

Justifications for Non-Consensual Medical Intervention: From Infectious Disease Control to Criminal Rehabilitation

Jonathan Pugh, Thomas Douglas

[This is a pre-publication version. The final article is forthcoming in *Criminal Justice Ethics*.]

Abstract: A central tenet of medical ethics holds that it is permissible to perform a medical intervention on a competent individual only if that individual has given informed consent to the intervention. Yet, it occasionally seems morally permissible to carry out non-consensual medical interventions on competent individuals for the purpose of infectious disease control (IDC). We describe two different moral frameworks that have been invoked in support of non-consensual IDC interventions, and identify five desiderata that might be used to guide the assessments of the moral permissibility of such interventions on either kind of fundamental justification. We then consider what these desiderata imply for the justifiability of carrying out non-consensual medical interventions that are designed to facilitate rehabilitation amongst serious criminal offenders. We argue that this analysis suggests that a plausible case can be made in favour of such interventions.

A central tenet of medical ethics holds that it is permissible to perform a medical intervention on a competent individual only if that individual has given informed consent to the intervention. However, in some circumstances it is tempting to say that the moral requirement to obtain informed consent prior to administering a medical intervention is outweighed. For example, if an individual's refusal to undergo a medical intervention would lead to the transmission of a dangerous infectious disease to other members of the community, we might claim that it would be morally permissible to administer the intervention, even in the absence of consent. Indeed, as we shall discuss below, there are a number of examples of public health authorities implementing compulsory or coercive measures for the purposes of infectious disease control (IDC).

The plausibility of the thought that non-consensual medical interventions might be justified when performed for the purpose of infectious disease control (IDC) raises the question of whether such interventions might permissibly be used to realise other public goods. In this paper we focus on one possibility: whether it could be permissible to non-consensually impose certain interventions that alter brain states or processes through chemical or physical means on serious criminal offenders. We shall suggest that some such interventions might be permissible if they safely and effectively served to facilitate the offender's rehabilitation, and thereby prevent criminal recidivism.ⁱ We refer to brain interventions intended to serve these objectives as 'neurocorrectives'.

Authorities in certain jurisdictions have already begun to explore the use of neurocorrectives. For instance, drug-addicted offenders are sometimes compelled to take medications that are intended to attenuate their addictive desiresⁱⁱ and sex-offenders in some jurisdictions may be compelled to take testosterone-lowering drugs as a part of their criminal sentence or as a condition of their parole.ⁱⁱⁱ Moreover, as our understanding of the neural correlates of violent behavior increases, it seems plausible to claim that we might develop effective neurocorrectives that involve the use of other pharmaceuticals^{iv} or novel technologies such as deep brain stimulation,^v transcranial magnetic stimulation,^{vi} and neuro-feedback.^{vii}

There has been increasing ethical debate over the deployment of neurocorrectives.^{viii} Much of this debate has focused on their safety and effectiveness. We shall not examine these issues here. We simply assume, for the sake of argument, that the interventions we consider are both effective and have negligible side-effects.^{ix} Other critics have questioned whether an offender could validly consent to undergoing such an intervention, if doing so were a condition of their early release from prison^x. In this paper, we side-step this contentious issue by assuming that the hypothetical recipients of neurocorrectives have *not* validly consented to

undergoing the intervention. Instead, we shall argue that the considerations standardly offered in support of non-consensual interventions in IDC also offer considerable support to the permissibility of some safe and effective non-consensual neurocorrectives.^{xi}

First, however, some definitions. We shall adopt a definition of ‘medical interventions’ that is perhaps broader than colloquial use, since we shall understand it to refer not only to typical medical procedures such as pharmacotherapies and surgical procedures—measures that involve some degree of bodily invasion—but also measures that restrict an individual’s freedom of movement and association, insofar as these are used for medical reasons. For instance, we shall understand quarantine to qualify as a ‘medical intervention’. We will use the term ‘IDC interventions’ to refer to medical interventions performed for the purposes of infectious disease control.

We shall use the term ‘non-consensual intervention’ to refer to any intervention performed without the valid consent of the recipient of the intervention. It is possible to distinguish a number of different kinds of non-consensual intervention. For instance, one might draw a distinction between compelled and coerced interventions. Following Feinberg, we may say that “. . . an option is closed by compulsion when one alternative has been made impossible”.^{xii} For instance, an individual may be compelled to undergo an intervention by being subjected to direct physical force. A compelled intervention is carried out without the recipient’s consent, and perhaps even despite his dissent. In contrast to compulsion, coercion does not make an alternative impossible, but rather destroys its appeal by increasing its cost.^{xiii} In contrast to the case of compulsion, in coercion, the individual does consent to the intervention, but the coercive pressure to which he has been subjected may *invalidate* that consent. We use the term ‘non-consensual’ intervention to refer both to interventions performed without consent, and to interventions performed with invalid consent.

I Non-Consensual Interventions in IDC

The putative requirement to obtain consent before performing a medical intervention is standardly thought to derive from a reason to respect the prospective patient's rights, which may include rights to personal autonomy and rights to freedom from bodily interference. Performing a medical intervention non-consensually is thought to be impermissible because it violates one or more of these rights.^{xiv} In the IDC cases of interest to us, this reason to respect the patient's rights plausibly remains in place (it has not been forfeited). However, there is also a powerful reason to impose the medical intervention, since it will help to forestall the spread of an infectious disease. In a number of historical cases, this reason has been thought by many to either outweigh reasons to respect the individual's rights, or to deprive these rights of their normal protective force, with the result that the intervention may permissibly be imposed non-consensually.

One type of medical intervention that has often been used for the purposes of IDC is vaccination. Vaccinations are amongst the most effective IDC interventions that societies can employ. For instance, broad vaccination coverage is largely responsible for the widespread immunity to measles and tetanus in modern society, and it led to the eradication of smallpox in the late twentieth century.^{xv} Barring adverse side-effects, which are uncommon for most vaccines, and cases where an individual is insusceptible to the negative effects of an infectious disease, being vaccinated is typically beneficial to the vaccinated individual, since an effective vaccine will normally confer immunity to a particular infectious disease. However, effective individual immunization also confers benefits to third parties; if a

sufficient proportion of the population is immunized against a disease then this can disrupt the spread of the disease from person to person. In such a scenario of ‘herd immunity’, the fact that a sufficient majority of the population is immune to a particular disease confers a degree of protection even to those who have not been vaccinated.^{xvi}

In light of this, some countries employ laws that could be described as introducing coercive pressure on individuals to undergo vaccination, insofar as they permit the imposition of costs for vaccine refusal.^{xvii} For instance, in the majority of states in the USA, children must normally complete a vaccination schedule as a condition of entering public school.^{xviii} In some cases, it might be claimed that if the operative reason undergirding an individual’s choice to undergo the vaccination is that they want to avoid these state-imposed costs of non-compliance (rather than, say, because they want the health benefits of being vaccinated), then that individual may qualify as being coerced in a manner that invalidates their consent.^{xix}

Vaccinations aim to prevent the spread of infectious disease by preventing individuals from becoming infected in the first place. However, states also impose other sorts of non-consensual medical intervention for IDC purposes. For example, in some jurisdictions it is legally permissible to compel the medical examination (and in rare cases even treatment) of persons suspected of carrying an infectious disease.^{xx xxi} The non-consensual examination and treatment of individuals carrying a dangerous infectious disease can aid IDC in two ways. First, effectively treating that individual will mean that that they will no longer be able to transmit the disease to others. Second, it can help to restrict the development of drug resistant pathogens, since one way in which a pathogen can develop drug resistance is by genetically evolving as a result of an incomplete treatment.^{xxii}

Although exerting coercive pressure to undergo vaccinations and even compelling medical examinations and treatments is permitted in some jurisdictions, these are somewhat controversial non-consensual IDC interventions. The use of non-consensual quarantine and

isolation however is often deemed to be less controversial (for reasons that we shall explore below). Both quarantine and isolation share the aim of separating certain individuals from the rest of the community in order to prevent or limit the transmission of infectious pathogens. The difference between the two lies in the diagnostic status of the detainee; in isolation, it is known that the detainee is infected, whereas quarantine involves the detention of individuals who have been (or are likely to have been) exposed to the infectious agent and are thus suspected of being (but not *known* to be) infected. Quarantine and isolation have long been used in response to outbreaks of infectious disease, particularly those for which no effective medical treatment is available.^{xxiii} For instance, quarantine was widely used in response to the recent outbreak of Ebola in West Africa.

II Approaches To The Moral Justification of Non-Consensual Medical Interventions

In this section we shall delineate two standard moral justifications for non-consensual medical interventions in IDC: the constrained consequentialist justification, and the self-defence justification. We then identify some desiderata for assessing IDC interventions that can be regarded as common ground between these two justifications, before, in the subsequent section, arguing that these desiderata support the justifiability of at least some non-consensual neurocorrectives.

