

Interventionist Advisory Brain Devices, Aggression, and Crime Prevention

Holmen, Sebastian Jon; Ryberg, Jesper

Published in:
Journal of Cognition and Neuroethics

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Holmen, S. J., & Ryberg, J. (2021). Interventionist Advisory Brain Devices, Aggression, and Crime Prevention. *Journal of Cognition and Neuroethics*, 8(1), 1-22. http://jcn.cognethic.org/jcnv8i1_HolmenRyberg.pdf

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact rucforsk@kb.dk providing details, and we will remove access to the work immediately and investigate your claim.

Journal of Cognition and Neuroethics

Interventionist Advisory Brain Devices, Aggression, and Crime Prevention

Sebastian Jon Holmen

Roskilde University

Jesper Ryberg

Roskilde University

Biographies

Sebastian Jon Holmen (ORCID [0000-0003-2774-941X](https://orcid.org/0000-0003-2774-941X)) is a PhD research fellow at the Department of Philosophy and Science Studies at Roskilde University.

Jesper Ryberg is Professor of Ethics and Law at Roskilde University.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). September, 2021. Volume 8, Issue 1.

Citation

Holmen, Sebastian Jon, and Jesper Ryberg. 2021. "Interventionist Advisory Brain Devices, Aggression, and Crime Prevention." *Journal of Cognition and Neuroethics* 8 (1): 1–22.

Interventionist Advisory Brain Devices, Aggression, and Crime Prevention

Sebastian Jon Holmen and Jesper Ryberg

Abstract

Some novel brain devices are able to predict neural events, making it possible for the device to advise its user to engage in the appropriate countermeasures before the event takes place. Other devices can automatically discharge such countermeasures on its user's behalf. In this paper, we consider some of the ethically questions that will arise if it becomes possible to combine such advisory and interventionist capabilities in a brain device to combat episodes of uncontrollable impulsive aggression. Specifically, if a device becomes available that can monitor and collect an offender's neural data, give him behavioural advice based on this data, and discharge countermeasures *unless* the offender actively keeps it from doing so, should such an interventionist advisory brain device be mandated to some offenders? In the following, we critically examine a range of plausible reasons to oppose such use related respectively to the device's capacity to monitor and collect an offender's brain data, its advice-giving feature, and its ability to discharge aggression-hampering treatment absent offenders' active dissent. We find that, surprisingly, none of the considered reasons can stand further scrutiny.

Keywords

Interventionist Advisory Brain Devices, Crime prevention, Explosive Aggression, Neurointerventions

1. Introduction

The ability of some novel brain devices to predict the occurrence of specific neural events, providing patients with the possibility of engaging in the necessary countermeasures or for an automated therapeutic activation system¹ to do so on their behalf, will potentially transform the way neurological disorders and neurodegenerative diseases will be treated in the future. For example, some devices tested on patients suffering from epileptic seizures use trained algorithms to detect the neuronal patterns related to a seizure and inform the patient through a handheld device how likely it is to materialize. This makes it possible for these patients to avoid many of these seizures

1. Such an automated system could, but need not be, similar to novel methods of brain stimulation in which an implant detects brain patterns related to the condition the implant is intended to counteract and automatically adjust the timing, location, and intensity of the delivered stimulation in response to these data (Malekmohammadi et al. 2016; Glannon and Ineichen 2016).

by pre-emptively administering the appropriate medication (Cook et al. 2013). It has been speculated that the ability of such advisory brain devices to forecast neural events may perhaps be employed in the future to prevent socially undesirable behaviour, such as uncontrollable aggressive and violent behaviour, which is similarly preceded by predictable neural changes (Gilbert 2015). In this paper, we will consider some of the ethical questions the development of such a device might raise.

Suppose that an advisory brain device is developed that is able to predict upcoming episodes of explosive aggression in individuals unable to control such violent impulses if they materialize, and could advise them to take the appropriate countermeasures (e.g., instruct them to leave a high-risk situation or simply to calm down) within a specified timeframe. Suppose, further, that this device had the additional feature of automatically being able to discharge electrical stimulation or medication to keep these explosive episodes from materializing *unless* the offender actively keeps it from doing so (e.g., by remote control) within the specified timeframe.² Finally, suppose that employing such *interventionist advisory brain devices* were effective in reducing re-offending rates among offenders suffering from such otherwise uncontrollable episodes of aggression, without them having to endure serious side-effects from being connected to the device or from its continued functioning. Should such devices then be mandated by the state to this group of offenders? Although other questions related to the ethics of employing brain devices and other neurointerventions on offenders are increasingly being addressed in the literature (e.g., Ryberg 2020; Birks and Douglas 2018; Douglas 2014; Ryberg 2012; Bublitz and Merkel 2014; Shaw 2014; Holmen 2020; Kirchmair 2019; Petersen and Kragh 2017; Ligthart et al. Forthcoming), the possibility of combining the advisory and interventionist capabilities of brain devices to combat recidivism has received virtually no scholarly attention.³

