

3-1-2022

## Neuroscience, Criminal Sentencing, and Human Rights

Elizabeth Shaw

Follow this and additional works at: <https://scholarship.law.wm.edu/wmlr>



Part of the [Criminal Procedure Commons](#), [Human Rights Law Commons](#), and the [Neuroscience and Neurobiology Commons](#)

---

### Repository Citation

Elizabeth Shaw, *Neuroscience, Criminal Sentencing, and Human Rights*, 63 Wm. & Mary L. Rev. 1409 (2022), <https://scholarship.law.wm.edu/wmlr/vol63/iss4/9>

Copyright c 2022 by the authors. This article is brought to you by the William & Mary Law School Scholarship Repository.  
<https://scholarship.law.wm.edu/wmlr>

# NEUROSCIENCE, CRIMINAL SENTENCING, AND HUMAN RIGHTS

ELIZABETH SHAW\*

## ABSTRACT

*This Article discusses ways in which neuroscience should inform criminal sentencing in the future. Specifically, it compares the ethical permissibility of traditional forms of punishment, such as incarceration, on the one hand, and rehabilitative “neurointerventions” on the other. Rehabilitative neurointerventions are interventions that aim directly to modify brain activity in order to reduce reoffending. Various jurisdictions are already using techniques that could be classed as neurointerventions, and research suggests that, potentially, an even wider range of rehabilitative neurointerventions may be developed. This Article examines the role of human rights (in particular, the moral right to mental integrity and the legal right against degrading treatment) as a constraint on the state’s use of neurointerventions. It also discusses the extent to which traditional forms of punishment, such as incarceration, interfere with the right to mental integrity.*

---

\* Elizabeth Shaw is a Senior Lecturer in the School of Law, University of Aberdeen. She received her LLB and LLM by research from the University of Aberdeen. She received her PhD from the University of Edinburgh in 2013. She is a co-director of the Justice Without Retribution Network. Her research interests are interdisciplinary, involving criminal law, philosophy and neuroscience. She is a co-editor of *Free Will Skepticism in Law and Society* (Cambridge University Press, 2019), and *The Routledge Handbook of Philosophy and Science of Punishment* (Routledge, 2021).

## TABLE OF CONTENTS

INTRODUCTION .....	1411
I. EXAMPLES OF NEUROINTERVENTIONS .....	1414
II. INTENDED AND UNINTENDED EFFECTS ON MENTAL INTEGRITY:	
PRISON VS. NEUROINTERVENTIONS .....	1417
III. PRACTICAL IMPLICATIONS .....	1431
CONCLUSION .....	1442

## INTRODUCTION

This Article will discuss ways in which neuroscience should inform criminal sentencing in the future. Specifically, it will compare the ethical permissibility of traditional forms of punishment, such as incarceration, on the one hand, and rehabilitative “neurointerventions” on the other. Rehabilitative neurointerventions are interventions that aim directly to modify brain activity in order to reduce reoffending.<sup>1</sup> As I will explain below, various jurisdictions are already using techniques that could be classed as neurointerventions, and research suggests that, potentially, an even wider range of rehabilitative neurointerventions may be developed. This Article will examine the role of human rights (in particular, the moral right to mental integrity<sup>2</sup> and the legal right against degrading treatment<sup>3</sup>) as a constraint on the state’s use of neurointerventions. It will also discuss the extent to which traditional forms of punishment, such as incarceration, interfere with the right to mental integrity.

Many jurisdictions rely heavily on incarceration as the default response to serious crime.<sup>4</sup>

---

1. For brevity, “neurointerventions” will be used to refer to rehabilitative neurointerventions. For similar use of the term neurointervention, see, for example, TREATMENT FOR CRIME: PHILOSOPHICAL ESSAYS ON NEUROINTERVENTIONS IN CRIMINAL JUSTICE 1-2 (David Birks & Thomas Douglas eds., 2018) [hereinafter TREATMENT FOR CRIME]; JESPER RYBERG, NEUROINTERVENTIONS, CRIME, AND PUNISHMENT: ETHICAL CONSIDERATIONS 1 (2019). The terms “neurocorrective” or “direct brain intervention” are also used in the literature as synonyms for rehabilitative neurointervention. *E.g.*, Thomas Douglas, *Nonconsensual Neurocorrectives and Bodily Integrity: A Reply to Shaw and Barn*, 12 NEUROETHICS 107, 107 (2019); Elizabeth Shaw, *Direct Brain Interventions and Responsibility Enhancement*, 8 CRIM. L. & PHIL. 1, 2 (2014).

