Neurointerventions, Self-Ownership, and Enforcement Rights

Peter Vallentyne, University of Missouri

Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice, edited by David Birks and Tom Douglas (Oxford University Press, forthcoming, 2018), pp. 124-39

Abstract (120-50 words): Individuals who have not intruded, and who do not risk intruding, upon the rights of others, normally are wronged by harmful non-consensual neurointerventions. Nonetheless, I argue that neurointerventions sometimes do not wrong the intervenee, namely when: (1) suitably valid consent has been given by the intervenee, or (2) the intervenee risks non-rightfully intruding upon the rights of others and the intervention is proportionate and necessary for suitably reducing the intrusion-harms she imposes, or (3) the intervenee is *not* psychologically autonomous and the intervention is in her interests. Moreover, in the second case, it *wrongs* an individual to impose harmful non-consensual *alternatives to neurointerventions* (e.g., incarceration) when they impose greater intrusion-harm on the individual and do not achieve a greater reduction in the relevant intrusion-harm she imposes.

Keywords (3-5): neurointerventions, self-ownership, consent, liability, autonomy

1. Introduction:

Under what conditions, if any, does it wrong an individual for the state, or private individuals, to administer to him a *neurointervention*? These are chemical, electrical, surgical, and other interventions that act directly on the brain. They may done for a variety of reasons: impulse and aggression control, drug addiction control, reduced/enhanced sexual drive, reduced delusions and hallucinations, enhanced capacity for guilt, empathy, and/or remorse, etc. They may be requested by the intervenee and administered with her full consent (where incarceration is not at issue), they may be administered with consent of the intervenee as a way of avoiding incarceration, or they may be imposed without the valid consent of the intervenee as a way of reducing the harms she risks imposing on herself or others.

I shall argue that neurointerventions need not wrong the intervenee when one of the following conditions is satisfied: (1) the intervenee is psychologically autonomous and has given valid consent, or (2) there is some chance that the intervenee will, in the future, impose unrectified non-rightful harms from rights-intrusions, and the intervention imposes no more harm than is proportionate and necessary for achieving a suitable reduction in such harm, or (3) the intervenee is not psychologically autonomous and the imposition is in her interests.

I shall focus on the question of whether it is possible *in principle* that a neurointervention not wrong the intervenee. My hope is that this will help set the stage for answering the important practical question of how common it is, under actual conditions, for neurointerventions not to wrong the intervenee.

As will become apparent, I believe that neurointerventions raise no fundamentally new moral issues. Their assessment is simply a matter of working out the implications of more general moral principles.

2. Rights, Intrusions, and Permissibility

Our central question concerns the conditions under which a neurointervention wrongs the intervenee. In this section, I clarify exactly what is at issue.

There are three (sometimes overlapping) ways that treating someone in a certain way can be (morally) impermissible. First, it may be *impersonally wrong* to treat that person in that way even if it does not wrong her or anyone else. I deny that there are any impersonal wrongs, but I cannot argue that point here. I merely set this possibility aside. Second, the action may *wrong a third party*. For example, if I promise my mother not to impose a neurointervention on Smith, my doing so may wrong my mother, even if it does not wrong Smith. Such cases can definitely arise, but I simply grant this and set such cases aside. Finally, the action may *wrong the person who is the direct object of the treatment*. This is the case on which I will focus.

A person is *wronged* by a given action just in case the action *infringes* her rights. It does not follow automatically that such an action is impermissible, since there may be an *overriding justification*. If there is, the right is merely infringed (not violated) and the action is not impermissible in virtue of the rights infringement (although it might be impermissible for other reasons). For example, pushing an innocent person gently in order to stop a bomb that will kill a million innocent people may be permissible, even though it wrongs her. A wronged individual is owed some kind of recognition of the failure to respect her rights. For example, she might be owed an apology and compensation for any harm suffered. In what follows, I focus solely on

whether the intervenee is wronged (her rights are infringed) and leave open whether there can be overriding justifications for such wronging. My general assumption is overriding justifications hold only in relative rare conditions—such as avoidance of a social catastrophe—but this will not be invoked below.

To explain the wrong done by neurointerventions, I shall appeal to rights-based *intrusions*. This is a more general category than rights-infringements. It includes all events that 'contravene' the protection offered by a right. If the event is not the result of an autonomous choice (e.g., when a person is windblown against another), then it is a *non-autonomous intrusion*, and it is neither permissible nor impermissible, and neither rightful nor wrongful (since only autonomous choices can have such status). If the event is the result of an autonomous choice, but the rightholder is *liable to the intrusion* because it is suitably necessary and proportionate (because she has conditionally forfeited some of the protection the rights provide), then it is a *rightful intrusion* against the rightholder, and it does not wrong her. If the event is the result of an autonomous choice and the rightholder is *not liable to the intrusion*, then it is a *wrongful intrusion* (i.e., a rights-infringement) against the rightholder, and it wrongs her.

Harms (setbacks to interests) need not involve intrusions (e.g., someone's suffering when you successfully court the one he loves), but when they do, they are *intrusion-harms*. Below, I will claim that the chance that one will impose unrectified non-rightful intrusion-harm can make one liable to neurointerventions (and other forms of defense).

3. Self-Ownership

At the core of my argument is the claim that psychologically autonomous agents have certain *rights of self-ownership* that protect their bodies and minds in certain ways. Neurointerventions, I

claim, wrong them when they infringe these rights, but need not wrong them when they do not.