It might be argued that non-consensual interventions in IDC can be morally justified by appealing to classic utilitarian reasoning. On such a view, these interventions would be permissible if they are predicted to bring about at least as much aggregate happiness as any

alternative course of action. However, few have been persuaded by such justifications. As T. M. Wilkinson points out, this sort of utilitarian thinking runs contrary to the belief, commonly held in biomedical ethics, that people have certain rights that “constrain the pursuit of the greater good”.^{xxiv} Many theorists reject the simple utilitarian approach because of its failure to accommodate this belief, amongst others; indeed, some may claim that this failure gives us reason to be sceptical of making any consequentialist considerations central to the justification of either medical or criminal justice interventions, the two types of intervention with which we are concerned in this paper. We shall now set out two alternative justifications to which they have turned, which are both compatible with the belief that rights should constrain the pursuit of the good, but which may nonetheless have been used to justify certain non-consensual interventions in IDC.

The first justification retains a broadly consequentialist approach but incorporates a richer account of the good than the utilitarian’s appeal to happiness, and sets certain side-constraints on the pursuit of this good. We shall call this the ‘constrained consequentialist approach’. Lawrence Gostin offers an example of it in his influential *Public Health Law: Duty, Power and Restraint*. In introducing his account, Gostin writes:

The prime objective of public health law is to pursue the highest possible level of physical and mental health in the population, consistent with the values of social justice.

He later expands on this claim by suggesting that the twin aims that motivate public health interventions are “to advance human well-being by improving health, and to do so particularly by focusing on the needs of the most disadvantaged”.^{xxv}

These passages suggest that Gostin deviates from utilitarianism in two respects. Though, like the utilitarian, he believes that interventions must promote individual well-being, his appeals to social justice and the needs of the most disadvantaged suggest that, unlike the utilitarian, he does not regard aggregate utility as the sole determinant of the good but instead believes that the *distribution* of well-being across individuals should be incorporated into the theory of the good. Moreover, the claim that the objective of public health law is to pursue high levels of health in a manner that is *consistent* with the values of social justice leaves open the possibility that the pursuit of good outcomes may be side-constrained—that is to say, there may be certain means to pursuing those outcomes that are ruled out, regardless how good their effects would be. The nature of the constraints will depend on the theory of social justice that one adopts in cashing out the framework. Notably in claiming that public health ought to focus on the ‘needs of the most disadvantaged’, Gostin himself seems to implicitly endorse a fundamentally prioritarian account of social justice. Prioritarians might set strict constraints on the pursuit of the good by claiming that our pursuit of general well-being should be constrained by the principle that we must not select any option in which the worst off individuals will be worse off than the worst off individuals in any alternative distribution.

A second approach to the justification of non-consensual IDC interventions treats the imposition of such interventions as an instance of justified self-defence. We may say that an agent is acting in self-defence if she is acting to prevent another (the ‘attacker’) from harming her.^{xxvi} In the context of IDC, it has been suggested that it may be legitimate to think of an infected (or potentially infected) individual as an ‘attacker’, and other members of the public as potential victims who may be justified in defending themselves against the attacker, by, for example, imposing an IDC intervention.^{xxvii}

In what follows, we will not seek to adjudicate between constrained consequentialist and self-defence justifications for IDC interventions. Rather, we will deploy five desiderata that are commonly used to guide decisions regarding the deployment of non-consensual IDC interventions, and that are compatible with either kind of fundamental justification. We take these desiderata to collectively constitute the standard approach to decision-making regarding IDC interventions. We also take them to constitute *indicators* rather than *determinants* of the moral permissibility of public health interventions. These desiderata do not determine what is the right and wrong in IDC; rather, they constitute a practically applicable set of heuristics that agents operationalising either of these fundamental justifications may apply in order to guide their assessments of permissibility. Furthermore, it should be noted that, though the desiderata we shall outline are consistent with both constrained consequentialist and self-defence justifications, their interpretation may depend on which of these fundamental justifications one accepts. As we wish to remain neutral between these fundamental justifications, we will, where necessary, leave open whether these desiderata should be interpreted in line with a constrained consequentialist or a self-defence approach.

III Five Desiderata To Guide The Moral Assessment of Non-Consensual Medical Interventions

The first desideratum is the gravity—by which we mean the moral weight— of the harm that the intervention aims to prevent. This may depend, *inter alia*, on the magnitude of the harm, its qualitative importance, the likelihood or certainty with which it will occur, and its impact on distributive fairness.^{xxviii} The graver the harm, the easier it will be to justify the intervention.

The second desideratum is the effectiveness of the intervention in preventing the harm. In IDC, this will depend on the nature of the infectious disease, the nature of the intervention, and the context in which the intervention is to be deployed. Consider, for instance, the attempt to limit the spread of infectious disease by imposing mass quarantine; this will only be effective in preventing the spread of an infectious agent if it is possible to identify people likely to be incubating the infection, and if those people comply with the conditions of quarantine. Furthermore, the infectious agent must be transmissible in its pre-symptomatic or early symptomatic stages for quarantine to be useful in preventing further transmission.^{xxix xxx}

A third desideratum is that they have low opportunity costs. The expenditure of economic resources on a public health intervention necessarily diverts funds from other potentially beneficial public projects, including projects that might themselves promote public health.^{xxx} The opportunity cost of an intervention is the value of the alternative use to which the resources invested in the intervention would otherwise have been put. The higher the opportunity cost, the more difficult it will be to justify the intervention. In many cases, determining the opportunity cost of an intervention is difficult because it is unclear how the resources would otherwise be spent. However, even in those cases it may be possible to make a (necessarily uncertain) estimate of the expected opportunity costs. Moreover, in other cases, opportunity costs are quite clear: consider the position of a committee that has been given a fixed budget to spend on fighting a new pandemic, and is assessing competing proposals for the use of those funds.

The fourth and fifth desiderata are less straightforward so will be described in greater detail. The fourth desideratum is that the IDC intervention is the ‘least restrictive’ of the available alternatives for preventing the harm. As the discussion in section I makes clear,

many effective IDC interventions can impose significant burdens on the individuals that they target. In overriding a competent individual's refusal to consent, we arguably infringe his right to autonomy, and we may, depending on the intervention, also infringe other moral rights such as rights to bodily integrity, and freedom of movement and association. In view of this, it has been suggested that public health authorities should realise their IDC goals through the use of 'least restrictive alternative' of the available options.^{xxxii xxxiii}

The definition of 'restrictiveness' here has important ramifications for the moral foundations of a 'least restrictive alternative' desideratum. It might be claimed that the restrictiveness of an intervention depends on the extent to which it infringes upon the recipient's moral rights.^{xxxiv} Notably, the understanding of rights that such an account might invoke need not assume that rights must trump all other considerations; an intervention may be permissible (on broadly consequentialist grounds or grounds of self-defence) even if it infringes another's rights, as long as it is the intervention that infringes rights to the *least extent* of all available (effective) interventions. Of course, how the 'extent' of rights infringements is to be determined is open to debate.

On this interpretation, the least restrictive alternative desideratum may be regarded as reflecting a deontological constraint on either the pursuit of the good (on constrained consequentialism), or self-protection (on a self-defence-based justification). In either case, it is an interpretation that is most naturally understood in deontological terms. However, some may be suspicious at the invocation of 'rights' that this interpretation makes. An alternative interpretation of the least restrictive alternative desideratum might instead claim that the restrictiveness of the intervention depends on the extent to which the intervention *harms* the subject, for instance by frustrating certain interests they may have. On this reading, the least restrictive alternative desideratum may be understood to be highlighting the harms that non-

consensual interventions can cause to the recipients of the intervention. On either a constrained consequentialist or self-defence justification, these harms inflicted by the intervention weigh against the harms that it averts (captured by the first desideratum) in determining its justifiability.

This brings us to the fifth desideratum, which is that the intervention must be proportionate to the threat that the recipient of the intervention poses to others. This desideratum is most strongly associated with self-defence justifications, where the desideratum (understood as a guide to practical decision-making) can be seen as a direct operationalisation of a proportionality requirement at the level of fundamental justification; self-defence accounts make the permissibility of protective measures dependant on the existence of proportionality. We will thus begin by briefly outlining how proportionality has been understood within self-defence justifications, before suggesting that a proportionality desideratum for non-consensual medical interventions can be accepted by proponents of a constrained consequentialist approach too, even though this approach includes no proportionality requirement at the level of fundamental justification.

Jeff McMahan succinctly captures the requirement of proportionality in self-defence in the following way:

The requirement of proportionality holds, roughly, that the harm inflicted in self-defence must not be excessive in relation to the threatened harm one seeks to avoid.^{xxxv}

In order to establish whether a harm inflicted in self-defence is excessive, one must first assess the relative gravity of the harm inflicted by an act of self-defence, and the harm that the attacker can be expected to inflict in the absence of self-defence. On the ‘equivalent harm view’ of proportionality, self-defence is proportionate only if the “harm that the force is intended to fend off is at least equivalent to the harm inflicted on the attacker”.^{xxxvi} On this

view, inflicting fatal force in self-defence is only proportionate when doing so is intended to fend off a harm that is itself as grave as death.