On the one hand, it seems clear that if such interventionist advisory brain devices prove effective in reducing re-offending, the possibility of preventing grievous harm to victims of future crimes of aggression (and their families) provides a strong moral

-
2. Either the offender could be informed in advance of the timeframe between the point at which the first piece of advice is given and the point at which the drug is being automatically discharged, or the action-guiding advice could be supplied via information to the carrier that the treatment will be initiated within a timeframe (for instance, the offender could receive a message along the lines: "Treatment will be initiated within one minute if the deactivation button is not pushed"). In the following, we will not engage in detailed speculation of what will constitute the optimal design of the device.
 3. The only ethical attention this possibility seems to have attracted is contained in a short comment from Jesper Ryberg (2015).

reason to employ them. Further, the offender himself (and his family) will be spared the deprivations that re-punishing the offender entails. Finally, and more generally, the resources that would otherwise be spent on re-punishing these offenders could be allocated to other (perhaps more morally desirable) projects. On the other hand, several important moral objections may be raised in response to a proposal to mandate these devices to some offenders. Such objections may relate to one or more of the three features we are here proposing the interventionist advisory device would have: (1) its monitoring and collection of information about an offender's neural environment; (2) its advice-giving feature; and (3) its ability to discharge countermeasures unless the offender actively keeps it from doing so. In the following sections, we will critically examine several plausible objections that could be raised in relation to each of these features. Specifically, the paper proceeds as follows. In Section 2, we critically examine whether the fact that employing the device involves the non-consensual monitoring and collection of information about parts of an offender's neural environment constitutes a plausible reason to oppose its use. Section 3 discusses whether it would be convincing to oppose the use of the device because it may present advice to the offender against his will. We shall then, in Section 4, discuss the intuitively plausible view that it is morally significant that the stimulation or medication delivered by the device would not be administered as the result of an active decision on the part of the offender, but would rather come about due to an act of omission. In Section 5, we consider whether there are good reasons for thinking it morally preferable not to prevent episodes of explosive aggression with brain stimulation or medication. Section 6 discusses whether it may be wrong to use such devices on offenders because it involves requiring them to have the device placed on their body against their will. The question of whether the potential impact of the advisory device on offenders' self-conception should lead us to judge against employing them will be confronted in Section 7. Finally, in Section 8, we summarize and conclude that, surprisingly, none of these *prima facie* plausible considerations can stand further scrutiny.

Before embarking on this discussion, however, a few comments are required regarding the scope of the paper. First, although we shall briefly comment on this question when concluding the paper, we do not discuss concerns in depth that may be raised narrowly in relation to the implementation of such a device. These will of course depend upon what the implementation requires precisely. The following discussion is nevertheless motivated by the assumption that ethical concerns regarding the mandated use of an interventionist advisory brain device will not be limited exclusively to the way such a system is implemented. Second, we do not wish to consider the potential misuse

of this device (e.g., the risk that it may be hacked).⁴ As with all other types of technology, an interventionist advisory brain device may of course be vulnerable to various kinds of misuse. In the following, however, we are interested in the arguments that might be raised against the mandated use of such a device under the assumption that it works properly.

2. Mental Privacy

To be able to present an offender with behavioural advice and (if he does not instruct it not to do so) to discharge countermeasures, the interventionist advisory device would need to gather information about the offender's real-time neural environment. It might, however, be suggested that the non-consensual monitoring and collection of information about parts of an offender's inner life constitutes a case of involuntary mind-reading, and as such is a violation of his moral right to mental privacy. Several commentators have indicated, that they believe this kind of rights violation by neurotechnological means to be a matter of great ethical concern (e.g. Lavazza 2018; Ienca and Andorno 2017), a view that we shall not dispute in the present paper.⁵ But should considerations regarding mental privacy lead us to reject the use of interventionist advisory devices on some offenders? In our view, one reason will suffice to show that the answer should be in the negative.

The reason is that it is plausibly not sufficient for a violation of mental privacy simply that the device monitors and collects information about the offender's neural environment. For a right to mental privacy to be violated by the device it seems, in our view, necessary that someone other than the offender should have, or gain, non-consensual access to the brain data the device collects. This is not something the device under consideration would be designed to do.⁶ Rather, as described in the introduction, the device would collect and process the information for the sole purpose of being able to present the offender with behavioural advice. By, admittedly rough, analogy, a surveillance camera installed in a person's home (even one placed there without his

4. For a discussion of this possibility, see for example Pycroft et al. (2016).

5. For some issues regarding the specification of the scope of such a moral right to mental privacy, see Ryberg 2017.

6. A possibility that might make a privacy-based objection relevant would be the storing of an offender's brain data by the state, but this is not what we are considering here.

consent) can hardly be said to violate his privacy if the person being filmed is the only one presented with the information observed by the camera.⁷ Thus, it seems that a concern for offenders' mental privacy – important as this may be in other contexts – is unable to block the use of the interventionist advisory devices on some offenders.

