Neurointerventions as Criminal Rehabilitation: An Ethical Review

Jonathan Pugh, Thomas Douglas

[This is a pre-publication version. The final version is forthcoming in J Jacobs and J Jackson (eds.) *The Routledge Handbook of Criminal Justice Ethics*, Routledge.]

According to a number of influential views in penal theory,[1] one of the primary goals of the criminal justice system is to rehabilitate offenders. Rehabilitative measures are commonly included as a part of a criminal sentence. For example, in some jurisdictions judges may order violent offenders to attend anger management classes or to undergo cognitive behavioural therapy as a part of their sentences. In a limited number of cases, *neurointerventions*—interventions that exert a direct biological effect on the brain—have been used as aids to rehabilitation, typically being imposed as part of criminal sentences, separate treatment orders, or conditions of parole. Examples of such interventions include medications intended to attenuate addictive desires in substance-abusing offenders[2] and agents intended to suppress libido in sex offenders.

This chapter reviews some of the ethical issues raised by the use of neurointerventions as aids to rehabilitation. It focuses specifically on cases in which the neurointerventions are administered to criminal offenders, in response to the

[1] A number of theorists who advocate consequentialist views on punishment have claimed that punishment should contribute to the rehabilitation of the offender. See, for example, (Bentham, 1890; Carlen, 1989; Cullen, 1982). However, this claim is also supported by certain non-consequentialist views. For example, see (Hampton, 1984; Morris, 1981).

[2] For example, in the United Kingdom, methadone treatments have been imposed by the courts as part of Drug Treatment and Testing Orders. See (Eley et al., 2002; Hough et al., 2003)

commission of a particular crime, and under the provisions of the criminal law. We will refer to neurointerventions intended to aid rehabilitation as 'neurocorrectives' when they are used in this way. We will have nothing to say about the use of neurointerventions as aids to offender rehabilitation within the context of ordinary medical or psychiatric care.

Whilst the use of neurocorrectives is currently rare, there are good reasons to take seriously the possibility that it might, in the future, become more prevalent. Recent developments in behavioural and social neuroscience suggest that we may soon have at our disposal a far wider range of neurointerventions capable of facilitating rehabilitation in some criminal offenders. For instance, recent research has suggested that pharmacological agents could be used to affect certain traits that are linked to criminal behaviour, such as aggression, impulsivity, and the willingness to inflict harm on others (Bond, 2005; Crockett et al., 2013, 2010, 2008; Donovan et al., 2000; Khanzode et al., 2006; Nevels et al., 2010). It is also possible that deep brain stimulation (Lu et al., 2009), transcranial magnetic stimulation (Young et al., 2010), and neuro-feedback (Sitaram et al., 2009) could be used in similar ways.

Whilst such neurointerventions could potentially provide powerful new means of facilitating rehabilitation, it might be argued that they raise ethical problems that more traditional criminal justice interventions do not, and it is these problems that provide the impetus for this review. Briefly, one of the main problems is that neurocorrectives constitute a form of medical intervention, and it is a standard tenet of medical ethics that it is it is permissible to perform a medical intervention on a competent individual only if that individual has given his informed consent to that intervention. However, as we shall see, it is not clear whether it is possible to obtain valid consent to such an intervention from an incarcerated individual. On the other

hand, it is also unclear whether the requirements for informed consent are the same in criminal offenders as in other populations; it is often thought permissible to override the wishes of criminal offenders in ways in which it would not be permissible to override the wishes of non-offenders, and penal theory furnishes a number of candidate justifications for this view. The use of neurocorrectives in criminal justice thus raises problems that lie at the intersection of medical ethics and penal theory.

Before beginning to outline these problems in more detail, it is necessary to highlight a distinction that has important ramifications for the ethical debate regarding the use of neurocorrectives. The distinction pertains to the extent to which offenders have a choice about whether to undergo the neurocorrective. On the one hand, the neurocorrective might be imposed as a condition of parole or early release from prison. In such a scenario, the offender would essentially be given the option to either agree to undergo the neurocorrective and serve less time in prison, or to refuse to undergo the neurocorrective and serve his original prison term. To date, this is the most common way in which neurocorrectives have been used.

On the other hand, a neurocorrective might be imposed as a *mandatory* part of a criminal sentence; in such a scenario, the offender would not be able to avoid undergoing the neurocorrective. Although uncommon, there is some precedent for this sort of use; for example, androgen deprivation therapy is mandated as a part of the criminal sentence of certain sexual offenders in Florida.[3]

With this distinction in mind, we shall begin, in section I, by assessing two ethical objections to what we will call 'the neurocorrective offer': the offer of a choice between (i) undergoing a neurocorrective and receiving a reduced prison term, or (ii)

[3] For discussion, see del Busto and Harlow (2011); Douglas et al. (2013), p. 396.

receiving the normal prison term without any requirement to undergo the neurocorrective. In assessing these objections, we will assume that the circumstances in which the neurocorrective offer is made are not coercive, or not coercive in a manner that would invalidate the offender's consent. In Section II, we shall consider an objection that challenges this assumption, and that many have taken to be the most powerful objection to the neurocorrective offer. Finally in section III, we shall examine attempts to respond to this objection by denying that the offender's valid consent is required for the morally permissible use of neurocorrectives. We also consider the implications of such a denial for the moral permissibility of making neurocorrectives a mandatory part of criminal sentences.[4]

I Non Consent-Based Objections to the Neurocorrective Offer

i) Slippery Slope Arguments

It might be argued that making the neurocorrective offer would be the first step on a slippery slope to the future use of neurocorrectives in more problematic ways. It is possible to distinguish two different 'slippery slopes' that might be invoked. First, it might be claimed that even if it would be morally permissible for the state to give offenders the option of undergoing a neurocorrective in return for a reduced sentence,

[4] It should be acknowledged that in some US states the chemical castration of sex offenders is 'mandated' not only in the sense that the law makes such interventions compulsory for these offenders, but also in the sense that the law does not allow judges discretion to exclude chemical castration from the sentence (*California Penal Code*, 2012, 645(b)) . In some statutes, the mandatory use of chemical castration is contingent upon a determination by a court appointed medical expert that the offender is an appropriate candidate for such an intervention (*Florida Statutes Annotated*, 2013, §794.0235; *Louisiana Revised Statutes*, 2013, §14:43.6A).Whilst this practice raises other important moral issues, we shall not consider them in this chapter.

this can be expected to lead to a situation in which the state begins to impose neurocorrectives as a mandatory part of criminal sentences. Second, it might be claimed that offering neurocorrectives that are safe and effective in facilitating rehabilitation can be expected to lead to unsafe or ineffective neurointerventions also being offered. The worry here is that the neurocorrective offer might be made to certain offenders even though the neurointervention in question would either not aid their rehabilitation, or would pose excessive risks.[5]

The persuasiveness of any slippery slope argument depends in part upon the warrant that one has for claiming that the predicted bad outcome can be expected to occur once one has taken the first step on the putative slippery slope. In the arguments that we are considering here, we must assess the reasons that we have for believing that the predicted bad outcomes mentioned above could be expected to follow from making the neurocorrective offer.