However, other views take further factors into account. For instance, it might be argued that assessments of proportionality should acknowledge that the attacker may be liable to suffer defensive force by virtue of his culpability (or, on some views, his mere responsibility)^{xxxvii} for the threat he poses, whilst the victim is not liable to suffer the threatened harm.^{xxxviii} On McMahan's justice-based account, this asymmetry makes it permissible, as a matter of justice, to ensure that the threatening party, rather than the potential victim, suffers any harm necessitated by the threat posed.^{xxxix} Moreover, on this view, the stringency of the proportionality restriction on the use of defensive force may vary in accordance with the threatening party's degree of culpability (or moral responsibility) for the threat that he poses.^{xl} For our purposes here, the important implication of justice-based views of proportionality is that they allow for the possibility that it can be justifiable for an agent acting in self-defence to harm a morally responsible attacker *more* than the attacker would otherwise harm her.

In contrast to self-defence justifications, the idea of proportionality is arguably not fundamental to most constrained consequentialist justifications of IDC interventions. However, assessing whether an intervention can be justified on the constrained consequentialist approach requires assessing the net harmfulness of the intervention, where the most significant harms are likely to be harms to others that the intervention can be expected to prevent, and the harms that the intervention imposes on the recipient. *Ceteris paribus*, an intervention will be justified, on this approach, only if the expected harm to others that is averted is greater than the harm inflicted on the recipient, and this, on one interpretation (the equivalent harm view), is precisely what the proportionality desideratum

requires. Satisfaction of the desideratum, thus interpreted, can therefore be regarded as an indicator of whether an IDC intervention can be justified on the constrained consequentialist approach.^{xli}

Certain constrained consequentialist approaches could also accommodate a proportionality desideratum that goes beyond the equivalent harm view and sanctions interventions that cause more harm than they prevent. For instance, a constrained consequentialist approach that incorporated desert into its account of the good could regard harms to culpable agents as less inimical to the good (because less underserved) than harms to innocent agents.^{xlii xliii}

It should be acknowledged that our use of the concept of proportionality here differs from the way in which proportionality is commonly invoked by retributivists in the criminal justice context. In this latter context, proportionality is commonly understood *retrospectively*; in order to ascertain whether a punishment is proportionate, we have to make a comparison between the harms that will be imposed by that punishment, and the gravity of the wrong involved in the offender's *past* action. In the case of retrospective proportionality assessments, we already know what crime the individual committed, and have to judge what sort of punishment is proportionate to that known harm. In contrast, we are invoking proportionality in a *prospective* sense; in order to ascertain whether a preventative intervention that we are now going to impose is proportionate, we have to make a comparison between the harm that will be imposed by this intervention, and the harm that the individual will otherwise bring about in their *future* action.^{xliv} This arguably introduces epistemic difficulties; we may face barriers to knowing the extent of the harm that the agent will bring about, as well as her degree of culpability for it.

IV The Justification of Non-Consensual IDC Interventions and Neurocorrectives

Having outlined the five desiderata that we take to constitute the standard approach to the assessment of nonconsensual IDC interventions, we shall now consider what this approach might imply regarding the justification of non-consensual *neurocorrectives*, assuming that it is apt for the moral assessment of such interventions. Our discussion of non-consensual IDC interventions in the previous section suggests that there may be some circumstances in which it may be permissible to impose medical interventions non-consensually. For the sake of argument, we shall assume that the non-consensual IDC interventions surveyed above are all permissible. We shall argue that considerations pertaining to the gravity of harm targeted, effectiveness, and opportunity costs count in favour of non-consensual neurocorrectives, just as they count in favour of the putatively permissible non-consensual IDC interventions surveyed above, including quarantine, coercive vaccination programmes and forced treatment. We shall then consider whether the least restrictive alternative and proportionality desiderata can be invoked to rule out the permissibility of neurocorrectives in a manner that cannot be applied to non-consensual IDC interventions. We shall argue that although some neurocorrectives would be more restrictive than other available methods for preventing recidivism, this is neither clearly the case for all neurocorrectives, nor clearly sufficient for establishing that these interventions would be impermissible. Similarly, we shall argue that some neurointerventions would not, in many cases, clearly be disproportionate to the harms they are intended to prevent.

In the preceding section, we outlined five desiderata that may be used to guide the moral assessment of non-consensual IDC interventions. One potential obstacle facing the

translation of a moral framework from public health ethics to the context of criminal justice is that interventions in the latter context may have different aims. For instance, on retributive approaches to criminal justice, the purpose of criminal justice is to ensure that the offender gets his ‘just deserts’. Such approaches are backwards-looking in the sense that they claim that the appropriateness of a particular correctional intervention depends ultimately on the offender’s past conduct (that which renders them deserving of hard treatment) not on the effects of the intervention. As such, retributivism naturally contrasts with ‘forward-looking’ consequentialist approaches to criminal justice, which claim that the appropriateness of a particular correctional intervention is to be determined by whether it will lead to good consequences. Strict retributivists are unlikely to be convinced that a public health moral framework is suitable for adoption within the criminal justice system insofar as it does not incorporate backwards-looking retributive elements that they deem to be central to criminal justice.

Naturally, we cannot settle debates about the justification and aims of criminal justice here. However, it should be acknowledged that rehabilitation has been understood to be a central goal of criminal justice on a wide range of not only consequentialist penal theories,^{xlv} but also by non-consequentialist moral education and paternalistic theories of punishment.^{xlvi} Furthermore, we do not mean to claim that rehabilitation is the *only* goal of criminal justice; rather, we are claiming that insofar as rehabilitation is aptly construed as an appropriate goal of criminal justice (as it is on many theories), then this framework may be used to guide our moral assessments of how we may permissibly intervene in order to promote the rehabilitation. This is quite compatible with claiming that criminal justice ought also to incorporate other retributive elements; those who make this claim might feel it necessary to impose other forms of intervention on offenders in order to meet these aims, and to supplement the framework we present here with backwards-looking considerations that limit

the sorts of interventions that we may impose in the name of retributivism. The comparative weight that the desiderata we present here will bear against other retributive desiderata (such as retrospective proportionality) will depend on the comparative weight that one affords to retributive aims and justifications of punishment more generally.

Furthermore, the desiderata that we have presented may be justified by either a constrained consequentialist approach or a deontologically-grounded theory of self defence. Notably, the latter can be understood as integrating both retributive and crime-prevention elements, since on this approach, force may only permissibly be used on aggressors (thus satisfying the retributivist's desert requirement), even if the primary purpose is to avert harm.^{xlvii} Moreover, even when the desiderata are understood as being undergirded by a constrained consequentialist theory, the approach is not vulnerable to some of the criticisms of advocating rehabilitation as a central unconstrained consequentialist aim of punishment. For instance, one such criticism is that an unconstrained consequentialist approach would seem to allow for *any* sort of harsh treatment that would achieve the goal of rehabilitation, including interventions that would traditionally be ruled out as disproportionate by retributive approaches. However, the constrained consequentialist approach that we have outlined can avoid this criticism by incorporating proportionality requirement that limits the sorts of interventions we may impose on offenders in the name of rehabilitation.^{xlviii}

Let us turn, then, to applying the five desiderata outlined in the previous section to the case of neurocorrectives'. Consider first the gravity of the harm that neurocorrectives seek to prevent. It seems plausible to claim that neurocorrectives could be used to prevent serious harms, since many crimes cause serious harm (both physical and mental) to their victims. Indeed, in some cases (most obviously in the case of murder), the harm caused to the victim may be comparable in seriousness to the harm of their contracting a lethal infectious disease.

Furthermore, criminal offending leads to other indirect costs for other members of society, since resources must be spent in order to, *inter alia*, provide support to victims of crime and apprehend and punish criminal offenders; in the UK, the cost to the taxpayer of reoffending is estimated to be £9.5 to £13 billion per year.^{xlix}

However, even if criminal offending often causes harm to its victims, it might be argued that there are nonetheless significant differences between the harms caused by criminal offending and the harms caused by the spread of infectious disease. First, it might be claimed that the *likelihood* that the harm will occur in the absence of intervention is typically higher in IDC than it is in criminal justice: we cannot predict whether a criminal offender will re-offend with the same degree of accuracy with which we can predict that an individual carrying an infectious disease will transmit it to others. Although various risk assessment instruments have been used to assist sentencing and release decisions in both the US and the UK^l, there are a number of well-documented problems with them.^{li} Recidivism can thus currently be predicted only with limited accuracy, and ‘false positive’ assessments are likely in this context.^{lii}

However, this does not mean that the harms associated with criminal recidivism can be predicted with less confidence than those associated with infectious disease; public health authorities sometimes have to use risk-prediction instruments that have comparably limited predictive accuracy. Consider cases in which the transmission of an infectious agent depends on the vector engaging in certain behaviours, such as in sexually transmitted infections. Assessing whether an infected individual poses a threat to public health, authorities must assess the risk of that agent’s engaging in unprotected sexual intercourse. It seems doubtful that they could make such an assessment with a significantly greater degree of certainty than an assessment of whether a violent criminal offender is likely to re-offend. Another case in

which public health authorities may face significant uncertainty in their predictions of threats to public health is in the implementation of quarantine. In quarantine, healthy individuals can be subjected to compulsory detention just because they have been exposed to an infectious agent, even though they themselves may not be infected, and the threat they actually pose to others is thus uncertain.