3. Unwanted Advice

If the non-consensual collection of brain data does not in itself constitute a convincing moral reason to oppose employing the advisory interventionist device on offenders, perhaps the advice-giving feature made possible by the collection of data does so. Specifically, should the fact that the advice presented to the offender – in the form of recommendations to take certain precautions – is unsolicited and might be presented to him against his will have us oppose the use of such devices? Generally speaking, there is no doubt that receiving unsolicited advice can sometimes be unwelcome to the advisee. We see no reason to think this might not also be the case with the behavioural advice provided by the interventionist advisory device. Upon further reflection, however, it is clear that the potential unwanted nature of the advice is not a convincing reason to oppose its proposed use.

First, the advice could be delivered to the offender in a way that makes it possible for him to avoid receiving it. If – as was the case with the use of advisory devices on patients with epilepsy cited in the introduction (Cook et al. 2013) – the advice is provided through a handheld device, for example, then the offender could simply place the device out of sight.

Second, even if we assume that the offender cannot avoid the advice from the device and that it is being presented to him against his will, it is hardly clear that this should raise a moral red flag. The reason is that it is very difficult to accept that it should generally be morally wrong to present someone with advice they do not want, particularly if it would be beneficial for the advisee to receive the advice.⁸ For example, it is surely difficult to

7. Although the non-consensual installation of the camera in the house might of course be wrong in itself because it, e.g., violates his property rights.

8. It might perhaps be objected that, in the present case, the advice would not be of benefit to the advisee because even if he decided not to act in accordance with it – for instance, by calming down or leaving the location – the risk that he would be involved in criminal conduct would nevertheless still be prevented by the drug that would subsequently be discharged. Therefore, the advice would not really place the advisee in a better situation than the one he would end up in anyway. However, this does not suffice to show that

see it as morally wrong to advise a person who is clearly agitated to try to calm himself down, even if he does not wish to receive this advice at the time at which it is given; or that one would be in the wrong to propose to a recovering alcoholic not to have the drink he has ordered in the event that he does not want this advice. Indeed, in both of these cases at least, it seems to us morally desirable to present the agitated person and the recovering alcoholic with such advice even if this is information they do not wish to receive. Many more examples could easily be cited. Furthermore, advice encouraging the advisee to act in ways that are beneficial for other people is at least sometimes also permissible, even if the advisee does not wish to receive it. It is, in our view, absurd to suggest that it would be wrong to advise your rich uncle to give a large portion of his wealth to combat extreme poverty if your uncle did not wish to be given this advice. The more general point is that, if the fact that someone does not wish to receive advice can be said to prohibit us from providing it, then many instances in which it is clearly desirable (or, at least, permissible) to present someone with advice they do not wish to receive would seem to be ruled out.

This is not to say that it can never be morally dubious to present someone with unwanted advice. For instance, if one is constantly bombarded with unwanted advice on how to act (say, every thirty seconds), this is sure to be highly disruptive of, for example, one's ability to direct one's own life. Similarly, if the interventionist advisory device were to constantly advise the offender to engage in countermeasures, this may be highly debilitating in similar ways and surely a strong moral reason to oppose the device's use. Whether some offenders might experience such an extreme stream of advice from the device, however, is ultimately an empirical question. But it seems unlikely to become a widespread practical problem given that the groups of offenders under consideration, i.e., those suffering from uncontrollable episodes of explosive aggression, presumably do not experience such episodes at a frequency that would result in constant behavioural advice from the device.⁹

the advice would not be beneficial. For instance, it might be the case that it would be more satisfactory to the advisee if he were to handle the situation himself by following the advice from the device, than if the aggressive outburst was prevented by the drug.

9. This presumption derives some plausibility from a study indicating that, on average, the highest number of episodes experienced during a single year by individuals suffering from intermittent explosive disorder (a disorder characterized by episodes of impulsive explosive aggression) is 27.8 episodes (Kessler et al. 2006).

All in all, based on the above considerations, we believe an objection to employing an interventionist advisory device on the basis of it potentially involving the presentation of advice to the offender against his will should be rejected.

4. Treatment Due to an Act of Omission

Suppose it is true, as has been argued in the previous sections, that the non-consensual collection of neural data and the advice-giving function of the brain device should not lead us to reject its use on some offenders. There might still be reasons related to the interventionist feature of the device, i.e., its ability to discharge countermeasures to prevent an explosive aggressive episode from materializing, that would speak against its use. One such reason relates to whether the intervention of the device would violate the offender's autonomy.