In relation to the first slippery slope argument—the putative slippery slope to the mandatory imposition of neurocorrectives—it seems that there is little warrant for such a belief. To see why, it is illuminating to consider the fact that those who press this slippery slope argument as the sole basis for rejecting the neurocorrective offer believe that the neurocorrective offer *considered in itself* would be morally permissible. The reason that they object to it is that they believe that making the neurocorrective offer would lead subsequently to the mandatory use of neurocorrectives in criminal justice. They thus draw an important moral distinction between the neurocorrective offer and mandatory imposition of neurocorrectives regarding the former as permissible *in itself* while the latter is not.

[5] This objection has been raised in relation to the use of chemical castration as a criminal remedy (Lewis, 1953, p. 229; Tancredi and Weisstub, 1986).

It seems that the most plausible way to defend this moral distinction is to appeal to the moral significance of the offender's consent; in the case of the neurocorrective offer, an offender's undergoing a neurocorrective would be *consensual*, whilst in the latter case it would be *non-consensual*. This is a morally significant difference; in obtaining an agent's valid consent to some intervention, we are respecting their autonomous choice.[6]

Yet, if those who press this slippery slope objection appeal to the moral significance of consent in order to justify their initial claim that the neurocorrective offer is not morally problematic in itself, then their objection loses its force. If there is a morally significant difference between consensual and involuntary uses of neurocorrectives, then it is unclear why the former should lead to the latter.[7] Indeed, there is a clear and plausible moral principle that could be invoked against any such move, namely:

> *The Consent Requirement*: The provision of neurocorrectives to criminal offenders would be impermissible without the valid consent of the offender.[8]

The consent requirement seems to provide us with a non-arbitrary moral principle that can allow us to clearly and consistently permit the neurocorrective offer whilst forbidding the mandatory imposition of neurocorrectives. Moreover, it is a principle that is in fact *likely* to be invoked, since it has the backing of a rich philosophical and

[6] In turn, the principle of respect for autonomy, which finds its roots in the philosophy of both Kant and Mill, is a cornerstone of liberal thought, as well as modern biomedical ethics. See Kant (2004); Mill (1991) See also Beauchamp and Childress (2001) .
[7] See Douglas (2010) for a related general objection to a class of slippery slope arguments.
[8] For discussion, see (Douglas, 2014, p. 107).

political tradition. This, however, is not to say that it *ought* to be invoked. In section III, we shall suggest that there may be ways of challenging the consent requirement.[9]

It might be objected, at this point, that the existence of a clear and principled ground for distinguishing two practices is not always enough to prevent a slippery slope from one to the other. In some contexts where slippery slope arguments are invoked, there is a fear that some morally acceptable practice might ultimately lead to a morally unacceptable practice because the difference between the acceptable and the unacceptable practice is merely one of degree. This raises the possibility that acceptance of the unacceptable practice might be made by a series of small steps, none of which will be significant enough to draw attention. Thus, though there is a moral difference between the two practices, this difference is not likely to attract sufficient attention to prevent the slide. However, in the case that we are considering here, the difference between offering a reduced prison term in return for agreement to undergo a neurocorrective and mandating a neurocorrective is a clear and qualitative difference, so it is difficult to see why such a transition should be expected to escape notice.

Consider now the second slippery slope argument, according to which offering safe and effective neurocorrectives is likely to lead ultimately to the offering of unsafe or ineffective neurocorrectives. Given that the difference between a safe and effective intervention and an unsafe or ineffective one is often only a matter of degree (for example, it may consist in a difference in the dose of a drug), there is a possibility that the latter would be introduced by a series of small steps. There is also historical evidence that might be adduced in support of this argument: humanity has a

[9] Notice though that if the consent requirement can be challenged in this context, then, in the absence of another explanation, this may put pressure on the claim that the use of mandatory neurocorrectives would be an outcome that we have moral reason to avoid.

disturbing track record of erroneously assuming the safety and effectiveness of purported neurotechnological 'solutions' to criminality, with psychosurgery being only the most notorious example (Greely, 2008).

These considerations suggest that the slippery slope argument adverting to the risk that neurointerventions will be used when unsafe or ineffective should be taken seriously. However, it should be noted that this argument cannot be fully assessed until we have a clear empirical understanding of the ways in which the use of neurocorrectives could be monitored, and the likely success of these techniques in preventing unsafe or ineffective use. This is something we do not currently have.

ii) Neuro-Interventions Would Not Achieve The Objectives Of Criminal Justice

One of the main justifications for using neurocorrectives is that they promise to be highly effective in facilitating offender rehabilitation.[10] However, whilst rehabilitation might plausibly be claimed to be a central aim of the criminal justice system, it might be argued that the criminal justice system also has deterrent and retributive objectives. If this is right, then one might worry that even if the use of neurocorrectives might allow the criminal justice system to be more effective in facilitating rehabilitation, the use of such interventions may fail to achieve, or perhaps even run contrary to, the deterrent and retributive objectives of criminal justice.

[10] Whilst it is uncontroversial that neurointerventions could be effective in rehabilitating offenders in the superficial sense that the offenders will become less disposed to commit criminal offences, Vincent has argued that they may be unable to rehabilitate offenders into more responsible agents, since the way in which such interventions restore certain mental capacities might adversely affect other factors that undergird responsible agency such as authenticity and personal identity. See Vincent (2014).

One way to argue that neurocorrectives would fail to achieve the deterrent and retributive objectives of criminal justice is to claim that they would cause insufficient suffering.[11] In order for an intervention to achieve the retributive and deterrent aims of criminal justice, that intervention must inflict a certain degree of suffering. It might be feared that neurocorrectives—especially those with few side-effects—would not inflict sufficient suffering to deter criminal offending by others. Furthermore, since retributivists believe that retribution requires punishing offenders *in proportion* to their desert, some (although, as we shall explain below, not all) retributivists might claim that neurocorrectives would fail to inflict enough suffering to adequately achieve the retributive objective of criminal punishment.

One problem with this objection is that it could arguably be avoided by purposely developing neurocorrectives that have, and are intended to have, serious negative side-effects; interventions with such effects would most appropriately be conceived as alternative punishments, rather than alternatives to punishment.[12] In response, though, it might be argued that there is something specific about the *kind* of suffering inflicted by incarceration that makes it suited to realizing the deterrent or retributive goals of criminal justice.[13] Yet, it is not clear that this is the case; as Ryberg points out, "several retributivists accept that punishment need not consist only in imprisonment or fines" and can instead include sanctions such as "home detention, community service, and electronic monitoring" (Ryberg, 2012, p. 240).[14]

[11] This is an objection that has been advanced by a US court against the use of chemical castration as a criminal remedy (Idaho Court of Appeals, 1991, 120 Idaho 953, 821 P.2d 1008).
[12] Jesper Ryberg develops this point in Ryberg (2012), pp. 239-240.
[13] Though see, for a response to this suggestion, Ryberg (2012).
[14] It might be argued that, intuitively, it would be impermissible to impose neurocorrectives *for* their negative side-effects. However, if this is so, and if Ryberg is correct to think that the permissibility of doing so is consistent with retributivism, it may be that the appropriate response is to reject retributivism. This would in turn undercut the objection that neurointerventions are insufficiently retributive.