2. See Thomas Douglas, *Criminal Rehabilitation through Medical Intervention: Moral Liability and the Right to Bodily Integrity*, 18 J. ETHICS 101, 119 (2014). Douglas suggests that the moral right to mental integrity is the “most promising” basis for objecting to neurointerventions. *Id.* For some discussion of the right to mental integrity in the context of neurointerventions, see, for example, Elizabeth Shaw, *The Right to Bodily Integrity and the Rehabilitation of Offenders Through Medical Interventions: A Reply to Thomas Douglas*, 12 NEUROETHICS 97, 98, 101 (2019); David Birks & Alena Buyx, *Punishing Intentions and Neurointerventions*, 9 AJOB NEUROSCIENCE 133, 136 (2018).

3. See European Convention for the Protection of Human Rights and Fundamental Freedoms art. 3, Nov. 4, 1950, 213 U.N.T.S. 221 [hereinafter ECHR] (“No one shall be subjected to torture or to inhuman or degrading treatment or punishment.”).

4. See Birks & Buyx, *supra* note 2, at 133.

The United States has the highest incarceration rate in the world,<sup>5</sup> with a prison population of 1,430,800 people.<sup>6</sup> With an incarceration rate of 144 prisoners per 100,000 inhabitants, England and Wales have the highest incarceration rate in Western Europe.<sup>7</sup> The prison population was 78,058 in March 2021 (having decreased due to COVID-19),<sup>8</sup> but it is predicted to increase to 98,700 by September 2026.<sup>9</sup> Imprisonment comes at a high economic cost to society—39 billion dollars of taxpayers' money is spent on incarceration each year in the United States.<sup>10</sup> Incarceration also comes at great personal cost to prisoners—according to a study of twenty-four high-income countries, the suicide rate among female prisoners is nine times higher than that of the general female population, and the suicide rate among male prisoners is three times that of the general male population.<sup>11</sup> Separation from family and friends, severely restricted freedom, and the stigma of incarceration, as well as dangerous and overcrowded conditions, are among the factors that can contribute to prisoner suicides.<sup>12</sup> It has been argued that if neurointerventions are developed to effectively aid criminal rehabilitation, then these interventions have the potential to generate economic savings and to spare offenders suffering by allowing them to avoid incarceration or be released earlier,<sup>13</sup>

---

5. *Highest to Lowest: Prison Population Total*, WORLD PRISON BRIEF, [https://www.prisonstudies.org/highest-to-lowest/prison\\_population\\_rate?field\\_region\\_taxonomy\\_tid=All](https://www.prisonstudies.org/highest-to-lowest/prison_population_rate?field_region_taxonomy_tid=All) [<https://perma.cc/WG8Y-4VPJ>].

6. E. ANN CARSON, PRISONERS IN 2019, NCJ 2551155, at 3 (2020), <https://bjs.ojp.gov/library/publications/prisoners-2019> [<https://perma.cc/XJ7T-G6FP>].

7. GEORGINA STURGE, UK PRISON POPULATION STATISTICS, HOUSE OF COMMONS LIB. NO. CBP-04334, 2020, at 24 (UK), <https://researchbriefings.files.parliament.uk/documents/SN04334/SN04334.pdf> [<https://perma.cc/M5TH-8NKD>].

8. MINISTRY OF JUSTICE, OFFENDER MANAGEMENT STATISTICS BULLETIN, ENGLAND AND WALES, 2020, at 2 (UK), [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/982382/Offender\\_Management\\_Statistics\\_Quarterly\\_Q4\\_2020.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/982382/Offender_Management_Statistics_Quarterly_Q4_2020.pdf) [<https://perma.cc/7QZP-EY4P>].

9. MINISTRY OF JUSTICE, PRISON POPULATION PROJECTIONS 2020 TO 2026, ENGLAND AND WALES, 2020, at 1 (UK), [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/938571/Prison\\_Population\\_Projections\\_2020\\_to\\_2026.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/938571/Prison_Population_Projections_2020_to_2026.pdf) [<https://perma.cc/3VT7-K3HM>].

10. Birks & Buyx, *supra* note 2, at 133.

11. Seena Fazel et al., *Suicide in Prisons: An International Study of Prevalence and Contributory Factors*, 4 LANCET PSYCHIATRY 946, 951 (2017).

12. Shaoling Zhong et al., *Risk Factors for Suicide in Prisons: A Systematic Review and Meta-Analysis*, 6 LANCET PUB. HEALTH e164, e166-70 (2021).

13. *E.g.*, Douglas, *supra* note 2, at 103; Birks & Buyx, *supra* note 2.

although, some fear that the dangers of mandatory neurointerventions may outweigh their advantages.<sup>14</sup> It is, therefore, of practical importance to compare these possible responses to crime and to consider how society can seek to reduce crime in a humane and effective way.