Self-ownership is a bundle or rights. For autonomous agents, its core is:

> *Control Self-Ownership*: Each autonomous agent initially (prior to consensual transfers and forfeitures) has *control rights* over her person. These consists of: (1) *a claim-right* that others not use her person without her valid consent, (2) *a liberty-right* to use her person: others are not wronged merely because she uses her person without their consent, (3) *a moral power to authorize use* of her person by others: no one (the right-holder or others) is wronged merely because another uses her person in a way to which she has given, and not retracted, valid consent for that use by that other person.[1]

For example, for an autonomous agent with full control self-ownership, (1) she is wronged, if someone else cuts her hair without her valid consent, (2) she does not wrong anyone merely because she cuts her own hair (using scissors, etc. that she is at liberty to use) without anyone's consent, and (3) no one is wronged merely because another cuts her hair (using scissors, etc. that he is at liberty to use) with her on-going valid consent.

In addition to the control rights over one's person, self-ownership rights include: (1) *moral powers to transfer* rights to others (e.g., by gift or sale), (2) *moral immunities to loss*, if one does not intrude upon the rights of others, (3) *rights to rectification*, if one's rights are infringed, and (4) enforcement rights to use force against another to prevent one's rights from being infringed.

Below I shall appeal to *robust*, rather than *full*, self-ownership. Full self-ownership requires that the rights be *maximal* in the following three senses: (1) *unrestricted in content*

(include all of the above rights), (2) *unconditional in applicability* (e.g., the rights apply even when infringement is necessary to avoid a social catastrophe), and (3) *absolute in force* (i.e., there are no overriding justifications; it is always impermissible to infringe a right). Below, I do *not* assume self-ownership is full. In certain extreme circumstances, the rights may not apply, or, if they do apply, it may be permissible to infringe them because of an overriding justification (e.g., large benefits to a large number of people). I assume, however, that the rights are *robust* in the sense that (1) they (like full self-ownership rights) are unrestricted in content (e.g., the right to use oneself includes the right to kill oneself), (2) they apply under all circumstances, except perhaps relatively extreme and uncommon one, and (3) they cannot be overridden, except perhaps by extremely large benefits.

Robust control self-ownership entails that no one need be wronged, if an autonomous agent performs, or gives *valid consent* to someone else to perform, the following kinds of actions on her: surgically removing an unwanted arm, performing a transsexual operation, administering gender reassignment medications, administering mind-destroying drugs, or killing her. The point is that the agent is charge of her person and needs no one else's permission to rightfully engage in such actions or to authorize others to do so. These rights are, of course, not uncontroversial, but I shall assume that individuals have them.

Assuming (as we shall) that we are dealing with embodied beings, self-ownership rights are primarily rights to one's body. They are thus primarily rights to bodily integrity. Of course, the moral significance of using a person's body depends very much on the impact on her mind. My working assumption will be that, in the actual world, the only way to use or affect a person's mind is by physically using or affecting the person's body (e.g., getting you to believe something by telling you something requires that your ears, or eyes, receive physical signals from me). If

this is false, then there will be additional (non-physical) ways of infringing someone's control self-ownership.

One immediate, and non-controversial, implication of robust control self-ownership is that autonomous agents with such rights (i.e., those who have not transferred them away or forfeited them) are wronged by non-consensual neurointerventions. We shall now examine (1) consensual neurointerventions, and (2) ways in which an agent can lose some of her rights of self-ownership in virtue of the risk of her non-rightfully intruding upon the rights of others.

4. Consensual Neurointerventions for Autonomous Agents

In this section and the next, we shall focus on (*psychologically) autonomous agents*, which are individuals with a suitably robust capacity for reflection on, and rational revision of, their beliefs, desires, and intentions. There are, of course, many deep issues that need to be sorted out here, but I shall not attempt to do so. As a working assumption, I will simply assume that most 'cognitively normal' adults between the ages of 21 and 65 are autonomous in the relevant sense but most children under two and most severely cognitively impaired individuals are not. Psychological autonomy, of course, typically develops in degrees and will typically generate certain decision-relative rights (e.g., a five-year old may have the right to decide how to comb her hair but not to get a tattoo). I will not address how degrees of autonomy are to be handled.

Consensual neurointerventions are neurointerventions done with the consent (e.g., the public expression of endorsement by one's will) of the intervenee. When the consent is *valid*, the intervention does not wrong the intervenee. Consent, however, need not be valid, and so we shall address the conditions of validity for consent.

Valid consent is consent that meets the relevant conditions to be morally transformative

(e.g., to make an otherwise wrongful action rightful). One condition of validity is that the consenter be *psychologically autonomous* in some suitable sense. This requires a kind of internal freedom that is absent in the consent of a young child, a severely demented person, or of a psychotic person. A second condition on the validity of consent is that it be *suitably informed*. There are different views about what this requires, but it requires at least that the person to whom the consent is given, and others with whom he is collaborating, not knowingly provide false information to the consentee that is significantly relevant to her decision as to whether to consent. A third condition is that the consent be *suitably free*. There are different views about what this requires, but, at a bare minimum, it requires that the consenter not be threatened with *impermissible harm* (to herself or those she cares about) if she does not consent.