Furthermore, even if incarceration were particularly well suited to achieving the retributive and deterrent objectives of criminal justice, concerns about insufficient retribution and deterrence would at most constitute a decisive objection to the use of neurocorrectives as the *sole* response to criminal offending. They would not militate against the use of neuro-interventions alongside traditional punishments. Indeed, there might be certain advantages in using neurocorrectives to realise the rehabilitative goals of criminal justice while traditional punishments remedies are used to deter criminal behaviour and mete out deserved suffering; separating rehabilitation from retribution and deterrence might allow these elements to be more closely tailored to the circumstances in a particular case.

As Ryberg notes, the above concerns regarding neurocorrectives causing *insufficient* suffering only raise problems for so-called 'positive retributivists' (Ryberg, 2012, p. 239). Whilst both positive and negative retributivists claim that it is morally unacceptable to punish offenders more severely than they deserve, those who endorse positive retributivism also make the further claim that it is morally unacceptable to punish offenders *less* severely than they deserve. It is this latter requirement that was invoked above. However, the requirement that offenders not be punished more severely than they deserve might also be thought to ground an objection to the use of neurocorrectives. Both negative and positive retributivists might object to the use of neurocorrectives on the basis that they would inflict more suffering than the offender deserves.

How are we to ascertain the degree of suffering that it would be excessive to impose on an offender? Of course, much will depend on the particular offence in question. However, one useful general comparison that can be made here is to the degree of suffering involved in incarceration. If retributivists accept the assumption

that the state may permissibly subject some criminal offenders to incarceration, then they cannot consistently claim that neurocorrectives involving suffering no more severe than that involved in incarceration constitute excessive punishment. Indeed, it seems that the restrictions of movement and association entailed by incarceration reliably cause significant suffering. They frequently damage existing personal relationships whilst making it difficult to form new ones, they seriously restrict sexual and reproductive freedoms, they make it impossible to pursue most careers, and they generally prevent the realisation of many life-plans (Douglas, 2014, pp. 113–115). Yet retributivists often believe that prolonged incarceration is a punishment that is proportionate to desert in the case of certain offenders.

This suggests that one could not consistently object to the administration of *all* neurocorrectives on the grounds that they would inflict excessive suffering, for many neurocorrectives would cause less suffering than incarceration. For instance, it seems plausible to suppose that the negative side-effects of many psychopharmaceuticals would be comparable in their quality, frequency and severity to those of existing routinely used mind-active drugs such as fluoxetine (Prozac) and citalopram. Those effects typically cause much less suffering than the effects of incarceration outlined above.

To these side-effects we also, of course, need to add the *intended* effects of the neurocorrective, which may also cause a degree of suffering. For example, we can imagine that an aggression-lowering drug would cause suffering to an individual to the extent that the individual enjoys and wishes to retain the aggressive urges that the drug attenuates. However, there seems no reason to suppose that the intended effects of all neurocorrectives would involve substantial suffering. Consider a drug that attenuates only the most extreme impulses towards violent aggression, and suppose

that it is imposed on an offender who dislikes and reflectively rejects those violent impulses. Such a drug could be expected to cause significantly less suffering than incarceration. As such, it is not clear that a sentence that involved the use of neurocorrectives would suffering that is disproportionate to desert in all cases.

We shall return to the question of the suffering that neurocorrectives cause, and how they compare to the suffering caused by incarceration, in section III. To conclude our discussion of this objection though, it is worth noting that even if certain neurocorrectives on offer to offenders might cause suffering that could appropriately be deemed excessive, we are considering this objection on the assumption that the offender could autonomously choose to reject the neurocorrective offer. Indeed, if we believe that the criminal justice system should respect criminal offenders as rational self-governing agents,[15] then we might plausibly claim that competent offenders should be free to make their own informed decision about whether the suffering that a certain neurocorrective might cause outweigh the benefits of a reduced prison sentence.

II Consent-Based Objections to the Neurocorrective Offer

In the previous section, we suggested that one reason why imposing neurointerventions as a mandatory part of a criminal sentence might be deemed more problematic than making the neurocorrective offer is that the latter approach would allow the offender to consent to the use of a neurocorrective, whereas the former would not. Furthermore, we suggested that one of the primary reasons that consent is deemed to be morally significant is that, in obtaining an agent's valid consent to some

[15] See Murphy (1973) for a defence of this claim.

intervention, we respect the agent's autonomy. In turn, we may understand autonomy to refer to an agent's capacity for self-determination or self-government. Autonomy is one of the central values in contemporary medical ethics; however, it is a highly contested phenomenon and admits of numerous interpretations that we lack the space to survey here.[16] For our purposes, though, we may understand an agent to be autonomous to the extent that they are able to make decisions about what to do in accordance with their own desires and values, and in the absence of external controlling influences.

In medical ethics, it is often claimed that in order for an individual's consent to be valid (and reflective of autonomous choice), it must, amongst other things, be made in the absence of coercion (Beauchamp and Childress, 2001; Faden and Beauchamp, 1986). In their influential account of informed consent in medical ethics, Beauchamp and Childress claim that coercion occurs "if and only if one person intentionally uses a credible and severe threat of harm or force to control another" (Beauchamp and Childress, 2001, p. 94). One possible objection to the neurocorrective offer is that an offender could not validly consent to undergoing the neurocorrective because the offer is inherently coercive (Farah, 2002). This argument has commonly been raised against the practice of offering sex offenders chemical or physical castration in return for a reduced prison sentence, which we shall henceforth refer to as the 'castration offer'.[17] Since this offer is structurally similar to (and is indeed an instance of) the neurocorrective offer, it will be useful to consider the literature on the castration offer here.

[16] For some influential theories see Dworkin (1988) Christman (1991); Ekstrom (1993). For a useful overview, see (Taylor, 2005)

[17] See European Committee for the Prevention of Torture and Inhuman and or Degrading Treatment or Punishment (2009); Green, (1986); McMillan, (2014a); Scott and Holmberg (2003); Vanderzyl (1994).

Vanderzyl offers a clear statement of the objection under consideration in the following passage:

> . . . offering a convicted offender castration as an alternative to a lengthy prison sentence constitutes an inherently coercive practice rendering truly voluntary consent impossible (Vanderzyl, 1994, p. 140).

Proponents of this sort of argument are not always clear about *why* this sort of offer involves coercion. However, one justification that is sometimes suggested is that the dire circumstances that offenders facing incarceration find themselves in means that they cannot be said to make a free choice when they choose to undergo the castration. For instance, Green claims that in such a scenario:

> Freedom of choice is impossible because the convict's loss of liberty constitutes a deprivation of such a magnitude that he cannot choose freely and voluntarily, but he is forced to give consent to an alternative he would not otherwise have chosen (Green, 1986, pp. 16–17).

McMillan suggests that another way in which the castration offer might be coercive is if the offer were made to an offender with the intention that the operant reason for the offender's accepting the intervention would be the fact that the state will not release him if he refuses. Although the offer increases the range of options open to the sex offender, McMillan suggests that it may be understood to be coercive in so far as it "relies upon the undesirability of indefinite incarceration to coerce (the offender) into being castrated" (McMillan, 2014a, p. 587).