Nonetheless, as we described above, individuals carrying sexually transmitted infections have been subjected to non-consensual treatment in the name of public health in certain jurisdictions, and quarantine is widely practiced across various jurisdictions. As such, on the assumption that these interventions are morally permissible, the mere fact that the prediction of criminal recidivism might be highly uncertain is not *in itself* sufficient for establishing that the IDC moral framework cannot appropriately be applied in an assessment compulsory neurocorrectives; the prediction of infectious disease risks may also be highly uncertain.

Alternatively, it might be argued that the *aggregate* magnitude of the harm that could be prevented through the use of non-consensual neurocorrectives is minimal in comparison to that which could be prevented through the use of non-consensual IDC interventions. The latter often aim to prevent the spread of serious infectious diseases that might otherwise have the capacity to spread throughout a large population. In contrast, whilst criminal re-offending might entail certain indirect economic costs, it seems that an individual offender is likely to only cause serious direct individual harm to a small number of people in comparison to the large numbers of people who would be caused such harms by the spread of an infectious disease.

The strength of this argument turns on two empirical claims; first, that the scope of the harm caused by individual offenders is likely to be small, and second, the claim that

permissible non-consensual IDC interventions prevent harms with large scope. However, it is not clear that either claim is universally true. First, it seems possible that a criminal offender could be likely to directly harm a large number of people, for instance if the offender has previously been convicted of terrorist offences. Second, non-consensual medical interventions might not always be intended to prevent serious direct harm to large numbers of people; they might, for instance, be imposed in response to infectious diseases that are typically non-lethal. For example, it has been argued that influenza vaccinations should be mandatory for all health care workers, even though some have suggested that comparatively few deaths are attributable to the forms of influenza against which such vaccinations protect.^{liii} As such, the potential scope of the harms that neurocorrectives might prevent may not differ significantly from those that might be prevented by interventions in IDC.

Consider now the second and third desiderata: effectiveness and opportunity cost. It is reasonable to believe that some neurocorrectives will score at least as well on these desiderata as putatively permissible IDC interventions. Admittedly, since many of the neurocorrectives discussed above are only in developmental stages there is no robust data on whether these interventions would be effective in preventing recidivism. However, the limited data on the use of chemical castration to prevent recidivism in sex offenders might leave us with cause for optimism in this regard. In an extensive meta-analysis, Lösel and Schmucker found a 11.1% recidivism rate amongst sex-offenders who had been castrated following the offence compared to a 17.5% recidivism rate amongst non-treated offenders.^{liv} Second, there seems no reason to suppose that the opportunity costs associated with neurocorrectives would be any greater than those associated with IDC interventions. In fact neurocorrectives would be very similar in kind to the sorts of pharmacological interventions that might be used in IDC, and thus likely to involve similar resource expenditure.

Accordingly, it seems that some neurocorrectives would score at least as well as putatively permissible IDC interventions on the desiderata of effectiveness and opportunity cost. This suggests that, if these IDC interventions are indeed permissible, then non-consensual neurocorrectives are not universally ruled out as impermissible by these three desiderata. Nevertheless, it might be argued that ‘the least restrictive alternative’ and proportionality desiderata rule out the use of neurocorrectives. We shall conclude by considering each of these desiderata in turn.

First, however, it is necessary to explore further the nature and implications of the ‘least restrictive alternative’ desideratum. Suppose first that we adopt the rights-based understanding of the desideratum explored in the previous section. An assessment of an intervention’s restrictiveness will then take into account the nature of the rights that the intervention violates, the frequency and duration of the violation, and the number of people whose rights are violated.^{lv} Some comparisons of restrictiveness will be relatively straightforward; for instance, two interventions might differ only on one of the above dimensions. To illustrate, compare the use of directly observed therapy (DOT) to compelled treatment in IDC; the subject of DOT is required to take the prescribed dosage of his medication under observation.^{lvi} Whilst DOT arguably infringes on the individual’s rights in important ways, it seems clear that DOT is less restrictive than compelled treatment within a hospital setting, holding fixed the number of doses and people affected, since it does not infringe upon the individual’s right to freedom of movement and association to the same extent, insofar as it does not involve the imposition of involuntary confinement.

Comparisons of the restrictiveness of different interventions become less straightforward when they differ on more than one dimension, for example, where one intervention violates right M but not right N, and the other violates right N but not right M. Similarly, comparisons

become less straightforward when one intervention violates the rights of more people, and the other violates the rights more extensively. Finally, on the harm-based interpretation of the desideratum, it might not always be clear how we should compare different kinds of harm.

Nevertheless, there may be cases in which it is possible to make a somewhat credible comparison of the restrictiveness of the available alternative interventions. In the context of IDC, if our aim is to prevent further transmission of an infectious pathogen, we might compare compelled treatment with DOT, for example, and it might seem clear that DOT is less restrictive.

In the context of using neurocorrectives in criminal justice, in order to determine whether or not such interventions represent the least restrictive alternative means to achieving the end of preventing recidivism, we must compare them with the other means available to us.

It might be argued that a comparison of neurocorrectives with alternative means of preventing criminal recidivism will reveal neurocorrectives to be *more* restrictive than some available alternatives. Let us assume that we are comparing (i) subjecting one individual to a non-consensual neurocorrective, and (ii) subjecting the same individual to incarceration of the least restrictive kind necessary to retain whatever anti-recidivist effect incarceration has.^{lvii} We shall also assume the rights-based understanding of the least restrictive alternative desideratum, although our conclusions would also hold on the harm-based understanding.

One right that is clearly relevant to this comparison is the right to freedom of movement and association. Incarceration seems to infringe on an offender's right to freedom of movement and association, assuming that there is such a right and that it is not waived by the commission of a criminal offence.^{lviii} Moreover, infringements of this right may need to be continued for many years if the anti-recidivist effect is to be maintained. Neurocorrectives, by contrast, plausibly involve only a lesser violation of this right. Even if the offender were

required to present at a specified location for repeated iterations of a neurocorrective, this would clearly be a less extensive infringement of this right than incarceration.

Nonetheless, it seems that there is a strong case for claiming that the use of non-consensual neurocorrectives would be more restrictive than other penal methods because such interventions would violate other rights that might be deemed to be more important than the right to freedom of movement and association. First, insofar as neurocorrectives would entail some degree of physical invasion, non-consensual neurocorrectives would violate the offender's right to bodily integrity in a manner that is comparable to some of the non-consensual medical interventions surveyed above, *viz.* vaccinations, as well as compulsory medical examination and treatment. Although many jurisdictions permit non-consensual quarantine and isolation, which both seriously infringe upon an individual's rights to freedom of movement and association in the name of public health, such jurisdictions often prohibit non-consensual medical treatment, which would plausibly involve a serious violation of a right to bodily integrity (although it is not clear how much more serious this violation is than that which is involved in the non-consensual medical examinations). A plausible explanation for the acceptance of this position is that infringements of the right to bodily integrity are seen as more grave than infringements of the right to freedom of movement and association, *ceteris paribus*.

However, it is controversial whether this is so. Indeed, Thomas Douglas has argued that infringements of the right to free movement and association are typically equally grave, given their importance to maintaining our most valuable personal relationships. Moreover, even if infringements of bodily integrity *are* more important, other things equal, this may be offset by a difference in the *extent* of the rights infringements. Incarceration involves very extensive intrusions on freedom of movement and association; by contrast, neurocorrectives

need not always involve extensive intrusions on bodily integrity (for example, of the sort involved in surgical procedures) but might rather involve relatively moderate forms of physical interference, such as, for example, the administration of a drug via nasal spray, or the administration of a very weak electric current (as in transcranial direct current stimulation).

There is thus scope to question whether all neurocorrectives are more restrictive than incarceration in virtue of their physical invasiveness alone. However, another way in which opponents of neurocorrectives might argue that non-consensual neurocorrectives would be more restrictive than incarceration is by appealing to what we might term the offender's right to mental integrity, or their right to have “. . . the freedom to think one's own thoughts and to have one's own personality”.^{lix} Whilst incarceration can undoubtedly bring about some mental effects, it seems that an important moral difference between incarceration and neurocorrectives is that when these effects are brought about during the course of incarceration they are mediated by psychological processes.^{lx} In contrast, neurocorrectives would bring about profound mental effects directly through biological modulation of the brain states on which mental states supervene. Furthermore, these effects may also be unintended when they are brought about during the course of currently employed penal methods; in contrast, the intended purpose of neurocorrectives is to bring about these mental effects. In view of the fact that neurocorrectives involve intentionally changing another's mental states, those who stress the importance of relational conditions of autonomy might claim that the use of neurocorrectives involves the exertion of third party *control* over another's mental states in a way that having non-intentional effects on a prisoner's mental states does not.^{lxi} It thus seems plausible to claim that the former are more mentally invasive and arguably autonomy-undermining than traditional penal methods.