Given the prominence ascribed to the value of personal autonomy in contemporary bioethics and beyond, it is not surprising that a central question regarding the coercive use of neurointerventions on offenders has been whether such use violates an offender's autonomy and, if so, under what conditions (if any) this is morally permissible (see, e.g., Ryberg 2020, chapter 2; Douglas et al. 2013; Caplan 2006). As might be clear, however, this does not seem to be a concern that could plausibly be raised in relation to the interventionist function of the brain device under consideration here. Specifically, the fact that the offender would have the option of preventing the device from discharging countermeasures to stop the aggressive episode from materializing would arguably ensure that his autonomy regarding whether to receive the intervention remains intact.¹⁰ However, it may be objected that the fact that the device in this regards operates as an opt-out system (in which countermeasures are discharged as a result of an act of omission

10. It may be objected that a person who is about to experience an episode of impulsive aggression may not be competent to decide whether to receive treatment. That is, it may be suggested that the impulsive aggression might cloud his decision-making to such an extent so as to make him non-autonomous. However, whether this is indeed a viable concern is ultimately an empirical question regarding whether the device is able to predict the occurrence of the aggressive episode prior to it affecting his decision-making capacities. Furthermore, and perhaps more importantly, while it can surely be ethically dubious to subject a non-autonomous individual to a treatment (e.g., if the said treatment is not in the individual's best interest), the absence of a capacity for autonomy means that doing so cannot plausibly be a violation of his autonomy. Therefore, if an offender would in fact be non-autonomous due to a heightened level of aggression in most (or all) cases in which he must decide whether to receive the forthcoming treatment, the device can hardly be said to violate his autonomy.

on the part of the offender) provides ground for doubting that receiving the intervention from the device is truly the offender's autonomous decision. Specifically, two concerns regarding opt-out systems – often voiced in the debate about implementing such system to increase the availability of organs for transplantation – might be raised against this feature of the interventionist advisory device. Let us consider each in turn.

First, an opt-out system should not be too difficult for persons to opt out of. If it is too difficult (or even practically impossible) for persons to leave an opt-out system, it may plausibly be argued that this system impedes autonomous decision-making. This is surely true, but it does not seem a relevant challenge to mount against a proposal of using advisory interventionist devices on some offenders. More precisely, since the device could allow the offender to reject receiving the proposed treatment with a push of, for example, a button on the device itself or a handheld device, it should hardly be labelled as too difficult to avoid.

Second, and perhaps more importantly, in debates on the ethics of organ procurement it is sometimes argued that an opt-out system relies on the ethically dubious notion of *presumed consent*. That is, it is inferred from a failure to opt out of the system that a person would have given his consent had he been asked to express an explicit view on the matter. One major concern is that assuming consent in this manner overlooks that the person might simply have failed to register his dissent due to, for example, ignorance regarding his registration in the system. Similarly, it could be argued that, when the device discharges its countermeasures, the fact that the offender has not instructed the device not to do so (i.e., he has not opted out) cannot be assumed to mean he would have consented to the treatment had he been asked to express an explicit view on the matter. This cannot be assumed because the offender may simply be ignorant of the fact that the device is going to discharge countermeasures. Should this concern lead us to reject the use of the device under consideration? We believe the answer should be in the negative. First, it is widely believed that it is often morally appropriate to presume consent for treatment from individuals from whom it is difficult or impossible to collect explicit consent if not doing so will result in grievous harm to them. This is, for example, why most of us believe it is usually morally uncontroversial to subject an unconscious victim of a traffic accident to medical treatment without her explicit consent. Similarly, it could plausibly be argued that, if an offender is not in a position to explicitly consent or dissent to the device's treatment due (for example) to ignorance, it may be morally permissible to presume his consent if his explosive aggressive episodes are likely to lead to tremendous harm to him. Therefore, even if it is indeed the case that the device presumes an offender's consent, it is not obvious that this is always morally wrong. Second, it is not

at all clear that the interventionist advisory brain device we have described would in fact presume consent to its treatment. As described above, the brain device would, through a handheld device or the like, inform the offender prior to each instance of discharging its aggression-hampering treatment that the treatment will commence unless he actively keeps it from doing so. Thus, the offender would under normal circumstances seem to be fully informed about the forthcoming treatment and well-positioned actively to decide for himself whether or not he wishes to receive it.¹¹ Consequently, if and when the device delivers its treatment, it does not seem to have presumed the offender's consent.

In summary, the fact that the countermeasures discharged by the advisory interventionist device would come about due to an act of omission on the part of the offender does not seem to be cause for moral concern.

5. Preventing Explosive Aggression with Stimulation or Drugs

It may, however, be argued that, even if it is true that the interventionist feature of the device under consideration does not violate an offender's autonomy when it discharges its countermeasures, there is another, more basic, problem with this feature related to the means the device employs to prevent aggressive episodes. Specifically, it may be argued that, regardless of whether the treatment violates autonomy, it is simply morally inappropriate to prevent episodes of explosive aggression by means of brain stimulation or aggression-hampering drugs. As Martha Farah has pointed out, using these techniques to reduce aggression instead of more traditional approaches such as anger management classes "renders the effect no less therapeutic. Yet many people's intuitions raise a flag here. And if not here, then at the thought of more permanent interventions such as implanted stimulators or neurosurgery to achieve the same goals" (Farah 2002, 1126). However, while such a means-based argument against employing interventionist advisory devices may have great intuitive appeal, further scrutiny reveals it to face at least two serious challenges.