In this section, we shall consider two ways in which defenders of the neurocorrective offer might respond to the objection from coercion. The first response that we shall consider involves denying that the neurocorrective offer is coercive. One might begin to support such a denial by arguing that the explanations provided for why the neurocorrective offer is coercive are inadequate. For instance, many theorists have pointed out that one cannot claim that the offer is coercive simply by appealing to the offender's loss of liberty, because not all deprivations of liberty entail that the choices made in such circumstances are coerced (Bomann-Larsen, 2013, p. 68; Rosati, 1994, p. 161; Ryberg, 2012, p. 236; Wertheimer and Miller, 2014). The following example is commonly used to illustrate this point: Suppose that a patient suffering from a terminal illness is offered a high-risk, but potentially life-saving experimental operation; intuitively, this patient could validly consent to this offer, even though she is making a choice between two unappealing alternatives from a position of severely restricted liberty. Ryberg and Petersen (2013) have also used this example to argue that McMillan's appeal to the state's intentions in making the castration offer cannot adequately explain why the offer is coercive; after all, a doctor who offered her patient the option of undergoing this operation might make the offer intending that the patient's 'operant reason' for accepting it will be that he will die without it; yet, we do not seem to believe that this would make the offer coercive, as McMillan's argument seems to imply.

In response to this line of argument, defenders of the objection from coercion might seek to refine their analysis of why the neurocorrective offer is coercive so that their view is no longer subject to counter-examples of this sort. However there may be a more robust way of denying that the neurocorrective offer is coercive: one could

appeal to the conceptual claim that offers cannot, by their nature, be coercive. In the philosophical literature on coercion, it is often claimed that a necessary condition of coercion is that the coercer credibly *threatens* their victim (Nozick, 1969, p. 83; Stevens, 1988, p. 452) . Indeed, Beauchamp and Childress invoke this sort of condition in their understanding of coercion, delineated above. Since threats and offers are generally understood as mutually exclusive categories, this view supports Wertheimer and Miller's claim that "genuine offers are inherently non-coercive and therefore cannot compromise consent qua coercion" (Wertheimer and Miller, 2014, p. 592).

If this is to be a satisfying reply to the objection from coercion, there will need to be a convincing account of the relevant difference between threats and offers that can explain why only the former can be coercive. There have been several different approaches to the distinction between threats and offers in the philosophical literature, and we lack the space to adequately explore them all here.[18] However, one account that has been particularly influential in this area that we shall briefly describe is the so-called 'baseline' approach.

Both threats an offers announce an intention to bring about some outcome that is conditional on the recipient's actions. According to the 'baseline' approach, threats and offers are to be distinguished by the relationship of the proposed outcome to the recipient's baseline state of affairs, that is, how well off they would normally be in the absence of the proposal. Typically, threats announce a conditional intention to make the recipient worse off than the baseline state of affairs (think of a highway man who threatens to shoot you if you do not hand over your wallet), whilst offers do not; indeed, in some cases, offers announce an intention to make the recipient better off.

[18] For a description and analysis of alternative approaches, see Anderson (2011)

Theorists disagree about which baseline state of affairs should be relevant to our assessment of whether a proposal is a threat or an offer. On morally neutral baseline accounts, we might assess whether a proposal is an offer or a threat simply by considering whether it would make the recipient worse off than she would normally be.[19] The problem with such accounts is that the recipient's status quo situation might already include coercive elements. For instance, in his seminal essay on the topic of coercion, Nozick asks us to consider a slave owner who beats his slave each morning (for no reason connected to the slave's behaviour), and who one day proposes to the slave that he will not to beat him on that day if he performs some task (Nozick, 1969, p. 450). Whilst the slave owner announces a conditional intention to make the slave better offer than he would normally be in the absence of the proposal, it nonetheless seems that the slave is being coerced here.

Partly in order to account for why this sort of example involves coercion, theorists have developed moralized baseline accounts, according to which the baseline state of affairs that is relevant to our assessment of whether a proposal is a threat or an offer is understood to incorporate certain moral conditions. For instance, on Wertheimer's moralized baseline account, a proposal amounts to a threat if the proposer announces the conditional intention to make the recipient worse off than the recipient *ought* to be, either by violating the recipient's rights, or by failing to fulfil an obligation that the proposer has to the recipient. Such threats are coercive when the recipient has no reasonable choice but to comply with the proposal (Wertheimer, 2014).

Indeed, Wertheimer and Miller (2014) have invoked this account of coercive threats in defence of the claim that the castration offer is not coercive. On the

[19] As Feinberg suggests, there can be different interpretations of this. See Feinberg (1989), pp. 218–224.

assumption that imposing the ordinary prison term would not violate the offender's rights, offering the offender a choice between that ordinary term and a shorter prison term plus castration cannot be coercive on Wertheimer's account, since it does not involve a threat to violate the offender's rights if he does not comply. If the offender does not agree to undergo the castration, he will simply face the prison term to which he has made himself liable through his criminal conduct. A similar claim could be made with regards to other instances of the neurocorrective offer.

However, it should be acknowledged that moralized baseline approaches to the distinction between threats and offers have been challenged on many counts. [20] For instance (Feinberg, 1989, pp. 218–228) and Cohen (1977) have suggested that moralized baseline accounts are not always congruous with our intuitions, whilst Zimmerman (1981) has suggested that moralized baseline accounts do not link up in the right way with the underlying idea that coercion is wrong because it undermines freedom.

Furthermore, these writers have also objected to the fundamental claim that only threats can be coercive. For instance, Zimmerman claims that a proposal amounts to an offer if the recipient would prefer to move from his normally expected (i.e. morally neutral) pre-proposal situation to the proposal situation. However, he claims that offers (as well as threats) can be coercive if the recipient of the offer would strongly prefer to move from his actual pre-proposal situation to some alternative, feasible pre-proposal situation that the proposing party is preventing the recipient from being in (Zimmerman, 1981, p. 133). Meanwhile, Feinberg argues that

[20] For Wertheimer's responses to some salient objections raised against his moralized baseline account, see Wertheimer (1989), pp. 244–255.

offers are coercive if the proposer only gives the recipient a choice between different evils, and if the 'differential coercive pressure' of the offer (that is the difference in the comparative worth of the projected consequences of the recipient's refusing the offer, and the recipient's accepting the offer) is sufficiently high (Feinberg, 1989, p. 234).

There is room for debate on the question of whether the neurocorrective offer would qualify as a coercive offer on these accounts. For it to do so on Zimmerman's account, there would have to be a plausible sense in which the offender's most preferred pre-proposal situation (in which she is not incarcerated) satisfies Zimmerman's feasibility requirement. Similarly, on Feinberg's account, there might be scope to question whether the differential coercive pressure of the neurocorrective offer is sufficiently high for the offer to qualify as coercive.