After briefly outlining examples of current and potential neurointerventions in Part II, this Article will then focus, in Part III, on an influential objection against neurointerventions—the idea that neurointerventions threaten mental integrity in a way that prison does not. Growing evidence of incarceration’s negative psychological and neurological effects presents a challenge for this objection.<sup>15</sup> Such evidence has been among the factors that have led some academics to raise the question of whether we should treat neurointerventions and prison on a par.<sup>16</sup> In defense of the original objection, it has been replied that even if, in some cases, prison infringes mental integrity to a similar extent as neurointerventions, the impact of prison on mental integrity is merely foreseen, whereas the threat that neurointerventions pose to mental integrity is intended.<sup>17</sup> Part III of this Article will critique this attempt to salvage the mental-integrity-based objection against neurointerventions. This Article will attempt to cast some doubt on whether the intention/side effect distinction has the *moral weight* in this context that has often been assigned to it. It will be argued that the intention/side effect distinction does not warrant regarding mandatory incarceration as being significantly less problematic than mandatory neurointerventions in cases where the adverse effects of prison and neurointerventions are similarly severe.

Part IV will discuss some of the practical implications of this argument. If we should regard the adverse effects of prison and neurointerventions as roughly morally equivalent (where they are similarly severe), one way to respond is to be equally willing to

---

14. *E.g.*, RYBERG, *supra* note 1; Shaw, *supra* note 1.

15. *E.g.*, Sjors Ligthart et al., *Prison and the Brain: Neuropsychological Research in the Light of the European Convention on Human Rights*, 10 NEW J. EUR. CRIM. L. 287, 299-300 (2019); GREGG D. CARUSO, REJECTING RETRIBUTIVISM: FREE WILL, PUNISHMENT, AND CRIMINAL JUSTICE 229-60 (2021).

16. *E.g.*, Douglas, *supra* note 1, at 117; James Edgar Lim, *Is Incarceration Better than Neurointervention? On the Intended Harms of Prison*, 9 AJOBNEUROSCIENCE 168, 169 (2018).

17. Birks & Buyx, *supra* note 2, at 134.

impose either and to allow the state to impose whichever is most effective at preventing crime. In contrast, this Article will favor a second, human-rights-based approach. I will suggest that the state has strong obligations to research techniques that minimize adverse so-called side effects or any adverse effects on mental integrity and to offer prisoners the choice where possible between prison and rehabilitative interventions. In the light of the philosophical discussion, this Article will make some suggestions about what would be the most ethically defensible interpretation and development of certain aspects of the legal right against degrading treatment.

### I. EXAMPLES OF NEUROINTERVENTIONS

As Areti Theofilopoulou notes, “[t]he most widely discussed case of neurointerventions in the context of criminal justice is that of anti-libidinal pharmacological agents administered to sex offenders.”<sup>18</sup> In the literature on neurointerventions, anti-libidinal drugs<sup>19</sup> that reduce the effects or production of testosterone are considered neurointerventions because they have a considerable impact on the brain and on aspects of the individual’s psychology such as sexual thoughts, desires, and motivations.<sup>20</sup> It has been suggested that the effects of anti-libidinal drugs on the brain are even more important in addressing sexual offending than the drugs’ physical effects on sexual functioning.<sup>21</sup> Even if someone’s ability to engage in penetrative sexual activity has been reduced/removed, the person can still find ways of committing serious sexual offenses if the desire/

---

18. Areti Theofilopoulou, *Punishment as Moral Fortification and Non-Consensual Neurointerventions*, 38 LAW & PHIL. 149, 155 (2019).

19. The terms “anti-androgens” or “chemical castration” are sometimes used to refer to such drugs. Anti-libidinal effects can be produced by hormonal testosterone-suppressing drugs such as cyproterone acetate, or by nonhormonal drugs such as some antipsychotics and selective serotonin reuptake inhibitor antidepressants. See Omer Khan et al., *Pharmacological Interventions for Those Who Have Sexually Offended or Are at Risk of Offending*, 2 COCHRANE DATABASE OF SYSTEMATIC REVIEWS. 1, 6 (2015).

20. See, e.g., David JE Byrne, *Neurointerventions, Recidivist Sex Offenders, and Situated Moral Agency: An Approach from the Margins*, 9 AJOB NEUROSCIENCE 158 (2018); Birks & Buyx, *supra* note 2, at 133; Jonathan Pugh & Hannah Maslen, *‘Drugs That Make You Feel Bad?’ Remorse-Based Mitigation and Neurointerventions*, 11 CRIM. L. & PHIL. 499, 500, 506-7 (2017).