First, it should be acknowledged that other treatment schemes, such as cognitive behavioural therapy, may turn out to be more effective in preventing explosive aggression in offenders than the proposed device. This is ultimately an empirical question, and one we are currently not in a position to answer. However, should it turn out that

11. As noted above, we will not enter into a more precise discussion of how the device should be designed to be able to deliver the advice to the offender most successfully, but obviously there are many possibilities (e.g., vibration of the handheld device; a spoken message; a particular ringtone; etc.).

the interventionist advisory device is both safe to use and the most effective way of preventing explosive aggression in offenders, it is difficult, in our view, to see why the fact that this effectiveness is ensured by brain stimulation or the discharge of a drug should be considered morally problematic. Second, if the wrongness of employing the advisory interventionist brain device on offenders arises from the wrongness of the means it uses to prevent aggressive episodes, these means must surely also be considered morally wrong to use in other cases as well. This would, however, seem to imply that using brain stimulation techniques or drug-based treatment schemes to, for example, treat individuals suffering from mental health problems should *generally* be considered morally dubious. Surely, few (if any) would accept this view.

However, there is an alternative way that an opponent of using drugs or brain stimulation to prevent explosive aggressive episodes could motivate this view. He or she could argue that what is morally important is not simply that stimulation or medication is being used to prevent explosive aggressive episodes; rather, while these episodes may be socially undesirable, they are non-pathological, and non-pathological conditions should not be treated by means of drugs or brain stimulation. However, while this variation of the objection is surely more plausible than the variation considered above, it still faces at least one crippling challenge. The challenge starts from the observation that there are countless examples where we accept the use of drugs to treat non-pathological states. It is, for example, not usually considered morally problematic to take a sleeping pill to avoid the occasional sleepless night. The same is the case with occasionally taking a pain reliever to treat a headache or a sore knee. But if one insists that it is wrong to treat non-pathological states with drugs, then these and many other similar cases should be taken to involve acts that are wrong to perform. However, surely an account which implies that clearly morally innocuous acts (such as taking a pain reliever to combat the occasional headache) should be morally dubious to engage in is itself highly dubious.

To sum up this section, what seem to us the two most plausible variations of an argument against employing interventionist advisory device turning on the allegedly morally problematic means it uses to prevent episodes of explosive aggression, both seem to face the challenge of becoming overinclusive. It is not clear, at least to us, whether and, if so, how one could specify the objection in a way that avoids this problem.

6. Being Coerced to Wear the Device

As has been argued above, one of the advantages of the proposed interventionist advisory device is that it plausibly does not violate offenders' autonomy when it discharges its countermeasures since it informs offenders of the fact that the treatment is about to commence and leaves them free to reject the said treatment should they wish to do so. It may, however, be argued that there is another way in which the device may be an affront to an offender's autonomy. Specifically, it may be suggested that, even if it is assumed (as we have) that the device would not cause offenders discomfort or other side-effects that may plausibly be debilitating to their ability to exercise their autonomy, requiring an offender to have the device placed on his body against his will is an autonomy violation in its own right.¹² However, while it is surely plausible to hold that individuals should usually be considered the final arbiters concerning what is placed on their bodies, it is not obvious that this shows that interventionist advisory devices should not be used on some offenders.

First, in the context of criminal justice, offenders are often required to place items on their bodies that they may not wish to have placed there, but such requirements are usually not considered morally questionable. Some offenders serving their time outside of prison may, for example, be required to wear an electronic tag (usually placed around their ankle) that monitors their location. And, while perhaps more controversial, some jurisdictions require inmates to wear prison uniforms while incarcerated.¹³ It is not clear, at least to us, whether there is a relevant difference between (presumably morally acceptable) practices requiring offenders to wear these objects on their body and requiring them to wear a brain device.

Second, and more generally, the criminal justice system is rife with practices that reduce offenders' autonomy but are nevertheless usually considered morally permissible (or even desirable). For example, it is usually accepted that incarceration, at least in some

12. This objection may plausibly be framed, not in terms of an autonomy violation, but as a violation of offenders' right to self-ownership (see, e.g., Thomson 1990). However, the challenges we offer seems to us to apply regardless of the objection's specific moral foundation.

13. It should, however, be noted that the reason why requiring offenders to wear prison uniforms is most often held to be morally controversial is not that being coerced to wear them violates their autonomy; rather, it is that prison uniforms stigmatize offenders. It is also worth pointing out that, if our proposed brain device could be placed somewhere discreet on the offender (such as behind his ear), it could hardly be said stigmatize offenders in a similar way.

cases, can be a morally appropriate response to wrongdoing, even though it involves restricting the control of offenders over their own life in the form of, *inter alia*, constraints on free movement and association. Arguably, relative to these and other constraints entailed by incarceration, the violation of offenders' autonomy by a device being placed on their body against their will seems, at least in our view, much less severe.