Interestingly, Feinberg's account of coercive offers might be understood to suggest a second line of response to the objection from coercion. Whilst Feinberg claims that certain offers can be coercive (as explained above), he also suggests that not all coercive offers invalidate consent. In particular, he suggests that coercive offers in which the proposing party had no role in creating the recipient's pre-proposal circumstances of vulnerability very rarely invalidate consent (Feinberg, 1989, p. 246). He describes the following case as an example of a coercive offer that does not invalidate consent. Suppose that a woman has a daughter who is dying, but that she cannot afford to pay for surgery that would save her life. Suppose further that a millionaire, who is in no way responsible for the fact that the woman is poverty stricken, offers to pay for the child's surgery if the mother agrees to have sex with him. According to Feinberg, the offer made by this 'lecherous millionaire' is coercive;

however, since the millionaire had no role in creating the woman's pre-proposal circumstances of vulnerability, the mother could, on Feinberg's view, nonetheless validly consent to the offer.

Accordingly, a second response to the objection from coercion that defenders of the neuorcorrective offer could make is to argue that even if the offer is coercive, the offender's consent to undergo the neurocorrective can nevertheless be valid. An initial problem with responding to the coercion-based objection to the neurocorrective offer in this way is that in order for an offer to be coercive without invalidating consent on Feinberg's account, the proposing party must have played no role in creating the recipient's pre-proposal circumstances of vulnerability. This condition is satisfied in Feinberg's lecherous millionaire case: the millionaire did not contribute to the childhood medical condition in virtue of which the woman is vulnerable. By contrast, the neurocorrective offer does not meet this condition. It seems most natural to understand the state as the party that makes the offer in this case, and the state has a clear role in creating the offender's pre-proposal circumstances of vulnerability: the state is responsible for incarcerating the offender, and the offender is in a position of vulnerability due to being incarcerated.

Arguably though, one might circumvent this problem by weakening Feinberg's conditions on consent-preserving coercive offers, so that a coercive offer could qualify as preserving the validity of the recipient's consent if the proposing party did not violate the recipient's rights in creating the recipient's pre-proposal circumstances of vulnerability.

However, even if this is a plausible move to make, we might question whether Feinberg gives a convincing argument for the claim that some coercive offers do not

invalidate consent. Moreover, even if such offers do not invalidate consent, one might wonder whether the neurocorrective offer could be wrong (i) because coercive, but not because its coerciveness invalidates consent, or (ii) for reasons other than its coerciveness. For instance, Bomann-Larsen (2013) has argued that although the neurocorrective offer is not coercive, the state, at least in some circumstances, may not be in the right normative position to make the offer; thus, in making the offer in these circumstances, the state may wrong the offender by failing to recognize him as a moral equal, and by violating his fundamental claim to moral respect.[21] In a similar vein, McMillan has suggested, following Feinberg's approach to coercion, that coercive offers that preserve consent can nonetheless cause the recipient "moral harm" (McMillan, 2014b, p. 596). Finally, Shaw endorses Bomann-Larsen's criticism of the neurocorrective offer, but also objects to it on the consequence-based grounds that offering neurocorrectives to offenders would "send out the message that all offenders who are offered the intervention stand in need of it, whether they consent to it or not" (Shaw, 2014, p. 13).

Clearly then, establishing that the neurocorrective offer does not preclude valid consent will not settle the debate regarding the moral permissibility of making the offer. However, it should be acknowledged that whilst the wrongs associated with exploitation, inappropriate offers and other moral harms may be significant, they are not readily understood as wrongs involving affronts to autonomy. As such, the responses that we have so far considered may be understood as ways in which defenders of the neurocorrective offer might seek to move the debate away from autonomy-based objections. This is a significant move, since affronts to autonomy are

[21] However, as Ryberg and Petersen (2013) point out in response to Bomann-Larsen, it seems that further work needs to be done in cashing out what the state is in a normative position to offer offenders, and why certain offers involve a violation of a fundamental claim to moral respect..

often understood to be particularly egregious in contemporary bioethics, given the paramount value placed on autonomy in this context (Beauchamp and Childress, 2001; Caplan, 2006, p. 117; Gillon, 2003).

In any case, suppose, contrary to what we suggested above, that the neurocorrective offer is indeed coercive, and that it is therefore not possible for an offender to validly consent to it. In order to defend the moral permissibility of the neurocorrective offer on this understanding, one would then have to refute the consent requirement that we introduced in our discussion of the first slippery slope argument, according to which neurocorrectives can only permissibly be provided with the valid consent of the offender who will undergo the intervention. In the final section, we shall turn our attention to arguments that have been used in an attempt to refute the consent requirement. Of course, such arguments also lend support to the claim that the *mandatory* imposition of neurocorrectives could also be morally permissible

III Denying the Need for Valid Consent

One way to argue against the consent requirement is to claim that the moral reasons we have to respect the offender's autonomy can be outweighed by other moral reasons. Whilst the value of autonomy plays a central role in medical ethics, there are a number of other moral values that are relevant to the penal context under consideration, including, *inter alia*, public welfare, liberty, justice and fairness. Indeed, it seems that these values can give rise to moral reasons to employ non-consensual neurocorrectives.

To illustrate, it seems plausible to claim that the state has strong reasons to employ neurocorrectives that are grounded in a concern for public welfare. These interventions could be highly effective at facilitating rehabilitation and preventing

recidivism, and might also be less costly than alternative means of realising these goals, such as incarceration. As well as protecting public welfare, one might also point out that neurocorrectives might benefit offenders themselves, either by allowing them to avoid more harmful corrective interventions, or because their rehabilitation is itself beneficial (Day et al., 2004, pp. 260–263; Ryberg, 2012, p. 232).

Of course, these welfare-based arguments against the consent requirement would be met with strong opposition, in view of the importance attributed to the principle of respect for autonomy (as we briefly discussed above). There is some precedent within psychiatry and infectious disease control for the imposition of nonconsensual medical interventions on competent adults. However, at least in the UK, individuals have a legal right to refuse medical interventions (including neurointerventions) that precludes the lawful imposition of such interventions on public interest grounds.[22] This might reflect a commitment to the view that, at least in relation to medical interventions, moral reasons to respect individual autonomy are more powerful than reasons to protect the public interest or the interests of the offender.

In any case, it is, as Ryberg argues, too simplistic to set up the debate as one in which we simply weigh autonomy-related concerns against the welfare-based reasons to allow these interventions. Ryberg (2013) argues that our moral judgment regarding the permissibility of coerced treatments in criminal justice should also incorporate penal-theoretic considerations. For instance, he points out that retributivists are likely to be unconvinced by welfare-based arguments in favour of the use of neurocorrectives if such interventions result in a reduction in the severity of punishment, assuming that offenders currently receive proportionate punishment.

[22] *Health and Social Care Act 2008,Part 2A Orders, 45e.*

Having said this, if offenders currently receive punishment that is too severe, Ryberg suggests that there may yet be a retributivist argument in favour of using such interventions (Ryberg, 2013, pp. 10–11). These points are not intended to settle the debate about whether and when neurocorrectives may permissibly be used. Rather, the point is that the debate needs to be expanded to include penal-theoretic considerations.