Perhaps it is plausible to maintain that neurocorrectives are more restrictive than incarceration in virtue of the mental interference that they involve, or in virtue of the combination of physical and mental interference. Again, however, it may be possible to respond by conceding that rights to mental integrity and bodily integrity are more important than rights to freedom of movement and association, while maintaining that the extensiveness of the violation of the latter right involved in incarceration exceeds the extensiveness of the violation of the former rights involved in the administration of some neurocorrectives. After all, just as some neurocorrectives might be administered through means that involve only moderate forms of physical invasion, so too might some neurocorrectives involve only mild or moderate forms of mental invasion: consider a neurocorrective that has a very local effect on one behavioural disposition (say, towards impulsive violence) which is not central to the offender's self-conception or personality.

Perhaps more importantly though, the use of non-consensual medical interventions in IDC suggests that the fact that an intervention is more restrictive than an alternative is not alone sufficient to establish that carrying out the intervention would be impermissible (this is why we referred to the 'least restrictive alternative' desideratum as a *desideratum*, not a requirement). One of the main problems with assessing the moral justification of different sorts of non-consensual interventions in IDC or criminal justice is that there will often be trade-offs between the desiderata that we delineated in the previous section. In particular, in many cases it is likely that there will be a trade-off between the effectiveness of an intervention and its restrictiveness. For instance, in the context of IDC, although DOT is less restrictive than compelled treatment, it also gives subjects more of an opportunity to avoid taking their medication. In such cases, we must judge that the least restrictive alternative is going to be sufficiently effective in achieving its aim to warrant our employing that intervention instead of a more restrictive alternative that is likely to be more effective.

The above considerations present a problem for opponents of neurocorrectives, since it is widely agreed that incarceration is ineffective at preventing recidivism amongst criminal offenders.^{lxii} A similar charge can be made against psycho-social rehabilitation. Consider, for example, the use of such programmes in the context of sexual offenders. Despite the wide use of these programmes amongst this population of offenders, a number of critics have argued that there is simply no good evidence that such programmes are effective in preventing re-offending.^{lxiii} With this in mind, even if a non-consensual neurocorrective is not the least restrictive means of preventing recidivism available, the use of such an intervention may nonetheless be morally permissible according to the moral framework outlined above, if this intervention is substantially more *effective* at preventing recidivism than other less restrictive means of achieving that end.

One way in which an opponent of neurocorrectives might respond to this point is to appeal to the proportionality desideratum. It might be argued that interventions with a degree of restrictiveness that is above a proportionality threshold would be ruled impermissible, even if they are the least restrictive of the available alternatives, or offer the best combination of restrictiveness and effectiveness. It could then be argued that neurocorrectives would, in virtue of their physical and mental invasiveness invariably fall above this threshold. Moreover, it might be argued that many IDC interventions are dis-analogous with neurocorrectives on this front; since many IDC interventions benefit their victim in some sense (in so far as they might treat the individual or prevent them from suffering an infectious disease), they are generally less restrictive, all things considered, than neurocorrectives, and are more likely to fall below a proportionality threshold of the sort that we are discussing here.^{lxiv}

At this juncture, it seems that the particular interpretation of the proportionality desideratum becomes particularly salient. Consider first the equivalent harm view of proportionality. On this view, the use of a non-consensual neurocorrective will only be proportionate if it prevents harms that are comparable to the harm caused by the intervention. As we have explained above, it seems that the use of some (although not all) non-consensual neurocorrectives might involve profound harms. It might be claimed that this restricts the scope of morally permissible non-consensual neurocorrectives to scenarios in which they are expected to prevent a similar degree of harm.

However, as we have seen, other interpretations of the proportionality desideratum allow that it can be permissible to impose harms on a threatening party that are greater than the harm that they themselves pose. On these accounts, if a criminal offender is believed to *culpably* pose a significant risk to the life of just one other person, then that may be enough to justify imposing significantly harmful preventative interventions on him. Crucially for our purposes, unlike many recipients of IDC interventions, it seems that the recipients of neurocorrectives can plausibly be understood to *culpably* pose a threat in a way that may plausibly justify imposing a harm on them that is greater than the harm that they themselves threaten to bring about. Moreover, as we have seen, significantly restrictive interventions have often been thought permissible in the context of IDC. These factors together casts significant doubt on the suggestion that neurocorrectives would invariably lie above the proportionality threshold.

Conclusion

Drawing on two theoretical frameworks, we have delineated five desiderata that are often thought apt to guide moral assessments of non-consensual medical interventions in the

context of IDC. We then employed these desiderata to assess the moral justifiability of non-consensual medical interventions performed for the purposes of criminal rehabilitation. We argued that at least some neurocorrectives could be highly effective at preventing considerable harms with few opportunity costs, thus satisfying three of the desiderata to a high degree. Having said that, we pointed out that, even though neurocorrectives need not always involve profound harms to the recipient of the sort that are described in the literature, they may not be the least restrictive alternative rehabilitative measure available. However, in assessing the non-consensual use of such interventions, considerations of restrictiveness must be weighed against the desideratum of effectiveness; even if less restrictive rehabilitative alternatives are available, we may nonetheless have sufficient reason to use neurocorrectives if they are more effective in preventing recidivism. Finally, we suggested that the fact that many serious criminal offenders are culpable for the threat of future harm that they pose suggests that significantly restrictive preventative interventions may satisfy the proportionality desideratum on plausible justice-based accounts of proportionality.

One concern that might be raised in response to our arguments is that they might plausibly be extended to justify the use of non-consensual crime-preventing medical interventions on individuals who have not previously committed a criminal offence, but who are still predicted to pose a threat of significant harm. This might be deemed problematic in our current political and legal culture, in which we are reluctant to punish people for their states of character, and instead believe that sanction should be imposed for conduct. Although we cannot offer a full treatment of this issue here, it should be noted that our arguments do not extend straightforwardly to this practice. Currently, one significant barrier to reliably predicting the risk of offending amongst non-offenders is that one of the main statistical predictors of future criminal offending is past criminal offending; for obvious reasons, this predictive factor would be absent in the risk assessment of non-offenders. Given the

potentiality for error and abuse, and the various reasons we have to try and minimize the extent to which the state may legitimately exercise coercive power over citizens, this epistemological barrier represents a sufficient reason to only permit the use of neurointerventions as a post-conviction sanction.

However, we suggest that this barrier is epistemological rather than moral. Suppose, completely fantastically, that we could, with 100% accuracy and reliability, predict with that an individual with no history of criminal offending will culpably bring about a significant harm, and that psycho-social rehabilitation would not be effective in preventing him from doing so. What should we do in the absence of the epistemological barrier? Whilst we should be wary of state abuse of this power, in the theoretically pure case of a thought experiment, we suggest that the ‘bullet’ of believing that it could be morally permissible to carry out a non-consensual neurocorrective on such an individual to prevent this harm becomes easier to bite; it is difficult to see how the mere fact that this person’s criminal offence lies in the future rather than the past should be an impediment to preventing the harm that they will certainly otherwise bring about.

Funding

We thank the Wellcome Trust [grant number 100705/Z/12/Z] and the Uehiro Foundation on Ethics and Education for their funding.

References

Ageing, Australian Government Department of Health and. “Strengthening Immunisation for Children.” Australian Government Department of Health and Ageing. Accessed February 5, 2015.
<http://immunise.health.gov.au/internet/immunise/publishing.nsf/Content/factsheet-strengthening-immunisation>.