I shall assume that the above three conditions, suitably construed, are necessary and sufficient for valid consent. It follows immediately that consent to a neurointervention can, but need not be, valid and thus that such intervention need not, but can, wrong the intervenee. Let me make this explicit, given my assumptions above about valid consent.

Consent to neurointerventions *is not valid,* when the consent is given by a non-autonomous individual (e.g., a severely demented person), on the basis of fraud (e.g., about the consequences of the intervention or of not consenting), or on the basis of threats to impose impermissible harms on the consenter or those she cares about, if she does not consent. In such cases, even consensual neurointervention wrongs the autonomous intervenees.

Consent to neurointerventions *can be valid* and thus the intervention need not wrong the intervenee. Suppose, for example, that an *autonomous* agent wants a neurointervention to reduce anxiety, reduce aggression, reduce obsessive-compulsive tendencies, increase/decrease her sexual drive, or enhance her capacity for guilt, empathy, and remorse. Suppose further that there

is no fraud or deception and that the agent is well-informed with respect to the consequences of the alternatives she has. The consent is thus, I claim, *suitably informed.* Finally, suppose that consenter correctly believes that non-consent (1) produces prudentially acceptable prospects, and (2) does not increase the chance that someone will set back her interests (e.g., physically harm her or a loved one). The consent is thus, I claim, is *suitably free.* Under these conditions, consent to a neurointervention is valid, and the intervenee is not wronged by the intervention.

The above two cases are easy extremal cases. The hard work is for the intermediate cases. I shall here address just one intermediate case: the case where the consenter has, in some sense, no prudentially acceptable alternative to consenting. To keep it simple, I shall address only the case where the consenter is fully autonomous, is suitably informed, and has true beliefs about the consequences of non-consenting. I shall thus focus on the requirement that the consent be suitably free.

Some have claimed that the freedom required for valid consent requires that the agent have an *acceptable alternative* to consenting, where an acceptable alternative is one that has prudentially sufficiently good consequences (e.g., a decent life) for her.[2] This conception of freedom for valid consent, however, has the implausible implication that where natural circumstances are dire (e.g., after a natural disaster), and no options are prudentially acceptable for an individual, her consent can never be valid. If that is so, then autonomous agents lose some moral control over their bodies and their lives. Those who want basic medical treatment, for example, will no longer have the moral authority to give permission to others for such treatment. This is very implausible.

One way of weakening the acceptable alternative requirement condition is to require only an acceptable alternative to consent *when it is feasible* for the person(s) to be authorized by the

consent (or her collaborators) to provide such an option. So, if the agent to be authorized by consent is not able to provide an acceptable alternative, then the consent may be free. But if the agent to be authorized could provide an acceptable alternative in the absence of consent, but will not do so (and this is known), then the consent is not free.

This is certainly an improvement over the original condition, but it is too weak in one respect and too strong in another. Consider the case of a judge sentencing an offender for petty theft and giving her the option of a neurointervention. Suppose that a judge truly announces that, if the offender does not consent to the neurointervention, she will be imprisoned for life in a comfortable well-run prison. This, I here stipulate, will give the offender a prudentially acceptable life, but the punishment will be morally impermissible, because it is excessive relative to the crime. Here, the consent is not, I claim, suitably free. It was obtained on the basis of the prospect of an impermissible harm (even though the result is prudentially acceptable). Thus, the requirement that there be a prudentially acceptable alternative to consent is too weak (on its own).

The prudentially acceptable requirement is also too strong. Suppose that a judge truly announces that, if a serial murderer does not consent to the neurointervention, she will be imprisoned for life under harsh conditions. Such imprisonment, I stipulate for the sake of the example, is not prudentially acceptable but is morally permissible. I claim that consent under such conditions can be suitably free. Although the consent is given in order to avoid prudentially unacceptable results, this does not undermine its freeness, given that those results would be permissible to impose. It is like a storeowner consenting to the removal of sexist signs in his store only because local citizens will otherwise (permissibly) boycott his store and put him out of business (which is prudentially unacceptable). It is also like someone consenting to a

neurointervention to avoid abandonment by his partner, where abandonment is prudentially unacceptable but morally permissible (e.g., because the partner has already made major sacrifices for their relationship).

The freeness of the consent is, I claim, determined by the moral permissibility of the actions taken, if consent is not given, by the agent receiving authorization from the consent (and of his collaborators)—not the prudential acceptability of those results. Of course, the moral permissibility of the actions taken will be a morally contested issue. That, I claim, is as it should be.

With respect to consensual neurointerventions for convicted criminals, where consent is given to avoid certain levels of incarceration or punishment, the validity of the consent depends (in part) on whether the consent was freely given, and that, I claim, depends on whether it was given in order to avoid impermissible setbacks to the convict's interests. Because I believe that many (and perhaps most) cases of incarceration/punishment in the modern U.S.A are *impermissible* and *wrong* the person incarcerated, I believe that, in such cases, consent given to avoid the incarceration/punishment is not valid because not suitably free. I believe that our rights to rectify impermissible behavior are limited to *the least harmful way* of reducing the person's (or those of others) unrectified non-rightful intrusion-harms. In general, incarceration (or at least our current practice thereof) does not meet this condition. First, there is incarceration for actions that did not harmfully wrong another (formerly gay sex, possession of marijuana, etc.). Second, even where the individual did wrongfully harm another, incarceration (at least for the periods typically used) is usually not the most effective way of obtaining as much rectification as possible. Of course, this depends on what is required for rectification. I hold that it is only compensation and public recognition of the wronging and harming. Usually, this is better

accomplished by docking the pay of the individual, requiring community service, and perhaps requiring counseling. (Incarceration may, of course, be required for various recalcitrant cases.) Given that this is typically less harmful to the individual than incarceration, incarceration is often needlessly harmful and thus wrongful and impermissible.

