience, 7 (3) (2016), pp. 140-149

Keyes. *Flowers for Algernon*. Heinemann, New York (1989)

K. Kiehl, M.B. Hoffman. *The Criminal Psychopath: History, Neuroscience, Treatment and Economics*  
*Jurimetrics*, 51 (2011), p. 355

M. Koenigs, J.P. Newman. *The Decision-Making Impairment in Psychopathy: Psychological and Neurobiological Mechanisms*. *Handbook on Psychopathy and Law* (2013), pp. 93-106

N. Levy. *Neuroethics: Challenges for the 21st century*. Cambridge University Press (2007)

N. Levy, T. Douglas, G. Kahane, S. Terbeck, P.J. Cowen, M. Hewstone, J. Savulescu. *Are You Morally Modified?: The Moral Effects of Widely Used Pharmaceuticals*. *Philosophy, Psychiatry, & Psychology*, 21 (2) (2014), p. 111

S. Ling, A. Raine. *The Neuroscience of Psychopathy and Forensic Implications*. *Psychology, Crime & Law*, 24 (3) (2018), pp. 296-312

B.J. Lovett, R.A. Sheffield. *Affective empathy deficits in aggressive children and adolescents: a critical review*. *Clinical Psychology Review*, 27 (1) (2007), pp. 1-13

H.L. Maibom. *Moral unreason: the case of psychopathy*. *Mind & Language*, 20 (2) (2005), pp. 237-257

A. Mews, L. Di Bella, M. Purver. *Impact Evaluation of the Prison-Based Core Sex Offender Treatment Programme*. Ministry of Justice (2017)

F.G. Miller, R.D. Truog. *Death, Dying, and Organ Transplantation: Reconstructing Medical Ethics at the End of Life*. Oxford University Press, New York (2011)

Ministry of Justice. *Compendium of Reoffending Statistics and Analysis* (2013)

<https://www.gov.uk/government/statistics/2013-compendium-of-re-offending-statistics-and-analysis>

D. Moreira, F. Almeida, M. Pinto, M. Fávero. *Psychopathy: A Comprehensive Review of its Assessment and Intervention*. *Aggression and Violent Behavior*, 19 (3) (2014), pp. 191-195

S. Müller, H. Walter, M. Christen. *When Benefitting a Patient Increases the Risk for Harm for Third Persons—The Case of Treating Pedophilic Parkinsonian Patients with Deep Brain Stimulation*. *International Journal of Law and Psychiatry*, 37 (3) (2014), pp. 295-303

National Collaborating Centre for Mental Health. *Diagnosis and Management of ADHD in Children Young People and Adults*, The British Psychological Society and The Royal College of Psychiatrists (2009)

- L. Nielsen. Taking Health Needs Seriously: Against a Luck Egalitarian Approach to Justice in Health  
Medicine, Health Care and Philosophy., 16 (2013), pp. 407-416
- M.C. Nussbaum. Women and Human Development. Cambridge University Press, New York (2000)
- D. Over. Rationality and the Normative/Descriptive Distinction. In D. Koehler, N. Harvey (Eds.), Blackwell Handbook of Judgment and Decision Making, Blackwell (2004), pp. 3-18
- N. Paulo, J.C. Bublitz. How (Not) to Argue For Moral Enhancement: Reflections on a Decade of Debate  
Topoi (2017), pp. 1-15
- I. Persson, J. Savulescu. Unfit for the Future: The Need for Moral Enhancement. Oxford University Press, Oxford (2012)
- A. Piquero, R. Paternoster, G. Pogarsky, T. Loughran. Elaborating the Individual Difference Component in Deterrence Theory. Annual Review of Law and Social Science, 7 (2011), pp. 335-360
- G. Pogarsky, A. Piquero. Can Punishment Encourage Offending? Investigating the “Resetting Effect”  
Journal of Research in Crime and Delinquency, 40 (1) (2003), p. 95
- J. Pugh. Enhancing Autonomy by Reducing Impulsivity: The Case of ADHD  
Neuroethics, 7 (3) (2014), p. 373
- M. Reichlin. The Moral Agency Argument Against Moral Bioenhancement. Topoi (2017), pp. 1-10
- M. Reimer. Psychopathy Without (The Language of) Disorder. Neuroethics, 1 (3) (2008), pp. 185-198
- D.B. Resnik. The Moral Significance of the Therapy-Enhancement Distinction in Human Genetics  
Cambridge Quarterly of Healthcare Ethics, 9 (3) (2000), pp. 365-377
- J. Ryberg. Is coercive treatment of offenders morally acceptable? on the deficiency of the debate  
Criminal Law and Philosophy, 9 (4) (2015), pp. 619-631
- R.T. Salekin, R. Rogers, K.W. Sewell. A review and meta-analysis of the psychopathy checklist and psychopathy checklist-revised: predictive validity of dangerousness. Clinical Psychology: Science and Practice, 3 (3) (1996), pp. 203-215
- G. Schaefer, J. Savulescu. Procedural Moral Enhancement. Neuroethics (2016), pp. 1-12