Another factor that complicates the moral assessment of the consent requirement is that, somewhat paradoxically, one might also raise an autonomy-based *objection* to the requirement. One of the main reasons that valid consent is deemed to be of such moral importance is rooted in the principle of respect for autonomy. In view of this, one might argue that valid consent is not required for the imposition of neurocorrectives on criminal offenders if those interventions can be understood as enabling autonomy. There may well be scope for understanding certain interventions as having these autonomy-enhancing effects. For instance, Arthur Caplan has suggested that the mandatory treatment of drug addicts may be justified on these grounds, since such addicts "do not have the full capacity to be self-determining or autonomous because their addiction literally coerces their behaviour" (Caplan, 2006, p. 118). Douglas et al. (2013) have suggested that similar considerations might count in favour of offering sex offenders the option to undergo chemical castration regardless of whether they could validly consent to it, since the desires that motivate many sexual offences (and that are attenuated by chemical castration) would qualify as impediments to autonomy on most plausible theories of autonomy.

However, it is not clear that such considerations would be decisive; the claim that an individual's future autonomy can permissibly be enhanced even at the cost of disrespecting his present autonomy would be controversial. Moreover, it is unlikely

that such an autonomy-based case for rejecting the consent requirement could be applicable to all criminal offenders (Douglas et al., 2013, pp. 400–410). The argument seems plausible with respect to cases in which the offender feels alienated from the impulsive desires that motivated his criminal act. However, whilst most theories of autonomy would claim that an offender would lack autonomy with respect to those desires, many offenders may not feel alienated from their desires in this way, and may instead endorse them. Most prominent theories of autonomy would claim that such offenders would be autonomous with respect to those desires.

Alternatively, it might be possible to provide a paternalistic justification of non-consensual neurocorrectives, according to which it may be permissible to impose such an intervention by virtue of the fact that it is somehow in the offender's own interests. The form of paternalism here is most naturally construed as a form of moral paternalism, according to which it is permissible to carry out paternalistic interventions on other agents in order to ensure their 'moral well-being'. On this view, it might be argued that, by imposing a neurocorrective, we alter an offenders moral traits or dispositions, and the moral paternalist will hold that it can be noninstrumentally good to possess certain moral traits or dispositions.[23] Thus, the moral paternalist might claim that, in imposing the neurocorrective, we benefit the offender in terms of their moral well-being. Indeed, according to the paternalistic theory of punishment defended by Herbert Morris, the justification of punishment is precisely to benefit the offender in this sort of way (Morris, 1981).

[23] See Dworkin (2005, p. 308) for discussion.

Of course, the strength of the moral paternalist's argument here depends a great deal on the claim that a person's welfare can be increased by changes in their moral traits and dispositions. In fact, the case of non-consensual neurocorrectives illustrates an important challenge to this view. It is likely that at least some criminal offenders would not endorse the changes in their moral traits and dispositions that neurocorrectives would bring about; yet, in the absence of endorsing these changes, it is difficult to see how the changes can be good *for* the offender. It seems that the moral paternalist can respond to this challenge in one of two ways; first, they may just bite the bullet and appeal to an objectivist account of welfare according to which there are certain goods that can improve a person's life even if the person herself does not endorse having those goods. Alternatively, Dworkin suggests that we might appeal to a form of what he calls 'legal moralism', according to which it may be justifiable to try and make a person's life go morally better *per se*, even if this does not make that person's life go morally better *for her* in the way that moral paternalism seems to require (Dworkin, 2005, p. 319).

We have outlined some public welfare-based, autonomy-based, and paternalistic arguments in favour of imposing non-consensual neurocorrectives, and suggested that they all face significant challenges. Rather than objecting to the consent requirement by appealing to such arguments, one might alternatively argue that consistency demands that we permit the non-consensual use of neurocorrectives in criminal justice. It is generally believed that the state may permissibly do things to criminal offenders without their consent that it could not permissibly do to others

without (and in some cases with) consent (Douglas, 2014, p. 105). The challenge that this observation raises for those who defend the consent requirement is to give an account of the moral difference between incarcerating an offender without their consent and imposing a neurocorrective without their consent that can explain why the former is morally permissible, but the latter morally impermissible.

One obvious difference is that neurocorrectives are likely to be physically invasive and thus may be said to violate the offender's right to bodily integrity in a way that incarceration does not. However, incarceration plausibly violates other important rights such as rights to free movement and association, or at least it would had the offender not forfeited these rights, or some aspect of them; as such, in order to defend the consent requirement by appealing to the right to bodily integrity, it seems that one would need to defend the claim that this right is more robust than rights to freedom of movement and association, such that it would take more serious criminal offending to forfeit the relevant rights to bodily integrity than to forfeit the relevant rights to free movement and association. One of us has recently argued that this claim is not as easy to defend as our intuitions might lead us to believe (Douglas, 2014, pp. 113–118).

Moreover, it might seem implausible that it is the bodily invasiveness of neurocorrectives that explains why, unlike incarceration, they may not permissibly be imposed without consent. The primary intended effects of neurocorrectives are *mental* rather than bodily. Indeed, some neurointerventions that might be used as neurocorrectives, such as transcranial electric brain stimulation, are only doubtfully or minimally physically invasive. It might thus seem more credible that the most significant moral difference between incarceration and neurocorrectives is to be located in the *mental* effects of the latter. Perhaps, then, the consistency argument

should be resisted not by appealing to a right to bodily integrity, but by appealing to a right to *mental* integrity (Douglas, 2014, pp. 113–118).

Farah seems to implicitly appeal to this sort of concept in observing that neurocorrectives might deny offenders the ". . . freedom to think one's own thoughts and have one's own personality" (Farah, 2002, p. 1126). However, there has been little research regarding whether there is a right to mental integrity, or concerning what sorts of mental influences might violate it. Bublitz and Merkel have offered an initial exploration of this putative right, suggesting that such a right exists, or is, "as a tacit assumption, woven into law's structure" (Bublitz and Merkel, 2014, p. 60); they also argue that nonconsensual interventions into other minds might violate this right if they substantially undermine mental self-determination by reducing the victim's mental control, or by exploiting pre-existing mental weaknesses (Bublitz and Merkel, 2014, p. 68). Whilst Bublitz and Merkel understand the right to mental integrity to afford protection against harms to mental life generally, for our purposes, the most interesting aspect of the purported right to mental integrity is that it is understood to afford protection against interventions that undermine an agent's *mental self-determination* (Bublitz and Merkel, 2014, p. 58). Mental self-determination here is to be understood in an interpersonal sense, to refer to the extent to which a person's mental powers are free from the undue influence of others.

One problem facing advocates of a right to mental integrity is how we should cash out the notion of 'undue influence' here. After all, we commonly interfere with other people's minds without this being a morally problematic instance of undermining their mental self-determination; for instance, when we enter into a

rational dialogue with another person, we will often have an effect on their mental states. How then, do we distinguish those sorts of interferences that undermine mental self-determination and those that do not? Anticipating this problem, Bublitz and Merkel claim that there is an important moral difference between interventions that operate directly on the brain (such as neuro-interventions of the sort that we are considering), and indirect interventions that are perceived sensually and processed by psychological mechanisms. The two types of intervention differ morally, they claim, because indirect interventions may respect mental self-determination, whilst direct interventions circumvent it. They write:

> Persons have most control over interventions whose sensual substrates they perceive, particularly those rising to the level of conscious awareness. . . Direct interventions, by contrast, are qualitatively different, presumably bypassing these psychological (not necessarily rational) processes altogether. Roughly one could say that *indirect interventions are inputs into the cognitive machinery our minds are adapted to process, whereas direct interventions change the cognitive machinery itself*. (Bublitz and Merkel, 2014, pp. 69"cit.