- Anderson, Roy M., and R. M. May. "Immunisation and Herd Immunity." *The Lancet* 335, no. 8690 (1990): 641–45.
- Anderson, Roy M., Robert M. May, and others. "Vaccination and Herd Immunity to Infectious Diseases." *Nature* 318, no. 6044 (1985): 323–29.
- Bayer, R., D. Wilkinson, and R. Bayer. "Directly Observed Therapy for Tuberculosis: History of an Idea." *The Lancet* 345, no. 8964 (June 17, 1995): 1545–48. doi:10.1016/S0140-6736(95)91090-5.
- Beauchamp, Tom L., and James F. Childress. *Principles of Biomedical Ethics*. 6th ed.. New York ; Oxford: Oxford University Press, 2009.
- Bentham, Jeremy. *Theory of Legislation*. 2nd ed. London: Trübner & Co, 1871.
- Bomann-Larsen, Lene. "Voluntary Rehabilitation? On Neurotechnological Behavioural Treatment, Valid Consent and (In)appropriate Offers." *Neuroethics* 6, no. 1 (April 2013): 65–77. doi:10.1007/s12152-011-9105-9.
- Brink, David O. "Utilitarian Morality and the Personal Point of View." *The Journal of Philosophy* 83, no. 8 (August 1, 1986): 417–38. doi:10.2307/2026328.
- Bublitz, Jan Christoph, and Reinhard Merkel. "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination." *Criminal Law and Philosophy* 8, no. 1 (January 1, 2014): 51–77. doi:10.1007/s11572-012-9172-y.
- Carlen. "Crime, Inequality and Sentencing." In *Paying for Crime*, edited by Carlen and Cook. Open University Press, 1989.
- Caruso, G. "Free Will Skepticism and Criminal Behavior: A Public Health-Quarantine Model." *Southwest Philosophy Review* 32, no. 1 (2016).
- Center for Disease Control. "CDC - School and Childcare Vaccination Surveys," 1999. <http://www2a.cdc.gov/nip/schoolsurv/schImmRqmtReport.asp?s=grantee&d=10&w=%20&t=2>.
- Chen, M. K., and J. M. Shapiro. "Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-Based Approach." *American Law and Economics Review* 9, no. 1 (April 23, 2007): 1–29. doi:10.1093/aler/ahm006.
- Crisp, Roger. "Medical Negligence, Assault, Informed Consent, and Autonomy." *Journal of Law and Society* 17, no. 1 (1990): 77–89. doi:10.2307/1409956.
- Crockett, Molly J., Annemieke Apergis-Schoute, Benedikt Herrmann, Matthew D. Lieberman, Ulrich Müller, Trevor W. Robbins, and Luke Clark. "Serotonin Modulates Striatal Responses to Fairness and Retaliation in Humans." *The Journal of Neuroscience* 33, no. 8 (February 20, 2013): 3505–13. doi:10.1523/JNEUROSCI.2761-12.2013.
- Crockett, Molly J., Luke Clark, Marc D. Hauser, and Trevor W. Robbins. "Serotonin Selectively Influences Moral Judgment and Behavior through Effects on Harm Aversion." *Proceedings of the National Academy of Sciences* 107, no. 40 (October 5, 2010): 17433–38. doi:10.1073/pnas.1009396107.
- Crockett, Molly J., Luke Clark, Golnaz Tabibnia, Matthew D. Lieberman, and Trevor W. Robbins. "Serotonin Modulates Behavioral Reactions to Unfairness." *Science* 320, no. 5884 (June 27, 2008): 1739–1739. doi:10.1126/science.1155577.
- Cullen, Francis T. *Reaffirming Rehabilitation*. Cincinnati, Ohio: Anderson PubCo, 1982.
- Cullen, Francis T., Cheryl Lero Jonson, and Daniel S. Nagin. "Prisons Do Not Reduce Recidivism The High Cost of Ignoring Science." *The Prison Journal* 91, no. 3 suppl (September 1, 2011): 48S – 65S. doi:10.1177/0032885511415224.

- Dennis, Jane A., Omer Khan, Michael Ferriter, Nick Huband, Melanie J. Powney, and Conor Duggan. "Psychological Interventions for Adults Who Have Sexually Offended or Are at Risk of Offending." *The Cochrane Database of Systematic Reviews* 12 (2012): CD007507. doi:10.1002/14651858.CD007507.pub2.
- Doggett, Tyler. "Recent Work on the Ethics of Self-Defense." *Philosophy Compass* 6, no. 4 (April 1, 2011): 220–33. doi:10.1111/j.1747-9991.2010.00382.x.
- Doshi, Peter. "Influenza: Marketing Vaccine by Marketing Disease." *BMJ* 346 (May 16, 2013): f3037. doi:10.1136/bmj.f3037.
- Douglas, Kevin S., Stephen D. Hart, Jennifer L. Groscup, and Thomas R. Litwack. "Assessing Violence Risk." In *Handbook of Forensic Psychology*, edited by Irving B. Weiner and Randy K. Otto, 4th Edition. John Wiley & Sons, Inc., 2013.
- Douglas, Thomas. "Criminal Rehabilitation Through Medical Intervention: Moral Liability and the Right to Bodily Integrity." *The Journal of Ethics* 18, no. 2 (June 1, 2014): 101–22. doi:10.1007/s10892-014-9161-6.
- Ehreth, Jenifer. "The Global Value of Vaccination." *Vaccine* 21, no. 7 (2003): 596–600. doi:10.1016/S0264-410X(02)00623-0.
- El Almin, N. A., M.T. Parra, R. Kim-Farley, and J.E. Fieding. "Ethical Issues Concerning Vaccination Requirements." *Public Health Reviews* 34, no. 1 (2012).
- Eley, S., K. Gallop, G. McIvor, K. Morgan, and R. Yates. "Drug Treatment and Testing Orders: Evaluation of the Scottish Pilots. Scottish Executive Central Research Unit." Scottish Executive Central Research Unit, 2002. <http://www.scotland.gov.uk/Publications/2002/10/15537/11660>.
- Faden, Ruth R., and Tom L. Beauchamp. *A History and Theory of Informed Consent*. New York: Oxford University Press, 1986.
- Farah, Martha J. "Emerging Ethical Issues in Neuroscience." *Nature Neuroscience* 5, no. 11 (November 2002): 1123–29. doi:10.1038/nn1102-1123.
- Fazel, S., J. P. Singh, H. Doll, and M. Grann. "Use of Risk Assessment Instruments to Predict Violence and Antisocial Behaviour in 73 Samples Involving 24 827 People: Systematic Review and Meta-Analysis." *BMJ* 345, no. jul24 2 (July 24, 2012): e4692–e4692. doi:10.1136/bmj.e4692.
- Feinberg, Joel. *The Moral Limits of the Criminal Law*. New York ; Oxford: Oxford University Press, 1984.
- . *The Moral Limits of the Criminal Law Volume 3: Harm to Self [electronic Resource]*. New York: Oxford University Press, 1989.
- Gensini, G. "The Concept of Quarantine in History: From Plague to SARS." *Journal of Infection* 49, no. 4 (November 2004): 257–61. doi:10.1016/j.jinf.2004.03.002.
- Golash, Deirdre. *The Case against Punishment: Retribution, Crime Prevention, and the Law*. New York: New York University Press, 2005. <http://www.loc.gov/catdir/toc/ecip0419/2004015007.html>.
- Gostin, Lawrence O. *Public Health Law : Power, Duty, Restraint*. California/Milbank Series on Health and the Public ; 3 Y. Berkeley, Calif; London: University of California Press, 2001.
- Hampton, Jean. "The Moral Education Theory of Punishment." *Philosophy & Public Affairs* 13, no. 3 (July 1, 1984): 208–38.
- Harrison, Karen. "Dangerous Offenders, Indeterminate Sentencing, and the Rehabilitation Revolution." *Journal of Social Welfare and Family Law* 32, no. 4 (December 1, 2010): 423–33. doi:10.1080/09649069.2010.539915.

- Ho, David K., and Callum C. Ross. "Cognitive Behaviour Therapy for Sex Offenders. Too Good to Be True?" *Criminal Behaviour and Mental Health* 22, no. 1 (February 1, 2012): 1–6. doi:10.1002/cbm.1818.
- Hough, M., A. Clancy, T. McSweeney, and Turnbull, P. "Impact of Drug Treatment and Testing Orders on Offending: Two-Year Reconviction Results." London: The Home Office, Research, Development and Statistics Directorate, 2003.
- Iseman, Michael D. "Treatment of Multidrug-Resistant Tuberculosis." *The New England Journal of Medicine* 329, no. 11 (1993): 784–91. doi:10.1056/NEJM199309093291108.
- Ja, Mitty, Macalino G, Taylor L, Harwell Ji, and Flanigan Tp. "Directly Observed Therapy (DOT) for Individuals with HIV: Successes and Challenges." *MedGenMed : Medscape General Medicine* 5, no. 1 (February 2003): 30–30.
- Leverick, Fiona. *Killing in Self-Defence*. Oxford Monographs on Criminal Law and Criminal Justice Y. Oxford: Oxford University Press, 2006.
- Lu, Lin, Xi Wang, and Thomas R. Kosten. "Stereotactic Neurosurgical Treatment of Drug Addiction." *The American Journal of Drug and Alcohol Abuse* 35, no. 6 (2009): 391–93. doi:10.3109/00952990903312478.
- McMahan, Jeff. *Killing in War*. Uehiro Series in Practical Ethics. Oxford: Clarendon Press, 2009.
- . "Proportionate Defense." *Journal of Transnational Law and Policy* 23 (2014).
- . "Self-Defense and Culpability." *Law and Philosophy* 24, no. 6 (November 1, 2005): 751–74.
- . *The Ethics of Killing : Problems at the Margins of Life*. Oxford: Oxford University Press, 2002.
- McMillan, John. "The Kindest Cut? Surgical Castration, Sex Offenders and Coercive Offers." *Journal of Medical Ethics* 40, no. 9 (September 1, 2014): 583–90. doi:10.1136/medethics-2012-101030.
- Meisel, A., L. H. Roth, and C. W. Lidz. "Toward a Model of the Legal Doctrine of Informed Consent." *The American Journal of Psychiatry* 134, no. 3 (March 1977): 285–89. doi:10.1176/ajp.134.3.285.
- Miller, David. *Social Justice*. Oxford: Clarendon Press, 1976.
- Morris, Herbert. "A Paternalistic Theory of Punishment." *American Philosophical Quarterly* 18, no. 4 (1981): 263–71.
- Nozick, Robert. "Coercion." In *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, edited by Sidney Morgenbesser, Patrick Suppes, and Morton White. New York: St. Martin's Press, 1969.
- Offley, Will. "Should Influenza Vaccination Be Mandatory for Healthcare Workers?" *BMJ* 347 (November 12, 2013): f6705. doi:10.1136/bmj.f6705.
- Pereboom, Derk. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press, 2014.
- . "Free Will Skepticism and Criminal Punishment." In *The Future of Punishment*, edited by T. Nadelhoffer, 49–78. New York: Oxford University Press, 2013.
- . *Living without Free Will*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, 2001.
- Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13, no. 2 (1984): 134–71.