- L.J. Schneiderman, N.S. Jecker. Should a Criminal Receive a Heart Transplant? *Medical Justice vs Societal Justice. Theoretical Medicine*, 17 (1) (1996), pp. 33-44
- R. Sellaro, M.A. Nitsche, L.S. Colzato. The Stimulated Social Brain: Effects of Transcranial Direct Current Stimulation on Social Cognition. *Annals of the New York Academy of Sciences*, 1369 (1) (2016), pp. 218-239
- E. Shaw. Direct Brain Interventions and Responsibility Enhancement. *Criminal Law and Philosophy*, 8 (1) (2014), pp. 1-20
- E. Shaw. The Right to Bodily Integrity and the Rehabilitation of Offenders Through Medical Interventions: A Reply to Thomas Douglas. *Neuroethics* (2016), pp. 1-10
- P. Smith, S. Marvin. Neurofeedback with Juvenile Offenders: A Pilot Study in the Use of QEEG-Based and Analog-Based Remedial Neurofeedback Training. *Journal of Neurotherapy*, 9 (3) (2006), p. 87.
- J. Specker, F. Focquaert, K. Raus, S. Sigrid Sterckx, M. Maartje Schermer. The Ethical Desirability of Moral Bioenhancement: A Review of Reasons. *BMC Medical Ethics*, 15 (2014), p. 67
- A. Tversky, D. Kahneman. Loss Aversion in Riskless Choice: A Reference Dependent Model *Quarterly Journal of Economics*, 106 (1991), pp. 1039-1061
- S. Vrecko. Just How Cognitive is "Cognitive Enhancement"? On the Significance of Emotions in University Students' Experiences with Study Drugs. *American Journal of Bioethics Neuroscience*, 4 (1) (2013), p. 4

## Endnotes

1

See Specker et al. (2014) for a useful overview of possible objections to moral bioenhancement.

2

The standard tool for diagnosing psychopathy is the 'Psychopathy Checklist Revised' (Hare, 2003). The nature of psychopathy has been disputed. For an overview of the history of this concept see Moreira et al. (2014).

3

For further discussion of psychopaths' difficulties engaging in practical reasoning see Maibom, 2005.

4

The risky nature of neurosurgery raises ethical issues that will not be discussed in this article.

5

Schaefer and Savulescu (2016) do not provide examples of potential substantive moral enhancement techniques, but these might include, e.g., altering serotonin levels in order to increase the aversion to personally harming others (Crockett et al., 2010).

6

Ahlskog cautions against cognitive enhancement for people who do not have a pre-existing capacity for emotional empathy. However, this would only exclude a minority of criminals from receiving cognitive enhancement, as most criminals have some capacity for empathy.

7

Earp et al. note that moral bioenhancement, unlike other forms of biomedical enhancement, is likely to benefit non-enhanced individuals. However, it should be noted that moral enhancement via cognitive enhancement might make non-enhanced people better off in one respect (if the recipient treats others more morally), but might also disadvantage non-enhanced people in another respect, e.g. if the cognitive enhancement places others at a competitive disadvantage.

8

This article will focus specifically on serious offenders (rather than badly motivated people in general) because the most serious types of immoral conduct are often also serious crimes. However, it should be acknowledged that non-offenders can be badly motivated and that offenders can be well-motivated.

9

Interventions that affect levels of the neurotransmitter dopamine can sometimes have this effect (Carter et al., 2011).

10

There is already some evidence that certain talking therapies aimed at enhancing empathy are associated with an increase in offending by psychopathic offenders (D'Silva et al., 2004).