21. See Henry T. Greely & Nita A. Farahany, *Neuroscience and the Criminal Justice System*, 2 ANN. REV. CRIMINOLOGY 451, 466 (2019).

motivation to do so remains.<sup>22</sup> Don Grubin explains the link between testosterone and the motivation to commit sexual offenses as follows:

In the male brain, testosterone receptors are most dense in hypothalamic nuclei, the amygdala and other areas of the limbic system, the prefrontal cortex, and the temporal cortex, all parts of the brain known to be involved in processing sexual stimuli or initiating or maintaining sexual behavior.... testosterone has effects on the responsiveness of both general and specific neurological arousal mechanisms, it influences the processing of sexual sensory stimuli, it impacts on motivation, attention, and mood, and it is associated with aggression and dominance, all of which are potentially relevant to sexually problematic behaviour. Therefore, treatments aimed at moderating the activity of testosterone ... can ... weaken the foundation on which sex offending sits.<sup>23</sup>

A variety of jurisdictions currently administer anti-libidinal drugs to sex offenders under either statutory provisions that specifically concern sexual behavior<sup>24</sup> or general mental health legislation.<sup>25</sup> Depending on the jurisdiction, these interventions may be offered on a voluntary basis within prison or in the community,<sup>26</sup> imposed as a compulsory treatment in the community or as a mandatory condition of parole,<sup>27</sup> or exchanged for early release.<sup>28</sup> It is not the purpose of this Article to engage in an analysis of legislation on this

---

22. *See id.*

23. Don Grubin, *The Pharmacological Treatment of Sex Offenders*, in *THE WILEY BLACKWELL HANDBOOK OF FORENSIC NEUROSCIENCE* 703, 711-12 (Anthony R. Beech et al. eds., 2018).

24. For example, California authorizes medroxyprogesterone acetate treatment for certain sexual offenders with victims under thirteen years old. CAL. PENAL CODE § 645 (West 2019).

25. For example, England and Wales permit the Secretary of State to regulate the administration of medicine to detainees for treatment of mental health disorders. Mental Health Act 1983, c. 20, § 58 (UK).

26. *See generally* Grubin, *supra* note 23, at 718 (identifying treatment options aimed at preventing recidivism and those taken in the course of psychological therapy).

27. *See* Lisa Forsberg & Thomas Douglas, *Anti-Libidinal Interventions in Sex Offenders: Medical or Correctional?*, 24 *MED. L. REV.* 453, 458, 461-62 (2016).

28. For example, in Western Australia, offenders' participation in sex offender treatment programs can be taken into account in decisions about releasing the offender. *Parole and Sentencing Legislation Amendment Act 2006* (WA) s5A (Austl.).



topic in a specific jurisdiction, but rather to discuss general moral principles and considerations relating to the human right against degrading treatment, which could be applied to such domestic laws.<sup>29</sup>

Concerns have been raised that, despite the widespread use of anti-libidinal treatments within criminal justice systems, the evidence concerning their effectiveness is not robust.<sup>30</sup> However, some studies have suggested that, at least for some offenders, anti-libidinal treatments are more effective than psychological therapies, such as cognitive behavioral therapy, although anti-libidinal interventions are usually administered in conjunction with psychological therapies.<sup>31</sup> Anti-libidinal treatments are also likely to be more effective at rehabilitating offenders post-release than incarceration alone.<sup>32</sup> Although incarceration prevents re-offending against members of the public (though not necessarily other prisoners) while the offender is detained, and some offenders subjected to longer sentences may “age out” of offending while incarcerated,<sup>33</sup> it has nevertheless been shown that incarceration has a “*criminogenic* effect” for a significant proportion of offenders<sup>34</sup> and likewise some group psychological behavior programs can also *increase* re-offending.<sup>35</sup>

Research suggests the potential development of a wider range of rehabilitative neurointerventions in the future. Potential treatments include using drugs that modulate the neurotransmitter

---

29. For a comparison of some of the domestic legal regimes governing the provision of sex offender treatments, see Lisa Forsberg, *Crime-Preventing Neurointerventions and the Law: Learning from Anti-Libidinal Interventions*, in TREATMENT FOR CRIME, *supra* note 1, at 44, 46-48, 52-53.

30. Khan et al., *supra* note 19, at 7.

31. Friedrich Lösel & Martin Schmucker, *The Effectiveness of Treatment for Sexual Offenders: A Comprehensive Meta-Analysis*, 1 J. EXPERIMENTAL CRIMINOLOGY 117, 135 (2005).

32. Daniel S. Nagin et al., *Imprisonment and Reoffending*, in 38 CRIME AND JUSTICE: A REVIEW OF RESEARCH 115, 138 (Michael Tonry ed., 2009).