In sum, it is not clear that one can consistently accept the use of incarceration and many other criminal justice practices that impede offenders' autonomy while rejecting the use of an interventionist advisory device on offenders on the basis of its being placed on their bodies against their will.¹⁴

7. The Impact of the Device on Offenders' Self-Conception

Suppose the interventionist advisory brain device, based on an offender's neural data, has predicted that the offender is about to experience an explosive aggressive episode and has advised the offender that it will discharge countermeasures unless he actively keeps it from doing so and that the offender has allowed the device to commence treatment. Suppose, further, that it is true (as we have argued in previous sections) that none of these steps should raise moral suspicion. There may yet be a reason to be sceptical of employing the device, because the changes to an offender's character induced by the treatment could potentially have an impact on their self-perception¹⁵ – that is, roughly, the experience offenders have of themselves after the intervention from the device. This could be said to be a relevant concern because some subjects, having received treatment through other brain devices (such as Deep Brain Stimulation), have offered such reports as "I don't feel like myself anymore" after the device was installed (Schüpbach et al. 2006, 1813; see also Baylis 2013, 514). Such post-intervention testimonials have been framed as experiences of loss of authenticity, self-estrangement, or self-alienation (e.g., Kraemer 2013; Gilbert 2018; Pugh, Maslen, and Savulescu 2017). There are, however, several reasons why an objection to the use of interventionist

14. To avoid misunderstandings, it should be underlined that we do not presuppose that the way the criminal justice system currently treats criminals is morally acceptable. In our view there are strong reasons against current trends of mass incarceration. All we are suggesting is that there some cases in which it is acceptable to use incarceration as a punishment.

15. A related worry is that the mere knowledge of having the device on or in one's body may have a negative impact on one's self-conception. Whether this would indeed be the effect of receiving our proposed implant is a question that requires further empirical scrutiny.

advisory brain devices based on the effect they may have on an offender's self-perception does not seem to us to be convincing.

An initial observation worth making is that self-estrangement and the like are not always experienced as negative by the affected person. In a recent study involving patients treated for Parkinson's disease by means of Deep Brain Stimulation, for example, the authors found that most of the patients considered the estrangement induced by the treatment to be restorative in nature (Gilbert et al. 2017). That is, roughly, to these patients, the changes to their character induced by the treatment were seen as restoring elements of their self that had been subdued by the disease (see also Pugh 2020). Thus, even if the impact of our proposed brain device might lead some or all of its recipients to experience self-estrangement, self-alienation, or inauthenticity, it cannot be straightforwardly concluded that this is a cause for moral concern. But even if we, *arguendo*, assume that offenders will generally not welcome changes to how they experience themselves post-intervention, it is not clear that this shows that the brain device should not be employed. At least, one cannot consistently accept the use of imprisonment, which studies have demonstrated to have numerous psychological effects on offenders that may plausibly affect their self-conception (e.g., Haney 2002, 82), while rejecting the use of the proposed device. As already noted, however, most of us believe that incarceration can at least in some cases be a permissible (or even desirable) way of responding to some kinds of wrongdoing.

More importantly, although it is of course ultimately an empirical question, there are reasons to speculate that the treatment delivered by the device under consideration will most likely not lead offenders to experience a loss of authenticity or the like. Much, of course, depends on what exactly causes such experiences to emerge, but the proposed device would at least not seem to give rise to some obvious source for such an experience. For example, the fact that the offender is fully informed about and has full control over the device when it delivers its treatment seems to make it unlikely that he will be troubled by the uncertainty of not knowing when the device is affecting his behaviour. These same features of the device may also ensure that the changes to his character induced by the device are not experienced as the product of an alien intrusion.

Furthermore, some subjects in fact report feeling *more* like themselves after receiving certain forms of neurotechnological treatment. The perhaps most cited example of this is the use of the antidepressant Prozac (see, e.g., Kraemer 2013, 486). It seems plausible, in our view, that a treatment meant to reduce impulsive aggression, like our proposed brain device, might well have a similar effect. After all, an offender's behaviour during episodes of impulsive aggression when he is not fully in control of how he acts is presumably not

experienced by the offender as reflecting who he really is. If this is true, then a device like the one proposed to give offenders more control over their behaviour may plausibly leave them feeling more in line with who they perceive themselves to be (see also Ryberg 2012, 233).

To conclude this section, we have offered reasons to doubt that the potential impact of the treatment from an interventionists advisory brain device on offenders' self-conception would be experienced negatively by the affected offenders, supposing that such effects are indeed likely to emerge. It has also been argued that there is, in fact, reason to doubt that the device would have such effects – indeed, the device may plausibly aid offenders in behaving in ways that are more in line with how they perceive themselves.