I do not claim to have established that the incarceration we impose is typically impermissible. This is a deeply controversial issue. I merely note that, if (as I believe) our incarceration practices are typically impermissible, then my argument that consent to neurointerventions can be valid as an alternative to incarceration (etc.) does not apply to actual practice. We would first have to change our incarceration practices to make them permissible.

Of course, under some conditions, neurointerventions may not wrong the intervenee even if done without her valid consent. We shall now address that case.

5. Non-Consensual Neurointerventions for Autonomous Agents

Let us now consider the case of *non-consensual* neuro-interventions for autonomous agents (which, for simplicity, we will understand to include consensual cases where the consent is not valid). As argued above, if the agent still has full control self-ownership, then non-consensual use of her body (or person) wrongs the agent. The agent, however, need not still possess all those rights. She may have *forfeited* some of the protective force of her rights in virtue of past rights-intrusions or possible future rights-intrusions. Of course, agents don't unconditionally lose rights in this manner. They can still be wronged in various ways (e.g., torture for fun). A standard view, which I endorse, is that some of their rights become conditional on not being 'necessary and proportionate' for achieving certain moral goals, such as the reduction of wrongful intrusion-harm. An individual who has so forfeited some of the protective force of her rights in this

manner is said to be *formally (or potentially) liable* to having her rights (rightfully) intruded upon.[3] If the relevant necessity and proportionality conditions hold, then the individual is *effectively liable* to intrusion.

A crucial question is: Liable against whom? *In the absence of a state*, a natural reply is liable against the person, or persons, on whom the individual would, in the absence of defensive action, impose unrectified wrongful intrusion-harm. These individuals are defending themselves against such harm. The wrongful intruder may also be liable against *third parties*, when the third parties have been *given permission* by the rightful defending parties to aid them. A common view is that, where there is a sufficiently just state, then (1) the intruder is liable against relevant *state agents*, and (2) the intruder is *not liable* (or only under more restricted conditions) against those he will intrude against.[4] Although I am skeptical that the mere existence of a sufficiently just state has such a normative impact, when the intruded upon individuals have not consensually given up their enforcement rights, I shall leave this matter open. When I write that intruders are liable to certain interventions, I shall mean liable (1) only against those they intrude upon (and against anyone they authorize to help them), or (2) only against relevant agents of the state, or (3) both.

Can an autonomous agent be effectively liable to a neurointervention? This is so when (1) the intervention advances a morally relevant goal for liability to intrusion-harm, (2) the intrusion-harm imposed by the intervention is necessary for this advancement, and (3) the harm it imposes on the agent is proportionate relative the advancement of the goal. I shall argue that each of these conditions can be satisfied for neurointerventions.

Could a neurointervention advance a moral goal relevant for liability? That, of course, depends on what the morally relevant goals are. I claim that the *reduction in (expected value of)*

unrectified non-rightful intrusion-harms imposed by an agent is a relevant goal.[5] Recall that non-rightful intrusions are wrongful when done autonomously and they are neither rightful nor wrongful when done non-autonomously. To start, let us focus on wrongful intrusions, which is the primary case for autonomous agents.

The reduction in unrectified wrongful intrusion-harm is, I claim, a relevant goal. Reducing wrongful intrusions—harmful or not, rectified or not—is a broader goal that some would endorse. Here I appeal only to the less controversial goal of reducing *unrectified* wrongful intrusion-*harms*.[6]

I here leave open what *rectification* requires. Retributivists claim that it requires the imposition of suitable suffering on the wrongdoer. Some claim that it requires rehabilitation. As indicated, my own view (not assumed here) is that it only requires compensation and suitable public recognition of the wronging and harms. I leave all this open, but I shall assume that we have a measure of the degree of (the expected value of the) unrectified intrusion-harm for each intrusion-harm. Of course, this is an idealization. If rectification has several distinct requirements, the measure will need to be multi-dimensional, and various complexities will arise. This is admittedly an important issue, but I here set it aside.

When I speak *unrectified* intrusion-harm, I mean intrusion-harm that is not rectified by the intruder, whether or not rectification is *owed* by the intruder. Thus, for example, an agent may not owe any rectification (e.g., compensation) for the intrusion-harm for which she is not agent-responsible (e.g., because she could not have foreseen it). Nonetheless, assuming that she won't provide any rectification, reducing the intrusion-harm that she will impose counts towards the goal of reducing her unrectified non-rightful intrusion-harm.

Could a neurointervention reduce unrectified non-rightful intrusion-harms for an

autonomous agent? It seems clear that it can. Consider certain and imminent dangers. Shooting a chemically treated dart (e.g., with a sedative or a paralytic) can temporarily alter an agent's brain states and reduce/stop a wrongful aggression. Moreover, surgery or chemical treatments may reduce the agent's disposition to impose non-rightful intrusion-harms or increase his disposition to rectify them if they are imposed.