- Rawls, John. *A Theory of Justice*. Cambridge, Mass: Belknap Press of Harvard University Press, 1971.
- Reynolds v. McNichols, 488 (10th Cir. 1973).
- Rice, M. E., and G. T. Harris. "Violent Recidivism: Assessing Predictive Validity." *Journal of Consulting and Clinical Psychology* 63, no. 5 (October 1995): 737–48.
- Rodin, David. *War and Self-Defense*. Oxford: Clarendon Press, 2002.
- Rosati, Connie S. "Study of Internal Punishment, A." *Wisconsin Law Review* 1994 (1994): 123.
- Ryberg, Jesper. "Punishment, Pharmacological Treatment, and Early Release:" Edited by Elliot D. Cohen. *International Journal of Applied Philosophy* 26, no. 2 (2012): 231–44. doi:10.5840/ijap201226217.
- Salmon, Daniel A., Stephen P. Teret, C. Raina MacIntyre, David Salisbury, Margaret A. Burgess, and Neal A. Halsey. "Compulsory Vaccination and Conscientious or Philosophical Exemptions: Past, Present, and Future." *Lancet* 367, no. 9508 (February 4, 2006): 436–42. doi:10.1016/S0140-6736(06)68144-0.
- Schabas, Richard. "Severe Acute Respiratory Syndrome: Did Quarantine Help?" *The Canadian Journal of Infectious Diseases & Medical Microbiology* 15, no. 4 (2004): 204.
- Shaw, Elizabeth. "Direct Brain Interventions and Responsibility Enhancement." *Criminal Law and Philosophy* 8, no. 1 (January 2014): 1–20. doi:10.1007/s11572-012-9152-2.
- Sitaram, Ranganatha, Andrea Caria, and Niels Birbaumer. "Hemodynamic Brain-Computer Interfaces for Communication and Rehabilitation." *Neural Networks: The Official Journal of the International Neural Network Society* 22, no. 9 (November 2009): 1320–28. doi:10.1016/j.neunet.2009.05.009.
- Smith, P., C. Goggin, and P. Gendreau. "The Effects of Prison Sentences and Intermediate Sanctions on Recidivism: General Effects and Individual Differences." Ottawa, Ontario, Canada: Solicitor General of Canada, Corrections Research Branch, 2002. <http://publications.gc.ca/collections/Collection/JS42-103-2002E.doc>.
- State of Delaware. "Title 16: Health and Safety Regulatory Provisions Concerning Public Health, Chapter 5. Contagious Diseases Generally. Subchapter II. Tuberculosis Control." Accessed November 2, 2015. <http://delcode.delaware.gov/title16/c005/sc02/index.shtml>.
- Taylor, James Stacey. *Practical Autonomy and Bioethics*. New York ; London: Routledge, 2009.
- Thibaut, Florence, Flora De La Barra, Harvey Gordon, Paul Cosyns, and John M. W. Bradford. "The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Biological Treatment of Paraphilias." *The World Journal of Biological Psychiatry : The Official Journal of the World Federation of Societies of Biological Psychiatry* 11, no. 4 (2010): 604–55. doi:10.3109/15622971003671628.
- UK Gov. "Reducing Reoffending and Improving Rehabilitation - Policy - GOV.UK." Accessed February 17, 2015. <https://www.gov.uk/government/policies/reducing-reoffending-and-improving-rehabilitation>.
- UK Health and Social Care Act 2008*, 2008.
- Uniacke, Suzanne. "Proportionality and Self-Defense." *Law and Philosophy* 30, no. 3 (December 17, 2010): 253–72. doi:10.1007/s10982-010-9090-x.

Vanderzyl, K. "Castration as an Alternative to Incarceration: An Impotent Approach to the Punishment of Sex Offenders." *Northern Illinois University Law Review* 15, no. 1 (1994).

Wilkinson, T. M. "Contagious Disease and Self-Defence." *Res Publica* 13, no. 4 (March 13, 2007): 339–59. doi:10.1007/s11158-007-9024-0.

Young, Liane, Joan Albert Camprodon, Marc Hauser, Alvaro Pascual-Leone, and Rebecca Saxe. "Disruption of the Right Temporoparietal Junction with Transcranial Magnetic Stimulation Reduces the Role of Beliefs in Moral Judgments." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 15 (April 13, 2010): 6753–58. doi:10.1073/pnas.0914826107.

Zimmerman, David. "Coercive Wage Offers." *Philosophy & Public Affairs* 10, no. 2 (1981): 121–45.

ⁱ Notice that it might be possible to prevent recidivism in ways that do not involve the rehabilitation of the offender. For instance, incapacitating an offender by incarcerating them can be understood as an anti-recidivist measure. Since it seems that neuro-interventions that served to incapacitate an offender would face additional moral problems in comparison to rehabilitative neuro-correctives, in this paper we shall consider only those interventions that might be used to prevent recidivism by facilitating rehabilitation.

ⁱⁱ Eley et al., "Drug Treatment and Testing Orders: Evaluation of the Scottish Pilots. Scottish Executive Central Research Unit"; Hough et al., "Impact of Drug Treatment and Testing Orders on Offending: Two-Year Recidivism Results."

ⁱⁱⁱ Thibaut et al., "The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Biological Treatment of Paraphilias."

^{iv} Crockett et al., "Serotonin Modulates Behavioral Reactions to Unfairness"; Crockett et al., "Serotonin Selectively Influences Moral Judgment and Behavior through Effects on Harm Aversion"; Crockett et al., "Serotonin Modulates Striatal Responses to Fairness and Retaliation in Humans."

^v Lu, Wang, and Kosten, "Stereotactic Neurosurgical Treatment of Drug Addiction."

^{vi} Young et al., "Disruption of the Right Temporoparietal Junction with Transcranial Magnetic Stimulation Reduces the Role of Beliefs in Moral Judgments."

^{vii} Sitaram, Caria, and Birbaumer, "Hemodynamic Brain-Computer Interfaces for Communication and Rehabilitation."

^{viii} Douglas, "Criminal Rehabilitation Through Medical Intervention"; Shaw, "Direct Brain Interventions and Responsibility Enhancement"; Ryberg, "Punishment, Pharmacological Treatment, and Early Release."

^x Vanderzyl, "Castration as an Alternative to Incarceration: An Impotent Approach to the Punishment of Sex Offenders"; Bomann-Larsen, "Voluntary Rehabilitation?"; Rosati, "Study of Internal Punishment, A"; McMillan, "The Kindest Cut?"

^{xi} Derk Pereboom has discussed the possibility of providing a non-retributive justification of incarceration that is compatible with free will skepticism by appealing to similarities between the justification of quarantine and incarcerating criminal offenders on the assumption that they do not deserve to be punished. See Pereboom, *Free Will, Agency, and Meaning in Life*; Pereboom, "Free Will Skepticism and Criminal Punishment"; Pereboom, *Living without Free Will*. More recently, Gregg Caruso, "Free Will Skepticism and Criminal Behavior: A Public Health-Quarantine Model." has developed Pereboom's model within a broader justificatory framework of public health ethics. We are sympathetic to many of the claims made in these works, particularly Caruso's claim that we should give greater priority to the prevention of recidivism in criminal justice. However, unlike Pereboom and Caruso, our interest in this paper is not to establish a non-retributive justification of incarceration but rather to investigate the extent to which the moral justification of non-consensual medical interventions in public health might be used to justify the use of comparable non-consensual interventions in criminal justice. Furthermore, we consider alternative accounts of the justification of non-consensual public health interventions to Caruso. In fact, the arguments we make here may be understood to extend Caruso's

conclusions regarding the non-retributive justification of incarceration to the use of neurocorrectives as anti-recidivist measures.

^{xii} Feinberg, *The Moral Limits of the Criminal Law Volume 3*, 190.

^{xiii} *Ibid.*, 191.

^{xiv} Beauchamp and Childress, *Principles of Biomedical Ethics*; Faden and Beauchamp, *A History and Theory of Informed Consent*; Crisp, "Medical Negligence, Assault, Informed Consent, and Autonomy"; Meisel, Roth, and Lidz, "Toward a Model of the Legal Doctrine of Informed Consent."

^{xv} Ehreth, "The Global Value of Vaccination"; Gostin, *Public Health Law*, 376; Center for Disease Control, "CDC - School and Childcare Vaccination Surveys."

^{xvi} Anderson and May, "Immunisation and Herd Immunity"; Anderson, May, and others, "Vaccination and Herd Immunity to Infectious Diseases."