8. Conclusion

Some novel brain devices currently being investigated for use in a clinical setting are able to predict and advise patients about the emergence of specific neural events, thus making patients capable of engaging in the appropriate countermeasures. Similar devices are able to automatically adjust the timing, intensity, and location of treatment to counteract unwanted neural events. Inspired by these developments, we have considered some important ethical questions related to employing a hypothetical brain device combining such advisory and interventionist features to reduce recidivism among offenders suffering from severe problems with impulsive aggression. This device would have three features: (1) it would monitor an offender's brain data to predict upcoming aggressive episodes; (2) it would offer the offender behavioural advice; and (3) it would, unless the offender actively kept it from doing so, administer treatment by discharging measures to ensure that the aggressive episode does not materialize. There are *prima facie* plausible moral reasons to oppose each of these features, but we have suggested that, on closer scrutiny, none of the reasons considered convincingly rules out mandating advisory interventionist brain devices to the specified group of offenders. This conclusion does not, however, suffice to show that such devices should indeed be used. There are, for example, important ethical questions regarding the implementation of the device that we have not addressed. Much of this discussion would seem to hinge on exactly how invasive a procedure would be needed to implement the interventionist advisory device. Surely, if it is only possible to implement such a device by means of invasive brain surgery, this would provide a very strong reason to oppose its use. If, on the other hand, technological

developments mean that the device could be placed on the exterior of an offender's skull (e.g., behind an ear) and assert its effect through a needle placed just below his skin, arguments against mandating the device relating to its implementation seem less appealing. We have argued that the features of the device outlined above should not be cause for moral concern; therefore, *if* ethically unproblematic means of implementation are indeed developed, it is not obvious to us what principled reasons may be offered to oppose the use of interventionist advisory brain devices on some offenders.

References

- Baylis, Françoise. 2013. "'I Am Who I Am': On the Perceived Threats to Personal Identity from Deep Brain Stimulation." *Neuroethics* 6 (3): 513–26. <https://doi.org/10.1007/s12152-011-9137-1>.
- Birks, David, and Thomas Douglas. 2018. *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.
- Bublitz, Jan Christoph, and Reinhard Merkel. 2014. "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination." *Criminal Law and Philosophy* 8 (1): 51–77. <https://doi.org/10.1007/s11572-012-9172-y>.
- Caplan, Arthur L. 2006. "Ethical Issues Surrounding Forced, Mandated, or Coerced Treatment." *Journal of Substance Abuse Treatment* 31: 117–20.
- Cook, Mark J., Terence J. O'Brien, Samuel F. Berkovic, Michael Murphy, Andrew Morokoff, Gavin Fabinyi, Wendyl D'Souza, et al. 2013. "Prediction of Seizure Likelihood with a Long-Term, Implanted Seizure Advisory System in Patients with Drug-Resistant Epilepsy: A First-in-Man Study." *The Lancet Neurology* 12 (6): 563–71. [https://doi.org/10.1016/S1474-4422\(13\)70075-9](https://doi.org/10.1016/S1474-4422(13)70075-9).
- Douglas, Thomas. 2014. "Criminal Rehabilitation Through Medical Intervention: Moral Liability and the Right to Bodily Integrity." *The Journal of Ethics* 18 (2): 101–22. <https://doi.org/10.1007/s10892-014-9161-6>.
- Douglas, Thomas, Pieter Bonte, Farah Focquaert, Katrien Devolder, and Sigrid Sterckx. 2013. "Coercion, Incarceration, and Chemical Castration: An Argument From Autonomy." *Journal of Bioethical Inquiry* 10 (3): 393–405. <https://doi.org/10.1007/s11673-013-9465-4>.
- Farah, Martha J. 2002. "Emerging Ethical Issues in Neuroscience." *Nature Neuroscience* 5 (11).