33. See TREATMENT FOR CRIME, *supra* note 1, at 2.

34. See Nagin et al., *supra* note 32, at 144-45.

35. AIDAN MEWS ET AL., MINISTRY OF JUSTICE, IMPACT EVALUATION OF THE PRISON-BASED CORE SEX OFFENDER TREATMENT PROGRAMME (2017) (UK), [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/623876/sotp-report-web-.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/623876/sotp-report-web-.pdf) [<https://perma.cc/7YJP-NHS4>]. For discussion, see Elizabeth Shaw, *Counterproductive Criminal Rehabilitation: Dealing with the Double-Edged Sword of Moral Bioenhancement via Cognitive Enhancement*, 65 INT'L J. L. & PSYCHIATRY (2019).

serotonin to reduce aggression and impulsivity<sup>36</sup> and using transcranial magnetic stimulation (TMS)—placing a device against the skull that transmits magnetic pulses to stimulate nerve cells in the brain—to increase empathy.<sup>37</sup>

Neurointerventions may cause a range of potentially worrying mental and physical effects and side effects.<sup>38</sup> This Article will focus on the impact of neurointerventions on *mental integrity*, which, as explained below, is distinct from physical symptoms or suffering caused by these interventions. The mental integrity objection is particularly significant because it can be argued that all neurointerventions affect mental integrity, whereas it is possible to imagine neurointerventions that have minimal physical side effects and cause minimal suffering.<sup>39</sup> If the impact of neurointerventions on mental integrity is problematic, then this would have wider implications than arguments based on the suffering or physical symptoms caused by neurointerventions.

## II. INTENDED AND UNINTENDED EFFECTS ON MENTAL INTEGRITY: PRISON VS. NEUROINTERVENTIONS

It has been argued that we can distinguish prison from neurointerventions based on the impact neurointerventions have on mental integrity.<sup>40</sup> Although it is often assumed that we have a moral right to mental integrity, the literature on the precise scope of this right is very limited.<sup>41</sup> It is not possible within this Article to give a complete account of the content of the right to mental integrity. Rather, this Article will highlight the key aspects of this

---

36. Christopher Chew et al., *Biological Interventions for Crime Prevention*, in TREATMENT FOR CRIME, *supra* note 1, at 11, 22-25.

37. *E.g.*, Carmen S. Sergiou et al., *The Effectiveness of Transcranial Direct Current Stimulation as an Intervention to Improve Empathic Abilities and Reduce Violent Behavior: A Literature Review*, 55 AGGRESSION & VIOLENT BEHAV. 1, 2 (2020).

38. *See, e.g.*, Shaw, *supra* note 35, at 4; Birks & Buyx, *supra* note 2, at 135-37.

39. *See* Peter Vallentyne, *Neurointerventions: Punishment, Mental Integrity, and Intentions*, 9 AJOB NEUROSCIENCE 131, 131-32 (2018). TMS may be a real-life example of a neurointervention that causes minimal physical side effects and minimal suffering. *See, e.g.*, Sergiou et. al, *supra* note 37, at 2.

40. *See* Shaw, *supra* note 2, at 98.

41. Douglas observes that “[v]ery little work has been done to determine ... what varieties of mental influence” would infringe this right. Douglas, *supra* note 2, at 119.

right relevant to the comparison between prison and neuro-interventions. The right to mental integrity can be infringed by intentionally interfering with a person's mental states through non-rational means.<sup>42</sup> The wrongfulness of violating mental integrity cannot be reduced to the wrongfulness of inflicting "the experiential harm of suffering," and mental integrity can be violated "without any experiential harm."<sup>43</sup> Rather, the wrongfulness of violating mental integrity lies in the disrespect that is shown for the victim's rational agency, and is connected with the intuition that interfering with mental states through nonrational means can be "manipulative."<sup>44</sup>

Although there is a widespread intuition that neurointerventions pose a greater threat to mental integrity than prison, this intuition has been challenged with reference to evidence that prison can cause a range of adverse psychological and neurological effects.<sup>45</sup> One possible response to that challenge involves drawing the traditional distinction between intended effects and side effects.<sup>46</sup> Neurointerventions can threaten mental integrity in two ways—through both their intended effects and their side effects.<sup>47</sup> Firstly, regarding the intended effects of neurointerventions: even if these interventions pursue legitimate aims, such as the rehabilitation and/or reform of offenders, they achieve these aims through means that bypass rational agency, thereby posing a threat to mental integrity.<sup>48</sup> Secondly, the side effects of neurointerventions might be thought to threaten mental integrity by interfering with offenders' valuable mental states and choices.<sup>49</sup> Neurointerventions might

---

42. Shaw, *supra* note 1, at 2, 7-8; Jan Christoph Bublitz & Reinhardt Merkel, *Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination*, 8 CRIM. L. & PHIL. 51, 73 (2014). One might wonder whether the right to mental integrity protects *all* mental states. Birks and Douglas tentatively suggest that interfering with all-things-considered "disvaluable" mental states may not be objectionable. David Birks & Thomas Douglas, *Two Ways to Frustrate a Desire*, 51 J. VALUE INQUIRY 417, 418, 428-30 (2017). That issue will not be explored in the present Article.

43. Birks & Buyx, *supra* note 2, at 136.

44. *See id.* at 134.