Shaw (2014) has offered an alternative account of the morally relevant difference between direct and indirect interventions. She suggests that certain direct interventions are problematic because they would serve to objectify offenders by rehabilitating them in a manner that fails to engage them in a rational moral dialogue.[24] She also suggests that permitting the use of neurointerventions in criminal

[24] Notice that in this paper, Shaw objects to Bublitz and Merkel's earlier work in which they endorse the claim that the moral difference between direct brain interventions and standard environmental influences is that the former involve a manipulator who takes over responsibility for the results of the intervention,

justice would serve to widen the inequality of power between offenders and the state (Shaw, 2014, p. 13).[25] However, unlike Bublitz and Merkel, Shaw's account of why direct interventions are more morally problematic than indirect interventions does not rule out the possibility that some direct interventions might yet be morally permissible. Shaw claims that whilst direct interventions are problematic if they serve to change the offender's values in a manner that does not include the offender in a moral dialogue, nonconsensual direct interventions may be permissible if they enable offenders to engage in a moral dialogue as a part of their rehabilitation. For instance, she suggests that permissible interventions might include enhancing an offender's powers of attention, or their ability to delay gratification (Shaw, 2014, pp. 14–15).

It can be argued, however, that the question of whether the state may consistently permit the nonconsensual incarceration of offenders without also permitting the nonconsensual use of neurocorrectives remains open. Whilst the analyses considered above have laid important foundations for addressing this question, there is still a great deal of work to be done in this area; for instance, with respect to Bublitz and Merkel's analysis, one could question whether the fact that an indirect intervention engages the recipient's psychological processes is always sufficient for establishing that the intervention respects the recipient's right to mental self-determination. Furthermore, we might also question whether some direct interventions might be permissible even if they bypass the offenders psychological processes, as Shaw suggests with regards to interventions that serve to enable the offender to participate in rational moral dialogue. Of course, there is an important

---

whilst the latter type of influence do not. See Bublitz and Merkel (2009) and Shaw (2014, pp. 8–12) for discussion.

[25] Shaw (2014, p. 13) goes so far as to claim that the use of such interventions would not be permissible even if the offender requested them, on the grounds that doing so would affect society's stance towards offenders as a group.

penal theoretic issue underlying both of these questions, namely, the issue of whether how important it is for criminal sanctions to respect the offender's right to mental integrity, and whether criminal sanctions should seek to address the offender 'in a rational moral dialogue'.[26] Whilst much of the debate regarding the use of non-consensual neurocorrectives has been grounded in the framework of medical ethics, it seems that addressing these penal theoretic questions would add a great deal to the discussion of the proper role of these interventions in criminal justice.

References

Anderson, S., 2011. Coercion, in: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy.

Beauchamp, T.L., Childress, J.F., 2001. Principles of biomedical ethics, 5th ed.. ed. Oxford University Press, Oxford.

Bentham, J., 1890. Theory of legislation [electronic resource], 6th ed. ed. Trübner, London.

Bomann-Larsen, L., 2013. Voluntary Rehabilitation? On Neurotechnological Behavioural Treatment, Valid Consent and (In)appropriate Offers. Neuroethics 6, 65–77. doi:10.1007/s12152-011-9105-9

Bond, A.J., 2005. Antidepressant treatments and human aggression. Eur. J. Pharmacol. 526, 218–225. doi:10.1016/j.ejphar.2005.09.033

Bublitz, J.C., Merkel, R., 2014. Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination. Crim. Law Philos. 8, 51–77. doi:10.1007/s11572-012-9172-y

Bublitz, J.C., Merkel, R., 2009. Autonomy and authenticity of enhanced personality traits. Bioethics 23, 360–374. doi:10.1111/j.1467-8519.2009.01725.x

California Penal Code, 2012.

Caplan, A.L., 2006. Ethical issues surrounding forced, mandated, or coerced treatment. J. Subst. Abuse Treat. 31, 117–120. doi:10.1016/j.jsat.2006.06.009

Carlen, 1989. Crime, Inequality and Sentencing, in: Carlen, Cook (Eds.), Paying for Crime. Open University Press.

[26] For defences of communicative theories of punishment, see Hirsch (1996) and Duff (2003)

Christman, J., 1991. Autonomy and Personal History. Can. J. Philos. 21. doi:10.1080/00455091.1991.10717234
Cohen, G.A., 1977. Robert Nozick and Wilt Chamberlain: How patterns preserve liberty. Erkenntnis 11, 5–23. doi:10.1007/BF00169842
Crockett, M.J., Apergis-Schoute, A., Herrmann, B., Lieberman, M.D., Müller, U., Robbins, T.W., Clark, L., 2013. Serotonin Modulates Striatal Responses to Fairness and Retaliation in Humans. J. Neurosci. 33, 3505–3513. doi:10.1523/JNEUROSCI.2761-12.2013
Crockett, M.J., Clark, L., Hauser, M.D., Robbins, T.W., 2010. Serotonin selectively influences moral judgment and behavior through effects on harm aversion. Proc. Natl. Acad. Sci. 107, 17433–17438. doi:10.1073/pnas.1009396107
Crockett, M.J., Clark, L., Tabibnia, G., Lieberman, M.D., Robbins, T.W., 2008. Serotonin Modulates Behavioral Reactions to Unfairness. Science 320, 1739–1739. doi:10.1126/science.1155577
Cullen, F.T., 1982. Reaffirming rehabilitation. Anderson PubCo, Cincinnati, Ohio.
Day, A., Tucker, K., Howells, K., 2004. Coerced offender rehabilitation – a defensible practice? Psychol. Crime Law 10, 259–269. doi:10.1080/10683160410001662753
Del Busto, E., Harlow, M.C., 2011. American Sexual Offender Castration Treatment and Legislation, in: Boer, D.P., Eher, R., Craig, L.E., Miner, M.H., Pfäfflin, F. (Eds.), International Perspectives on the Assessment and Treatment of Sexual Offenders: Theory, Practice and Research. Oxford: Wiley-Blackwell, pp. 543–571.
Donovan, S.J., Stewart, J.W., Nunes, E.V., Quitkin, F.M., Parides, M., Daniel, W., Susser, E., Klein, D.F., 2000. Divalproex treatment for youth with explosive temper and mood lability: a double-blind, placebo-controlled crossover design. Am. J. Psychiatry 157, 818–820.
Douglas, T., 2014. Criminal Rehabilitation Through Medical Intervention: Moral Liability and the Right to Bodily Integrity. J. Ethics 18, 101–122. doi:10.1007/s10892-014-9161-6
Douglas, T., 2010. Intertemporal Disagreement and Empirical Slippery Slope Arguments. Utilitas 22, 184–197. doi:10.1017/S0953820810000087
Douglas, T., Bonte, P., Focquaert, F., Devolder, K., Sterckx, S., 2013. Coercion, Incarceration, and Chemical Castration: An Argument From Autonomy. J. Bioethical Inq. 10, 393–405. doi:10.1007/s11673-013-9465-4
Duff, R.A., 2003. Punishment, Communication, and Community. OUP USA Studies in Crime and Public Policy.
Dworkin, G., 2005. Moral Paternalism. Law Philos. 24, 305–319.
Dworkin, G., 1988. The theory and practice of autonomy. Cambridge University Press, Cambridge.
Ekstrom, L.W., 1993. A Coherence Theory of Autonomy. Philos. Phenomenol. Res. 53, 599–616.
Eley, S., Gallop, K., McIvor, G., Morgan, K., Yates, R., 2002. Drug treatment and testing orders: evaluation of the Scottish pilots. Scottish Executive Central Research Unit.
European Committee for the Prevention of Torture and Inhuman, or Degrading Treatment or Punishment, 2009. Report to the Czech Government on the visit to the Czech Republic carried out by the European Committee for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment (CPT).