^{xvii} In contrast to attaching costs to vaccine refusal, Australia offers financial incentives to citizens to get vaccinated. See Ageing, "Strengthening Immunisation for Children." For comprehensive reviews of vaccination policies, see Salmon et al., "Compulsory Vaccination and Conscientious or Philosophical Exemptions"; El Almin et al., "Ethical Issues Concerning Vaccination Requirements."

^{xviii} Center for Disease Control, "CDC - School and Childcare Vaccination Surveys."

^{xix} Nozick, "Coercion"; Feinberg, *The Moral Limits of the Criminal Law*; Zimmerman, "Coercive Wage Offers"; Beauchamp and Childress, *Principles of Biomedical Ethics*.

^{xx} Reynolds v. McNichols, 488 (10th Cir. 1973).

^{xxi} State of Delaware, "Title 16: Health and Safety Regulatory Provisions Concerning Public Health, Chapter 5. Contagious Diseases Generally. Subchapter II. Tuberculosis Control," § 524–525; *UK Health and Social Care Act 2008*.

^{xxii} Gostin, *Public Health Law*, 415; Iseman, "Treatment of Multidrug-Resistant Tuberculosis."

^{xxiii} Gensini, "The Concept of Quarantine in History."

^{xxiv} Wilkinson, "Contagious Disease and Self-Defence," 343.

^{xxv} Gostin, *Public Health Law*, 21.

^{xxvi} Doggett, "Recent Work on the Ethics of Self-Defense."

^{xxvii} Wilkinson, "Contagious Disease and Self-Defence."

^{xxviii} There is, of course, much disagreement about what distributive fairness requires, but on many accounts, harms to worse off individuals detract more from distributive fairness than harms to better off individuals.

^{xxix} Schabas, "Severe Acute Respiratory Syndrome."

^{xxx} Note that it is the marginal effectiveness that should be considered; that is, the effectiveness that the intervention will add to other means to preventing harm that will be pursued regardless. Thus, for example, suppose that either a medical intervention or an education programme would alone lower the risk of harm from infectious disease compared to a baseline in which neither intervention were pursued. But suppose that the education programme will be pursued regardless, and that the medical intervention adds nothing to the effectiveness of the education programme—i.e., that the risk of harm is not reduced by adding the medical intervention to the education programme. In this context, the effectiveness of the medical intervention should be regarded as zero.

^{xxxi} Gostin, *Public Health Law*, 64–68.

^{xxxii} *Ibid.*, 68.

^{xxxiii} Relatedly, the literature on self-defence frequently invokes a 'condition of minimal force'; permissible acts of self-defence must involve the use of only the *minimum force* that is necessary to avert the threat. This can be regarded as one gloss on the 'least restrictive alternative' desideratum. See Wilkinson, "Contagious Disease and Self-Defence."

^{xxxiv} Gostin, *Public Health Law*, 68.

^{xxxv} McMahan, *The Ethics of Killing*, 412.

^{xxxvi} Although Uniacke herself argues against the equivalent harm view, she cites both Rodin, *War and Self-Defense*, 48–9; Leverick, *Killing in Self-Defence*, 153. as advocating it.

^{xxxvii} McMahan suggests that moral responsibility and culpability can be distinguished in the following way. A agent's degree of culpability is a function of a number of variables, including, *inter alia*, whether the agent poses the wrongful threat intentionally, recklessly, or negligently, whether the agent has an excuse and how strong that excuse is, and the magnitude of the threatened harm. However, even if the agent is fully excused for posing a threat of harm, they may still be liable to defensive harm if they bear some degree of *moral*

responsibility for posing a threat of harm. For instance, McMahan suggests that an agent who poses a threat of harm because he is acting under irresistible may still be morally responsible for their act even if they are not culpable; this is sufficient to render them liable to some degree of defensive harm on his view, although the proportionality restriction on that defensive harm will be more stringent by virtue of the fact that they are not culpable. See McMahan, *Killing in War*, 159–173.

^{xxxviii} Uniacke, “Proportionality and Self-Defense,” 261.

^{xxxix} McMahan, “Self-Defense and Culpability,” 766; McMahan, *The Ethics of Killing*, 401–411; McMahan, *Killing in War*, Ch. 4.

^{xl} McMahan, *The Ethics of Killing*, 412; McMahan, *Killing in War*, Ch.4; McMahan, “Proportionate Defense.”

^{xli} It is merely an indicator, rather than a determinant, because the intervention may also cause or prevent other harms that would need to be included in the consequentialist calculus, and because deontological side constraints may rule the intervention out even if it satisfies this desideratum.

^{xlii} See Miller, *Social Justice for an influential desert-based theory of social justice*; See Rawls, *A Theory of Justice* for criticisms of desert-based theories.

^{xliii} Indeed, a consequentialist theory of this stripe could hold that harms to culpable agents contribute positively to the overall good.

^{xliiv} McMahan draws a similar distinction in his discussion of the difference between proportionality in punishment and self-defence. See McMahan, “Proportionate Defense,” 22–23.

^{xli v} Bentham, *Theory of Legislation*; Cullen, *Reaffirming Rehabilitation*; Carlen, “Crime, Inequality and Sentencing.”

^{xli vi} Hampton, “The Moral Education Theory of Punishment”; Morris, “A Paternalistic Theory of Punishment.”

^{xli vii} Golash, *The Case against Punishment*, ch 5.

^{xli viii} Notice that even if one holds that criminal justice should not at all aim at rehabilitation, our argument will still have interesting implications. That public health ethics justifications for non-consensual interventions seem also to work for CPNs suggests that there is a *prima facie* case for using CPNs. Thus, even if one thinks that the latter could not permissibly be incorporated within criminal justice systems, our arguments suggest that there may be reasons to use CPNs *outside* of criminal justice, for instance, as part of some kind of separate public protection system.

^{xlix} UK Gov, “Reducing Reoffending and Improving Rehabilitation - Policy - GOV.UK.”

^l Douglas et al., “Assessing Violence Risk”; Fazel et al., “Use of Risk Assessment Instruments to Predict Violence and Antisocial Behaviour in 73 Samples Involving 24 827 People”; Harrison, “Dangerous Offenders, Indeterminate Sentencing, and the Rehabilitation Revolution.”

^{li} See Douglas et al., “Assessing Violence Risk” for discussion.

^{lii} A further higher order problem with actuarial instruments is that there are a number of difficulties in measuring their predictive accuracy. See Rice and Harris, “Violent Recidivism.”

^{liii} Offley, “Should Influenza Vaccination Be Mandatory for Healthcare Workers?”; Doshi, “Influenza.”

^{li v} It should be acknowledged that the figure quoted above refers to a sample that includes offenders who have received either chemical or physical castration.

^{li v i} Gostin, *Public Health Law*, 68.

^{li v i i} DOT has commonly been used for patients suffering from tuberculosis, and it is also occasionally used in the treatment of HIV. See Ja et al., “Directly Observed Therapy (DOT) for Individuals with HIV”; Bayer, Wilkinson, and Bayer, “Directly Observed Therapy for Tuberculosis.”

^{li v i i i} Note that the least restrictive kind of incarceration compatible with it retaining its anti-recidivist effect would plausibly be far less restrictive than prevailing kinds of incarceration, which arguably involve harms and rights violations that are gratuitous from the point of view of preventing recidivism, and indeed may serve to encourage it. It might be objected at this point that taking incarceration as the relevant comparator is illicit, since criminal justice systems frequently have at their disposal less restrictive means of preventing recidivism, such as psycho-social rehabilitation programmes. We will respond to this objection in our discussion below by appealing to the importance of the effectiveness desideratum and its relation to the least restrictive alternative desideratum.

^{lviii} Of course, on most accounts of punishment, the prevention of recidivism is not the only justification for incarcerating criminal offenders. Incarceration might also serve retributive and deterrent purposes. However, in cases of civil commitment, it seems that the primary justification of incarceration is to prevent recidivism.

^{lix} Farah, "Emerging Ethical Issues in Neuroscience," 1126.

^{lx} Bublitz and Merkel, "Crimes Against Minds."

^{lxi} Taylor, *Practical Autonomy and Bioethics*, chap. 1.

^{lxii} Cullen, Jonson, and Nagin, "Prisons Do Not Reduce Recidivism The High Cost of Ignoring Science"; Chen and Shapiro, "Do Harsher Prison Conditions Reduce Recidivism?"; Smith, Goggin, and Gendreau, "The Effects of Prison Sentences and Intermediate Sanctions on Recidivism: General Effects and Individual Differences."

^{lxiii} Ho and Ross, "Cognitive Behaviour Therapy for Sex Offenders. Too Good to Be True?"; Dennis et al., "Psychological Interventions for Adults Who Have Sexually Offended or Are at Risk of Offending."

^{lxiv} However, there are other contexts in which interventions that seem comparably restrictive to neurocorrectives are used non-consensually for the prevention of similarly problematic harms. For instance, many jurisdictions have mental health legislation that allows competent but mentally disordered individuals to be non-consensually treated with mind-altering drugs when they constitute a threat to themselves or others.