45. Lighthart et al., *supra* note 15, at 298-300.

46. *See* Birks & Buyx, *supra* note 2, at 136.

47. *Id.* at 140.

48. *See* Shaw, *supra* note 1, at 2; Bublitz & Merkel, *supra* note 42.

49. *See* Birks & Buyx, *supra* note 2, at 136.

eliminate or diminish some legitimate desires offenders may have.<sup>50</sup> For example, as mentioned above, sex offenders are sometimes given testosterone-lowering drugs to try to reduce their deviant sexual desires.<sup>51</sup> This treatment also interferes with their capacity to experience legitimate sexual desires, impairing their ability to engage in harmless sexual relationships.<sup>52</sup> Aggression-lowering neurointerventions, which might be developed in the future, would aim at decreasing the likelihood of criminal violence, but could also interfere with offenders' legitimate desires and choices, such as a passion for boxing, a decision to join the army, or the decision to use force in self-defense.<sup>53</sup>

In contrast, it is often assumed that prison only threatens mental integrity through its side effects and not through its intended effects.<sup>54</sup> The intended effects of prison—for example, communicating censure and being a forum in which an offender can reform—are considered unproblematic, as the communication of censure and attempt to reform offenders involves engaging with offenders' rational agency.<sup>55</sup> The side effects of prison might interfere with offenders' valuable mental states and choices by causing or worsening offenders' mental health problems, either from separation from friends and family, from the constraint on liberty, or from the under-stimulating prison environment.<sup>56</sup> However, according to the traditional argument, the side effects of prison are not intended.<sup>57</sup> The judge, when sentencing someone to prison, does not intend to cause the person to experience mental health problems.

Many theorists believe that harms or rights infringements that are caused intentionally are harder to justify than those that are

50. *Id.*

51. Khan et al., *supra* note 19.

52. Birks & Buyx, *supra* note 2, at 136.

53. See John Harris, *Moral Enhancement and Freedom*, 25 *BIOETHICS* 102, 105 (2011).

54. For a critical discussion of this assumption and of its supposed moral relevance, see Adam J. Kolber, *Unintentional Punishment*, 18 *LEGAL THEORY* 1 (2012). For a long time, this assumption has featured in discussions of the justifiability of prison versus “treatments” for crime. See, e.g., Lawrence Stern, *Freedom, Blame, and Moral Community*, 71 *J. PHIL.* 72, 82 (1974) (conceding that “punishment does not aim to subvert rational and moral capacity” but denying that this is enough to justify preferring traditional punishment to treatment (emphasis added)).

55. See R. A. DUFF, *PUNISHMENT, COMMUNICATION, AND COMMUNITY* 28, 30, 108-09 (2001).

56. See Zhong et al., *supra* note 12.

57. See Kolber, *supra* note 54, at 2-3.

foreseen side effects.<sup>58</sup> If one subscribes to that view, one could therefore argue that neurointerventions are harder to justify than prison because neurointerventions interfere with mental integrity intentionally (as well as through side effects), whereas prison only interferes with mental integrity through side effects.<sup>59</sup> To clarify: this line of argument does not entail that causing adverse side effects is unproblematic, but only that it is easier to justify than causing *intended* harms or rights infringements.<sup>60</sup> Theorists who endorse this line of argument need not endorse current prison conditions, which are often inhumane.<sup>61</sup> Instead, these theorists typically focus on humane prison conditions, which nevertheless unavoidably have some negative side effects, and argue that *if* society only employed humane forms of imprisonment, then neurointerventions, with the same severity of negative side effects, would be harder to justify than humane incarceration.<sup>62</sup> For example, David Birks and Alena Buyx state, “when we discuss incarceration, we are imagining it to involve a diminution of the offender’s liberty, but without the overcrowding and risk of assault commonly experienced by incarcerated persons in the United States and United Kingdom. We call this minimally decent incarceration.”<sup>63</sup> In the minimally decent incarceration context, harms to mental integrity are considered “unintended but merely foreseen” while the identical harms are considered “intended” in the rehabilitative neurointervention context.<sup>64</sup>

Birks and Buyx made an important contribution to the debate about the relative impact of neurointerventions and prison on mental integrity by attempting to show that the intended effects of neurointerventions on mental integrity are even more problematic

---

58. For useful overviews, see generally Warren S. Quinn, *Actions, Intentions, and Consequences: The Doctrine of Double Effect*, 18 PHIL. & PUB. AFFS. 334 (1989); Dana Kay Nelkin & Samuel C. Rickless, *So Close, Yet So Far: Why Solutions to the Closeness Problem for the Doctrine of Double Effect Fall Short*, 49 NOÛS 376 (2015).