So, (1) a person can become formally liable to intrusion-harm, when it suitably reduces (the expected value of) his unrectified, wrongful intrusion-harms, and (2) neurointerventions can be effective in reducing such intrusion-harms. The two remaining crucial questions are: Can neurointerventions be necessary? Can they be proportionate? I shall argue that they can be.

There are different versions of the necessity condition, but the following captures the core idea: Imposing an intrusion-harm to an individual is *necessary* for an agent (e.g., the state) to achieve a given reduction in expected value of his unrectified non-rightful intrusion-harms just in case the agent has *no alternative action* that (a) equals or exceeds this reduction, (b) imposes no wrongful intrusion-harms on others, (c) is no worse for the agent (e.g., in terms of costs), and (d) is better for the offender. The core idea is roughly that the achieved reduction cannot be achieved in a way that is better for the formally liable person without imposing costs on others.

I shall assume that something like this condition is the relevant necessity condition for liability. It seems clear that sometimes neurointerventions can be necessary, in this sense, for reducing unrectified non-rightful intrusion-harms (e.g., where shooting a paralytic dart is the only way to stop a murder). The crucial question, then, is whether neurointerventions can be *proportionate* to the achieved reduction in unrectified wrongful intrusion-harms.

I assume (as is relatively common) that limits of proportionality increase with: (1) an increase in the expected value of the individual's future unrectified wrongful intrusion-harm, (2)

an increase in the individual's agent-responsibility for such intrusion-harm (e.g., the extent to which it was the foreseeable result of a resistible autonomous choice), and (3) an increase in the individual's degree of culpability (i.e., her agent-responsibility for acting wrongly, if she so acted).[7]

It seems clear that neurointerventions for an individual can be proportionate, given that (1) the achieved reduction in the expected value of the individual's future unrectified wrongful intrusion-harm can be quite high, (2) the individual can be highly agent-responsible for intrusion-harm she imposes, (3) the individual can have a high degree of culpability, and (4) the intrusion-harm imposed on the individual by a neurointervention can be quite low. I shall now defend this claim against two objections: (1) that proportionality imposes deontic constraints that neurointerventions violate, and (2) that proportionality imposes retributive considerations that neurointerventions violate.

One might hold that proportionality imposes certain *deontic constraints*, which rule out, at least typically, neurointerventions irrespectively of how harmful they are to the individual. Many neurointerventions may, for example, impose losses of *the capacity for rational reflection and choice* or impose losses of *mental integrity* (e.g., by subliminally fundamentally altering his preferences or beliefs). One might claim that such impositions are disproportionate independently of the harm they impose on the individual. This is certainly a possible, although non-standard, view. It is, however, implausible. A standard view sets the limits of proportionality based on the harms (setbacks to interests) of the individual. Those harms are sensitive to how much any loss of rational autonomy or mental integrity, for example, matters to the particular individual in the particular circumstances. I will now argue that, imposing a general deontic constraint, independently of the magnitude of the harms to the individual both gives too little,

and too much, protection to the liable individual.

A deontic constraint for proportionality gives *too little* protection to the liable individual, because it will rule out certain equally effective defensive actions, even though they are less harmful to him than alternatives otherwise deemed proportionate. Suppose, for example, that the defender has only two defensive options: (1) imposing a loss of mental integrity that imposes a minor harm to the individual, or (2) physically attacking him, which imposes a major, but proportionate, harm on him. In this case, the only proportionate defense is to impose the major harm, which is worse for the attacker than the minor harm from loss of mental integrity. Including the deontic constraint implausibly provides less relevant protection to the liable individual in this case.

The deontic constraint can also provide too much protection to the liable individual. Suppose, for example, that the defender has only two defensive options: (1) imposing a loss of mental integrity that imposes a minor harm to the individual (as above), or (2) physically attacking him, which a major harm to the individual, where this is *disproportionate* (unlike the above case). In this case, if there is a deontic constraint against imposing a loss of mental integrity, neither defensive action is proportionate, and the defending agent would wrong the attacker, if she imposed the minor harm on the attacker. This seems quite mistaken. If a minor harm is not disproportionate when it does not impose a loss of mental integrity (etc.), then surely it is not disproportionate when it does. What matters is the harm to the individual. Imposing certain deontic restriction inappropriately limits the defensive options open to defenders.[8]

Let us now consider a second objection: that neurointerventions can be disproportionate because they violate *retributive* considerations (e.g., that bad deeds be suitably punished) relevant to proportionality. This issue has been carefully, and correctly analyzed, by Ryberg[9] and

by Pugh and Douglas[10], and here I will merely focus on the core issue. First, like Ryberg, I'm skeptical that retributive considerations are relevant to proportionality. Instead, proportionality is based on the expected value of the wrongful intrusion-harm that the individual will impose, his agent-responsibility for such harm, and his culpability. These issues are, of course, related to what is retributively deserved. If retributivism takes the weak form of setting the *maximum* intrusion-harm that does not wrong an individual, with *no positive moral reason* for any intrusion-harm, then there need not be a significant difference between the two approaches. Both agree that some neurointerventions—those that are excessively harmful—are disproportionate. No one disagrees with this. The question is whether all neurointerventions are disproportionate, and the weak retributive approach does not entail this.

Consider, then, the strong form of retributivism, which also sets a (perhaps pro tanto) *minimum morally required* intrusion-harm for the individual. This view seems very implausible, but more importantly, it is fully compatible with the proportionality of neurointerventions. If the neurointervention is not harmful enough, one can supplement it with additional harmful treatment.