11

Other biases may decrease deterability (see Pogarsky and Piquero, 2003 on the gambler's fallacy). Individuals vary in their susceptibility to different biases (Piquero et al., 2011).

12

Similarly, in the context of using methylphenidate as part of rehabilitation programmes for offenders with ADHD, Ginsberg et al. (2013) observe that, in order to be effective, this intervention must be combined with traditional forms of rehabilitation, e.g. learning, cognitive therapy and work training.

13

Affective empathy has been defined as “feelings that are more congruent with another's situation than with [one's] own situation” (Hoffman 2001), as opposed to perspective-taking (or cognitive empathy), which is the ability to understand what someone else is thinking or feeling (Lovett and Sheffield, 2007).

14

It is hard to imagine that an intervention could be developed that reliably ensured that recipients only suppressed their empathy when it was morally appropriate to do so.

15

Related, but distinct ethical issues have been discussed, to a very limited extent, in the context of providing treatments that restore physical health (including sexual function) to sex offenders (Schneiderman and Jecker, 1996) and providing treatments for Parkinson's disease, when the treatments, as a side-effect, may cause compulsive antisocial behaviour (Muller et al. 2014).

16

Re J (A Minor) (Wardship: Medical Treatment) [1990] 3 All ER 930. According to the British Medical Association, offenders are currently entitled to the same level of care as the general public (BMA 2012, p704, p85). For an argument that offenders should, in addition, be granted ‘an absolute right to rehabilitation, see Seigafo 2017.

17

Aintree University Hospital v James [2013] UKSC 67.

18

R (Jenkins) v HM Coroner for Portsmouth [2009] EWHC 3229 (Admin).

19

A NHS Trust v X [2014] EWCOP 35.

20

However, rationing decisions are usually made at a higher level than the individual doctor. In the UK, the Secretary of State has a duty to promote a “comprehensive health service” under the National Health Service Act 2006, s1. However, this duty can only be carried out within the constraints of the “resources available”. *R v Secretary of State for Social Services ex p Hincks* (1980) 1 BMLR 93, 95. Local authorities (and, more rarely, individual professionals) can be found liable for harm to a patient due to negligent rationing decisions, see, e.g., *DHSS v Kinnear* (1984) 134 NLJ 886. This article will discuss any ethical obligation the state might have to provide cognitive enhancement to offenders, all other things being equal. It is beyond the scope of this paper to discuss financial constraints that might be relevant in this context.

21

For further defence of this idea, see Darwall (2002) and Nussbaum (2000).

22

For discussion of the distinction between biostatistical and normative approaches, see Resnik, 2000.

23

National Collaborating Centre for Mental Health (2009, 74–76).

24

Ultimately this individual turned her life around due to a combination of medication, support from friends and therapists, and her own efforts. She gained a masters degree, pursued a professional career, entered a happy marriage and became a loving and supportive mother to her own child.

25

See also the right to integrity of the person, Charter of Fundamental Rights of the European Union (2000/C 364/01), article 3.

26

The prohibition of torture and inhuman or degrading treatment or punishment, European Convention of Human Rights, article 3.

27

Mental Health Act 1983, s3. This is also the case in several other commonwealth jurisdictions (Caplan, 2008).

28

MHA 1983, s1.

29

Shaving a prisoner's head was held to constitute degrading treatment in *Yankov v Bulgaria* (2003) 40 EHRR 36

30

It is covered by the right to integrity of the person, Charter of Fundamental Rights of the European Union (2000/C 364/01), article 3

31

[2014] UKSC 38, at para. 22.

32

[1993] 1 All ER 821, 885.

33

In addition to distress about being deprived of one's cognitively enhanced capacities, discontinuing the intervention might also cause withdrawal symptoms (e.g. disturbed sleep patterns). However, withdrawal symptoms are relatively rare with the drugs proposed for cognitive enhancement, e.g. modafinil or methylphenidate, when given at the doses currently used to treat ADHD and sleep disorders.

34

*D v UK* (1997), 24 EHRR No. 423.

35

The ECtHR does not prohibit all deportations that would interrupt medical treatment. Such deportations only breach art 3 in limited circumstances, where the consequences for the patient are particularly severe. Similarly, it might be argued that it is not justified to impose a complete prohibition on non-consensually withdrawing cognitive enhancement in all circumstances.

36

Cf. Muller et al. 2014 for a discussion of such pre-commitments in the context of DBS causing compulsive behaviour.