59. See Nelkin & Rickless, *supra* note 58, at 376, 378-79.

60. Birks & Buyx, *supra* note 2, at 136.

61. See, e.g., Heather Ann Thompson, *What’s Hidden Behind The Walls of America’s Prisons*, THE CONVERSATION (June 4, 2017, 9:45 PM), <https://theconversation.com/whats-hidden-behind-the-walls-of-americas-prisons-77282> [<https://perma.cc/4JB6-6FNP>].

62. See Birks & Buyx, *supra* note 2, at 134.

63. *Id.* (citation omitted); see also Douglas, *supra* note 2, at 105-06.

64. Birks & Buyx, *supra* note 2, at 141.

than previously assumed.<sup>65</sup> Birks and Buyx argue that the intended effects of neurointerventions are not limited to achieving punishment goals.<sup>66</sup> They argue that what theorists previously categorized as side effects of neurointerventions, such as dampening offenders' legitimate desires, can actually be recharacterized as *intended* effects.<sup>67</sup> In contrast, according to Birks and Buyx, prison's supposed side effects, such as deterioration in mental health, genuinely are side effects.<sup>68</sup> They argue that if someone who administers a drug knows that the drug produces more than one effect, then the person who administers it intends to produce all the known effects of the drug.<sup>69</sup> So, if a testosterone-lowering drug is known to diminish both (1) the desire to commit sexual offenses and (2) legitimate sexual desires, the person who administers the drug cannot claim only to have intended to diminish the desire to commit sexual offenses.<sup>70</sup> According to Birks and Buyx, they must have intended to diminish both (1) and (2); whereas traditionally, the diminishing of (2) would have been considered a mere side effect.<sup>71</sup> If Birks and Buyx's argument succeeds, it could make neurointerventions seem even more problematic than they already appear on the traditional version of mental integrity-based objection.<sup>72</sup> If intentionally diminishing *illegitimate* sexual desires using brain interventions is hard to justify because it is a manipulative violation of mental integrity, intentionally diminishing *legitimate* sexual desires might seem even harder to justify.<sup>73</sup>

Birks and Buyx's recharacterization of the diminishment of legitimate desires as being an intended effect depends on the claim that the effects of an action can only be considered "side effects" if they are *causally* related to what one intended to do, rather than

---

65. *See id.* at 133-36, 141.

66. *Id.* at 137-41.

67. *Id.* at 141.

68. *Id.* Birks and Buyx do not reach a conclusion about whether neurointerventions are objectionable *all things considered*. Their aim is limited to arguing that, in cases where neurointervention and prison have a similarly severe impact on mental integrity, neurointerventions are *in one respect* harder to justify. *Id.* at 134, 140-41.

69. *Id.* at 140.

70. *See id.*

71. *See id.* at 135, 140.

72. *See id.* at 134-35, 141.

73. *See id.*



having a *constitutive relationship* to what one intended to do.<sup>74</sup> To illustrate this point, they draw on a famous example used by Philippa Foot in which, during a cave exploration, the largest explorer gets stuck in the only exit and cannot be moved.<sup>75</sup> The only way the other explorers can survive is to blow him up with dynamite in order to escape the cave.<sup>76</sup> (This example is introduced in order to understand the meaning of the term “intention,” not in order to discuss whether the explorers’ actions are ethically permissible.<sup>77</sup>) The surviving explorers cannot argue that they did not intend to kill their companion, but merely meant to blow him to pieces, because blowing him to pieces is too “close” to killing him—blowing him to pieces *constitutes* killing him, rather than merely *causing* his death.<sup>78</sup> Birks and Buyx apply this reasoning to the example of the testosterone-lowering drug that diminishes both (1) illegitimate sexual desires and (2) legitimate/valuable sexual desires, and they conclude that: “The state of affairs of diminishing [the offender’s] testosterone by administering [the neurointervention] is constitutive of the state of affairs of [the offender] being less likely to have a valuable sexual desire.”<sup>79</sup>

However, it could be objected that there is a plausible analysis of why the explorers, in the cave example, cannot dismiss the death of their companion as a side effect, which is not analogous to the way Birks and Buyx analyze the testosterone example. In the cave example there is an *intentional action*—lighting the dynamite—which *causes* an *intended effect*—someone being blown to bits.<sup>80</sup> Their motive for causing him to be blown to bits is to escape, and their companion being blown to bits is *intended as a means* of escaping.<sup>81</sup> The surviving explorers try to claim that lighting the

---

74. *See id.* at 138-39.

75. *Id.* at 138-39 (citing PHILIPPA FOOT, VIRTUE AND VICES AND OTHER ESSAYS IN MORAL PHILOSOPHY 21-22 (D. Z. Phillips ed., 1978)).

76. *Id.*

77. *Id.* at 21.

78. *See id.* at 21-22.

79. Birks & Buyx, *supra* note 2, at 140.

80. *See* FOOT, *supra* note 75, at 21-22.