- Gilbert, Frederic. 2015. "A Threat to Autonomy? The Intrusion of Predictive Brain Implants." *AJOB Neuroscience* 6 (4): 4–11. <https://doi.org/10.1080/21507740.2015.1076087>.
- Gilbert, Frederic. 2018. "Deep Brain Stimulation: Inducing Self-Estrangement." *Neuroethics* 11 (2): 157–65. <https://doi.org/10.1007/s12152-017-9334-7>.
- Gilbert, Frederic, Eliza Goddard, John Noel M. Viaña, Adrian Carter, and Malcolm Horne. 2017. "I Miss Being Me: Phenomenological Effects of Deep Brain Stimulation." *AJOB Neuroscience* 8 (2): 96–109. <https://doi.org/10.1080/21507740.2017.1320319>.
- Glannon, W., and C. Ineichen. 2016. "Philosophical Aspects of Closed-Loop Neuroscience." In *Closed Loop Neuroscience*, edited by Ahmed El Hady. Elsevier/Academic Press.
- Haney, Craig. 2002. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." *U.S. Department of Health & Human Services*. <http://img2.timg.co.il/CommunaFiles/19852476.pdf>.
- Holmen, Sebastian Jon. 2020. "Respect, Punishment and Mandatory Neurointerventions." *Neuroethics* (May). <https://doi.org/10.1007/s12152-020-09434-8>.
- Ienca, Marcello, and Roberto Andorno. 2017. "Towards New Human Rights in the Age of Neuroscience and Neurotechnology." *Life Sciences, Society and Policy*. <https://doi.org/10.1186/s40504-017-0050-1>.
- Kessler, Ronald C., Emil F. Coccaro, Maurizio Fava, Savina Jaeger, Robert Jin, and Ellen Walters. 2006. "The Prevalence and Correlates of DSM-IV Intermittent Explosive Disorder in the National Comorbidity Survey Replication." *Archives of General Psychiatry* 63 (June): 718–32.
- Kirchmair, Lando. 2019. "Objections to Coercive Neurocorrectives for Criminal Offenders – Why Offenders' Human Rights Should Fundamentally Come First." *Criminal Justice Ethics* 38 (1): 19–40. <https://doi.org/10.1080/0731129X.2019.1586216>.
- Kraemer, Felicitas. 2013. "Me, Myself and My Brain Implant: Deep Brain Stimulation Raises Questions of Personal Authenticity and Alienation." *Neuroethics* 6 (3): 483–97. <https://doi.org/10.1007/s12152-011-9115-7>.
- Lavazza, Andrea. 2018. "Freedom of Thought and Mental Integrity: The Moral Requirements for Any Neural Prosthesis." *Frontiers in Neuroscience*. <https://doi.org/10.3389/fnins.2018.00082>.

- Ligthart, Sjors, Tijs Kooijmans, Thomas Douglas, and Gerben Meynen. Forthcoming. "Closed-Loop Brain Devices in Offender Rehabilitation: Autonomy, Human Rights, and Accountability." *Cambridge Quarterly of Healthcare Ethics*.
- Malekmohammadi, Mahsa, Jeffrey Herron, Anca Velisar, Zack Blumenfeld, Megan H. Trager, Howard Jay Chizeck, and Helen Brontë-Stewart. 2016. "Kinematic Adaptive Deep Brain Stimulation for Resting Tremor in Parkinson's Disease." *Movement Disorders* 31 (3): 426–28. <https://doi.org/10.1002/mds.26482>.
- Petersen, T.S., and K. Kragh. 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. <https://doi.org/10.1136/medethics-2016-103492>.
- Pugh, Jonathan. 2020. "Clarifying the Normative Significance of 'Personality Changes' Following Deep Brain Stimulation." *Science and Engineering Ethics* no. 0123456789. <https://doi.org/10.1007/s11948-020-00207-3>.
- Pugh, Jonathan, Hannah Maslen, and Julian Savulescu. 2017. "Deep Brain Stimulation, Authenticity and Value." *Cambridge Quarterly of Healthcare Ethics* 26 (4): 640–57. <https://doi.org/10.1017/S0963180117000147>.
- Pycroft, Laurie, Sandra G. Boccard, Sarah L.F. Owen, John F. Stein, James J. Fitzgerald, Alexander L. Green, and Tipu Z. Aziz. 2016. "Brainjacking: Implant Security Issues in Invasive Neuromodulation." *World Neurosurgery* 92: 454–62.
- Ryberg, Jesper. 2012. "Punishment, Pharmacological Treatment, and Early Release." *International Journal of Applied Philosophy* 26 (2): 231–44. <https://doi.org/10.5840/ijap201226217>.
- Ryberg, Jesper. 2015. "Predictive Brain Devices, Therapeutic Activation Systems, and Aggression." *Ajob Neuroscience* 6 (4): 36–38. <https://doi.org/10.1080/21507740.2015.1094548>.
- Ryberg, Jesper. 2017. "Neuroscience, Mind Reading and Mental Privacy." *Res Publica* 23 (2): 197–211. <https://doi.org/10.1007/s11158-016-9343-0>.
- Ryberg, Jesper. 2020. *Neurointerventions, Crime, and Punishment: Ethical Considerations*. New York: Oxford University Press.
- Schüpbach, M., M. Gargiulo, M.L. Welter, L. Mallet, C. Béhar, J.L. Houeto, D. Maltête, V. Mesnage, and Y. Agid. 2006. "Neurosurgery in Parkinson Disease: A Distressed Mind in a Repaired Body?" *Neurology* 66 (12): 1811–16. <https://doi.org/10.1212/01.wnl.0000261103.50365.5b>.

Shaw, Elizabeth. 2014. "Direct Brain Interventions and Responsibility Enhancement." *Criminal Law and Philosophy* 8 (1): 1–20. <https://doi.org/10.1007/s11572-012-9152-2>.

Thomson, Judith Jarvis. 1990. *The Realm of Rights*. Cambridge, MA: Harvard University Press.