I conclude, then, that, on almost any plausible view of rectification, necessity, and proportionality, there will be at least some cases where a non-consensual neurointervention on an autonomous offender does not wrong her—for example, where the neurointervention is necessary and proportionate to prevent certain and imminent unrectified wrongful intrusion-harm. For concreteness, let us now identify some additional kinds of cases where this is so.

First, an agent can become liable to a non-consensual neurointervention where it is *less than certain* that he will impose imminent, unrectified wrongful intrusion-harm. The lower the chance, the more difficult it is to satisfy the necessity and proportionality requirements (since

they are based on the expected values involved), but if the harm from the infringement is large enough and the harm from the intervention low enough, those conditions can be satisfied.

More controversially, an agent can, I claim, become liable to a neurointervention even when the intrusion-harm is *not imminent*. The *temporal nearness* of the future rights-infringement is irrelevant. What matters is the *chance* of it occurring. If indeterminism is true, events in the distant future are typically less likely than comparable events in the near future (e.g., because they depend on more intermediate chancy conditions). Moreover, even if determinism is true, our *evidence* about what will (or may) happen in the distant future is much more limited that our evidence about events in the near future. Thus, in practice it may be rare, relative to our evidence, for an individual to be liable to a neurointervention to prevent some temporally distant unrectified wrongful intrusion-harm. In principle, however, this is certainly possible. For example, this will sometimes be true for individuals with highly aggressive tendencies who are very unlikely to fully rectify the wrongings they commit. Even if the intrusion-harms they impose are not imminent, they can be liable to neurointerventions (e.g., testosterone-lowering) to reduce them.

I further claim that the prevented unrectified intrusion-harm need not be wrongful. It may simply be *non-autonomous* (e.g., when the wind unexpectedly blows an agent's body against another, or when an infant or demented person fires a gun at someone). These are rights-intrusions, since they treat others in ways that would be unjust if they were the result of autonomous choices. They are not infringements, and thus not wrongful, because they are non-autonomous. Nonetheless, an individual can, I claim, be liable to a neurointervention to prevent unrectified, non-autonomous intrusion-harm. Let me give an example.

Suppose that an autonomous agent is prone to periodic loss of the capacity for

psychologically autonomous choices and, during such periods, she typically intrudes upon the rights of other by engaging in violent behavior. Suppose that the timing of occurrences is not predictable. Because these rights-intrusions are not autonomous, they are not rights-infringements (and do not wrong their victims). Of course, the agent, while autonomous, may owe others a duty to take step to reduce the chances of this happening, and failure to take those steps would infringe the rights of others. Even if the agent has no duty to take suitable steps, however, if she fails to take such steps, she becomes, I claim, liable to others taking steps to suitably reduce the chance of the intrusions. Others do not need to stand by and wait for the individual to lose her autonomy and intrude upon others. Under such conditions, I claim, an autonomous individual who is sufficiently likely to non-autonomously impose unrectified, intrusion-harm in the future (e.g., because she is unlikely to take her medications), is liable, when necessary and proportionate, to a certain amount of intrusion-harm. Thus, if non-consensual neurointerventions are necessary and proportionate (and I see no reason to think that they cannot be), administering them non-consensually does not wrong the agent.

Of course, one might object that it is never *proportionate* to impose intrusion-harm on an individual to prevent his *non-autonomous* intrusion-harm. After all, the individual is not agent-responsible for the intrusion-harm, nor for acting wrongly. I agree that the proportionality limits for reducing non-autonomous intrusion-harm are lower (often radically) than those for reducing wrongful intrusion-harm for which the individual is agent-responsible (especially when culpable). I deny, however, that those limits are zero. This is, however, a controversial issue that we cannot resolve here.[11]

Above we considered how neurointerventions can reduce *primary* non-rightful intrusion-harms. They can also increase the extent to which the intervenee *rectifies* such harms. At least

sometimes a neurointervention will make someone more likely to provide owed rectification (compensation, repentance, rehabilitation, submission to punishment, etc.). This both reduces the *unrectified* intrusion-harm from primary wrongings and eliminates the *secondary wronging* of failing to provide owed rectification (when such rectification is owed).

Thus, there is a wide range of cases in which non-consensual neurointerventions will not wrong autonomous agents.

6. Neurointerventions for Non-Autonomous Individuals

Let us now consider neurointerventions for non-autonomous beings. For simplicity, we set aside cases where an autonomous agent loses her capacity for autonomy only briefly (e.g., general anesthesia or a temporary coma). We shall consider cases such as young children and people with radical cognitive impairments (e.g., severely demented).

A basic question is whether a non-autonomous individual can ever be wronged. To be wronged (in the sense relevant here) is to have a right infringed. Some people think that non-autonomous individual have no rights, on the ground that rights only protect autonomous wills.[12] This has the crazy implication that infants and severely demented adults have no rights and thus cannot be wronged. Of course, people can treat them in ways that are wrong (impermissible), but the claim is that they are *not wronged* by such treatment. If it is wrong, it is an impersonal wrong or it wrongs someone else. Non-autonomous individuals are, for example, owed no special rectification (e.g., compensation).