European Committee for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment (CPT), n.d. Report to the German Government on the visit to Germany carried out by the European Committee for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment (CPT) from 25 November to 7 December 2010 [WWW Document]. URL (accessed 10.24.13).

Faden, R.R., Beauchamp, T.L., 1986. A history and theory of informed consent. Oxford University Press, New York.

Farah, M.J., 2002. Emerging ethical issues in neuroscience. Nat. Neurosci. 5, 1123–1129. doi:10.1038/nn1102-1123

Feinberg, J., 1989. The Moral Limits of the Criminal Law Volume 3: Harm to Self [electronic resource]. Oxford University Press, New York.

Florida Statutes Annotated, 2013.

Gillon, R., 2003. Ethics needs principles--four can encompass the rest--and respect for autonomy should be "first among equals." J. Med. Ethics 29, 307–312. doi:10.1136/jme.29.5.307

Greely, H.T., 2008. Neuroscience and Criminal Justice: Not Responsibility but Treatment. Univ. Kans. Law Rev. 56, 1103–1138.

Green, W., 1986. Depo-Provera, Castration, and the Probation of Rape Offenders: Statutory and Constitutional Issues. Univ. Dayt. Law Rev. 12, 1.

Hampton, J., 1984. The Moral Education Theory of Punishment. Philos. Public Aff. 13, 208–238.

Hirsch, A. von, 1996. Censure and Sanctions. Oxford University Press.

Hough, M., Clancy, A., McSweeney, T., Turnbull, P., 2003. Impact of Drug Treatment and Testing Orders on Offending: Two-Year Reconviction Results.

Idaho Court of Appeals, 1991. State v. Estes,.

Kant, I., 2004. Critique of practical reason. Dover, Mineola, NY.

Khanzode, L.A., Saxena, K., Kraemer, H., Chang, K., Steiner, H., 2006. Efficacy profiles of psychopharmacology: divalproex sodium in conduct disorder. Child Psychiatry Hum. Dev. 37, 55–64. doi:10.1007/s10578-006-0019-4

Lewis, C.S., 1953. Humanitarian Theory of Punishment, The. Res Judicatae 6, 224.

Louisiana Revised Statutes, 2013.

Lu, L., Wang, X., Kosten, T.R., 2009. Stereotactic neurosurgical treatment of drug addiction. Am. J. Drug Alcohol Abuse 35, 391–393. doi:10.3109/00952990903312478

McMillan, J., 2014a. The kindest cut? Surgical castration, sex offenders and coercive offers. J. Med. Ethics 40, 583–590. doi:10.1136/medethics-2012-101030

McMillan, J., 2014b. Surgical castration, coercive offers and coercive effects: it is still not about consent. J. Med. Ethics 40, 596–596. doi:10.1136/medethics-2013-101507

Mill, J.S., 1991. On Liberty, in: Gray, J., Smith, G.W. (Eds.), J.S. Mill, On Liberty, in Focus. Routledge.

Morris, H., 1981. A Paternalistic Theory of Punishment. Am. Philos. Q. 18, 263–271.

Murphy, J.G., 1973. Marxism and Retribution. Philos. Public Aff. 2, 217–243.

Nevels, R.M., Dehon, E.E., Alexander, K., Gontkovsky, S.T., 2010. Psychopharmacology of aggression in children and adolescents with primary neuropsychiatric disorders: a review of current and potentially promising treatment options. Exp. Clin. Psychopharmacol. 18, 184–201. doi:10.1037/a0018059

Nozick, R., 1969. Coercion, in: Morgenbesser, S., Suppes, P., White, M. (Eds.), Philosophy, Science, and Method: Essays in Honor of Ernest Nagel. New York: St. Martin's Press.
Rosati, C.S., 1994. Study of Internal Punishment, A. Wis. Law Rev. 1994, 123.
Ryberg, J., 2013. Is Coercive Treatment of Offenders Morally Acceptable? On the Deficiency of the Debate. Crim. Law Philos. 1–13.
Ryberg, J., 2012. Punishment, Pharmacological Treatment, and Early Release: Int. J. Appl. Philos. 26, 231–244. doi:10.5840/ijap201226217
Ryberg, J., Petersen, T.S., 2013. Neurotechnological Behavioural Treatment of Criminal Offenders—A Comment on Bomann-Larsen. Neuroethics 6, 79–83. doi:10.1007/s12152-011-9146-0
Scott, C.L., Holmberg, T., 2003. Castration of sex offenders: prisoners' rights versus public safety. J. Am. Acad. Psychiatry Law 31, 502–509.
Shaw, E., 2014. Direct Brain Interventions and Responsibility Enhancement. Crim. Law Philos. 8, 1–20. doi:10.1007/s11572-012-9152-2
Sitaram, R., Caria, A., Birbaumer, N., 2009. Hemodynamic brain-computer interfaces for communication and rehabilitation. Neural Netw. Off. J. Int. Neural Netw. Soc. 22, 1320–1328. doi:10.1016/j.neunet.2009.05.009
Stevens, R., 1988. Coercive offers. Australas. J. Philos. 66, 83–95. doi:10.1080/00048408812350261
Tancredi, L.R., Weisstub, D.N., 1986. Technology assessment: its role in forensic psychiatry and the case of chemical castration. Int. J. Law Psychiatry 8, 257–271.
Taylor, J.S. (Ed.), 2005. Personal autonomy : new essays on personal autonomy and its role in contemporary moral philosophy. Cambridge University Press, Cambridge.
Vanderzyl, K., 1994. Castration as an Alternative to Incarceration: An Impotent Approach to the Punishment of Sex Offenders. North. Ill. Univ. Law Rev. 15.
Vincent, N.A., 2014. Restoring Responsibility: Promoting Justice, Therapy and Reform Through Direct Brain Interventions. Crim. Law Philos. 8, 21–42. doi:10.1007/s11572-012-9156-y
Wertheimer, A., 2014. Coercion, 2nd Ed. ed. Princeton University Press, Princeton.
Wertheimer, A., Miller, F.G., 2014. There are (STILL) no coercive offers. J. Med. Ethics 40, 592–593. doi:10.1136/medethics-2013-101510
Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R., 2010. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proc. Natl. Acad. Sci. U. S. A. 107, 6753–6758. doi:10.1073/pnas.0914826107
Zimmerman, D., 1981. Coercive Wage Offers. Philos. Public Aff. 10, 121–145.