It seems clear to me that infants and severely demented adults can be wronged in that they can be owed some kind of rectification by those who treat them in certain ways. More generally, I believe that the sentient beings (or at least sentient human beings) have the same

rights over their bodies (and persons) as autonomous beings, except those rights protect their *interests* (e.g., wellbeing) rather than their wills. This protection can take different forms, but the simplest, and most plausible, I think, is where an autonomous intrusion wrongs them when it is against their interests (cf. without their valid consent) but not when it is in their interests (cf. with their valid consent). Thus, sticking a needle in an infant does not wrong her, when it is a highly effective and important vaccine with no ill effects, but does wrong her when it merely cause pain with no advancement of the infant's interests. The individual who wrongs the infant owes her some rectification (e.g., compensation) for the wronging. Obviously, there are lots of issues that need careful discussion, but I shall simply assume that something like this is true and appeal to it below.[13]

Let us now identify several conditions under which non-consensual neurointerventions for non-autonomous individuals do *not* wrong them. First, of course, is the case where the neurointervention *is beneficial* for the individual. It might be that, without the neurointervention, she is likely, in the future, to harm herself directly (e.g., by cutting herself) or indirectly (by harming others and then being wrongly beaten up or by permissibly being incarcerated in ways that are worse for her). These are cases where the neurointervention does not wrong the intervenee on paternalistic grounds. Although paternalism is controversial for autonomous agents, it is not for fully non-autonomous individuals (although it can be for partially autonomous beings, which are here ignored).

Harmful neurointerventions on non-autonomous individuals wrong them if they are not necessary and proportionate for reducing the *future* unrectified intrusion-harms that the non-autonomous person will impose. If, however, they are sufficiently likely to impose (non-autonomous) intrusion-harms on others, and the neurointervention is necessary and proportionate

to the reduction in such harms it achieves, then a neurointervention need not wrong them. Of course, crucial here is the claim that risks of future *non-autonomous* intrusion-harm can make one liable in the same way (although subject to more stringent proportionality requirement) as *wrongful* intrusion-harm. In the previous section, I claimed that this was so for autonomous agents with episodes of non-autonomous intrusions (e.g., when unexpectedly windblown). Here, I assume that it is also so for non-autonomous agents.

Of course, a crucial issue here is proportionality for non-autonomous intrusion-harms. As discussed above, for non-autonomous intrusions, proportionality is much more restrictive than the typical autonomous case. This is because, for non-autonomous intrusions, the intruders are neither culpable (responsible for acting wrongly) nor responsible for any intrusion-harm. My own view is that it is proportionate when the intrusion-harm imposed is no greater than the non-autonomous intrusion-harm avoided, but my most basic claim is that the proportionality limit is greater than 0 in such cases.[14]

I conclude that non-autonomous individuals are not wronged by neurointerventions that benefit them (in the long-run) nor by harmful neurointerventions that are necessary and proportionate to reduce their future unrectified, non-autonomous intrusion-harms to others.

7. Conclusion

I have argued that neurointerventions need not wrong the intervenees. They do not do so when either: (1) The intervenee is psychologically autonomous and (a) she has given valid consent—which can arise even when permissible incarceration will be imposed, if consent is not given, or (b) the neurointervention is necessary and proportionate to reduce the intervenee's unrectified non-rightful intrusion-harms. (2) The intervenee is not psychologically autonomous and (a) the

neurointervention is in her interests, or (b) the neurointervention is necessary and proportionate to reduce the intervenee's unrectified non-autonomous intrusion-harms. I have argued that each of these cases is possible.

I have not addressed the *important practical question* of *how common* it might be for neurointerventions not to wrong the intervenee. That depends both on the moral question about the conditions for necessity and proportionality for reducing intrusion-harms, and on lots of empirical facts that I am not qualified to assess: How much intrusion-harm, if any, does a particular neurointervention impose on the intervenee? How much does a particular neurointervention reduce the intervenee's expected future unrectified non-rightful harm from rights-intrusions? What are the consequences of alternatives to neurointerventions? Thus, although my argument leaves open the current practical implications for neurointerventions, it points to the relevant information for answering these questions.

I have focused on the in-principle moral status of particular neurointerventions. This sets aside a wide range of important practical issues that are morally relevant for the *adoption of laws*.[15] First, our knowledge of how the world works, and how neurointerventions in particular work, is very limited, and we are subject to various biases in the formation of our beliefs. Second, individuals with power often abuse that power. Third, individuals are not perfectly rational individuals. These and other considerations make it wise in general to proceed cautiously. Before making large scale changes in our laws, or practices, we should normally try out various small scale changes to learn more about how things work. We may learn that in practice, at least for the foreseeable future, the potential problems are just too great. If so, it may be morally impermissible for us to adopt laws that authorize state officials to impose harmful, non-consensual (and perhaps even consensual) neurointerventions.[16]

[1] For more on control self-ownership, see (1) John Christman, The Myth of Property (New York: Oxford University Press, 1994); (2) Peter Vallentyne and Bas van der Vossen, 'Libertarianism', *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall2010/entries/libertarianism/ and (3) Peter Vallentyne, 'Left-Libertarianism: A Primer', in Left Libertarianism and Its Critics: The Contemporary Debate, edited by Peter Vallentyne and Hillel Steiner (Palgrave Publishers Ltd., 2000): pp. 1-20. Self-ownership, even full self-ownership, is compatible with significant duties to aid others in virtue of using or appropriating natural resources. For more on this left-libertarianism view, see the last two pieces just cited.

[2] See, for example, Serena Olsaretti, Liberty, Desert, and the Market: A Philosophical Study (Cambridge: Cambridge University Press, 2004). For related discussion, see Alan Wertheimer, Coercion (Princeton: Princeton University Press, 1987) and Alan Wertheimer, Consent to Sexual Relations (Cambridge, MA: Cambridge University Press, 2003).

[3] For further discussion of how an individual can become liable to neurointerventions, see Thomas Douglas, 'Criminal Rehabilitation through Medical Intervention: Moral Liability and the Right to Bodily Integrity' Journal of Ethics 18 (2014), pp. 101–122.

[4] There is also the issue of whether individuals can become liable to intrusions by state officials when such intrusions have not been authorized by the primary victims of the individual. I think not, but leave this open as well.

[5] My full view is that the relevant goal can (but need not) include reductions in the unrectified wrongful intrusion-harms *of others* (which will sometimes justify deterrence). Here, however, I appeal only to the less controversial goal of reducing such harms imposed by the person whose

liability is being assessed. I also believe that reduction in the relevant intrusion-harms is the *only* relevant goal for liability, but I do not assert that here. See my "Self-Defense against Rights-Intrusions (Non-Culpable and Culpable)" in The Ethics of Self-Defense [tentative title], edited by Christian Coons and Michael Eric Weber (Oxford University Press, forthcoming 2016).

[6] Thus, I do not address *harmless* wrongful intrusions or harmful ones that the intruder will *fully rectify* in the absence of any special defensive action.

[7] Agents can be agent-responsible for intruding upon someone's rights but not culpable, if, for example, they freely and knowingly intruded but all their evidence suggested that the intrusion was permissible (because the person was liable or because there was an overriding justification).

[8] For further discussions of this important issue, see Martha J. Farah, 'Neuroethics: The Practical and the Philosophical' Trends in Cognitive Sciences, 9 (2005), pp. 34-40; Jan Christoph Bublitz and Reinhard Merkel, 'Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination', Criminal Law and Philosophy 8 (2014), pp. 51–77; Ryberg, 'Punishment, Pharmacological Treatment, and Early Release'; Douglas, 'Criminal Rehabilitation through Medical Intervention"; Thomas Douglas, 'Neurointerventions, Nudges and Mental Interference', this volume, ch. ?.

[9] Ryberg, 'Punishment, Pharmacological Treatment, and Early Release'

[10] Douglas 'Criminal Rehabilitation through Medical Intervention'.

[11] For additional support for the view that the risk of non-autonomous intrusions can make someone liable, see Judith Jarvis Thomson, 'Self-Defense', *Philosophy and Public Affairs* 20 (1991), pp. 300-302); Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford University Press, 2011); and Helen Frowe, *Defensive Killing* (Oxford: Oxford University

Press, 2014). For the opposing view, see Michael Otsuka, 'Killing the Innocent in Self-Defense', *Philosophy and Public Affairs* 23 (1994), pp. 74-94; Jeff McMahan, *Killing in War* (Oxford: Oxford University Press, 2009), e.g., pp. 170-181.

[12] See, for example, the discussion of the will theory in Matthew H. Kramer, N.E. Simmonds, and Hillel Steiner, A Debate over Rights (Oxford: Oxford University Press, 1998).

[13] More generally, I would argue that a plausible conception of rights, for both autonomous and non-autonomous sentient beings, will be a *choice-prioritizing* conception—combining interest-protecting and choice-protecting elements, with priority for the choice-protecting elements for autonomous agents. On this view, a rightholder is wronged by a given intrusion if and only if (1) she has freely and informedly *dissented* against it, or (2) (a) she has *not* freely and informedly *consented* to it, and (b1) the intruder had a reasonable and morally permissible *opportunity to request her consent* to the intrusion and receive a reply, or (b2) the intrusions is against her interests. For autonomous agents, this allows interests to be relevant when there is no dissent and no suitable opportunity to obtain consent (e.g., to push someone out of the way of a bus). For non-autonomous beings, only the interests matter, since their dissent and consent is never free and informed in the relevant sense. Here I ignore these complexities. I discuss an early version of this view in my 'Libertarianism and the State', Social Philosophy and Policy 24 (2007), pp. 187-205 and I revise it in my "Paternalism and Political Philosophy: Libertarian Perspectives", *Routledge Handbook on the Philosophy of Paternalism*, edited by Kalle Grill and Jason Hanna (Routledge Press, forthcoming 2017).

[14] I would argue that, although non-autonomous individuals can be liable to defensive attack against *primary* intrusions, they *do not owe any rectification* for their past intrusions, because

they are not agent-responsible for them. Thus, they are not liable to defensive attack against *secondary* intrusions (i.e., failure to fulfill a duty rectify a past intrusion), since they have no duties to rectify. I believe, however, that, if they were formerly autonomous and had outstanding duties of rectification when they become non-autonomous, they may be liable to rectification measures even though they have no current duty of rectification. Obviously, this complex issue requires more analysis.

[15] See, for example, Emily McTernan, 'Treating Crime: Lessons from History', this volume, ch. To be added, and Douglas 'Criminal Rehabilitation through Medical Intervention'.

[16] For helpful comments, I thank David Birks, Tom Douglas, two anonymous referees, and the audience at Neurointerventions and Treatment for Crime conference at the University of Oxford (2015).