81. *Id.*

dynamite has a *side effect*—the death of their companion—that is separate from the intended effect.<sup>82</sup> However, the surviving explorers' argument fails because the intended effect—their companion being blown to bits—is constitutive of the alleged side effect—the death of their companion—so his death is considered an intended effect.<sup>83</sup>

In contrast, in the testosterone example, there is an *intentional action*—administering the testosterone-lowering drug—which has the *intended effect* of diminishment of illegitimate sexual desires.<sup>84</sup> The motive for administering the drug is to prevent re-offending, and reducing illegitimate sexual desires is the means used to achieve this.<sup>85</sup> Those administering the drug claim that administering it has a *side effect*—diminishment of legitimate sexual desires—which is separate from the intended effect—diminishment of illegitimate sexual desires.<sup>86</sup> For the next move to be parallel to my analysis of the cave example, one would need to claim that the *intended effect*—diminishment of illegitimate sexual desires—is constitutive of the *alleged side effect*—diminishment of legitimate sexual desires.<sup>87</sup> But that claim is false. The illegitimate sexual desire for example, the desire to have nonconsensual sex with a minor, is a separate thing from the legitimate desire, to have consensual sex with one's adult partner.

Birks and Buyx claim that (1) lowering testosterone by administering the drug is constitutive of (2) being less likely to have a valuable desire.<sup>88</sup> They claim that these two states of affairs are analogous to the two states of affairs that have a constitutive relationship in the cave case: (1) blowing him up and (2) killing him.<sup>89</sup> However, on my analysis, the analogue of lowering testosterone by administering the drug is, in the cave case, the intentional action of lighting the dynamite, which is not constitutive of the

---

82. *See id.*

83. *See id.*

84. *See* Birks & Buyx, *supra* note 2, at 139-40.

85. *Id.*

86. *Id.* at 135.

87. *See supra* notes 70-71 and accompanying text (discussing the intended and unintended effects of testosterone-lowering drugs).

88. *See* Birks & Buyx, *supra* note 2, at 139-40.

89. *Id.* at 138-40.



death of the explorer—it is purely the *cause* of his death.<sup>90</sup> My analysis therefore implies that the relationship between (1) lowering the offender's testosterone by administering the drug *causes* (2) the effect that the offender is less likely to have a valuable desire, and (1) does *not constitute* (2) the effect. On my analysis of the cave case, the two states of affairs that have a *constitutive* (rather than purely causal) relationship in the cave case are (1) the explorer being blown to bits and (2) his death. The reason why the explorers cannot deny intending to kill the trapped explorer is that they intended to do an act—lighting the dynamite—in order to cause an intended effect—their companion being blown to bits—that is constitutive of his death; this is the same as intending to kill or intending to cause death.

Although it is not within the scope of this Article to conclusively establish that my analysis is correct,<sup>91</sup> I will provide some reasons for preferring my analysis to Birks and Buyx's. Firstly, Birks and Buyx's analysis has the highly counterintuitive implication that medicines never have known side effects, only intended effects.<sup>92</sup> They acknowledge that:

[O]ne implication of this account is that whenever a drug is administered to a patient with the knowledge that it is constitutive of the state of affairs of the patient experiencing side effects, the side effects cannot be unintended. It is only if the doctor is unaware that the drug has a particular side effect that it could not be intended.<sup>93</sup>

However, it is counterintuitive to suggest that, for example, when a doctor prescribes an antihistamine to relieve itchiness, which has the side effect of causing drowsiness, the doctor intends to make the patient drowsy. On my account, drowsiness and relief from itchiness are two separate effects, which are *not constitutively* related to each other (unlike being blown up and dying), and which are *causally*

---

90. See FOOT, *supra* note 75, at 21-22.

91. This would require engaging with the vast literature on intentions and side effects and with debates about the different ways in which actions, and their relationship to their effects, can be described.

92. See Birks & Buyx, *supra* note 2, at 135.

93. *Id.* at 141 n.34.

related to the administration of the drug. However, relief from itchiness is an intended effect, whereas drowsiness is an unintended side effect. Few doctors would accept the idea that when they go into work every day they are constantly forming *intentions* to cause their patients nausea, dizziness, disability, pain, et cetera. The intuition that medicines can have known but unintended side effects seems at least as powerful as the intuitions about the cave example upon which Birks and Buyx rely heavily, as they appeal to such intuitions in lieu of providing complete definitions of the concepts of “constitutive” and “causal.”<sup>94</sup> So, if the analysis of intentions and side effects should account for powerful intuitions, rather than implying that these intuitions should be discarded, then this provides some reasons for preferring my account to Birks and Buyx’s. My analysis is consistent with strong intuitions about the cave case (that the explorers cannot intend to blow their companion to bits without intending to kill him) and does not require us to discard the powerful intuition that medicines can have known unintended side effects.

This Article will now turn to the second part of Birks and Buyx’s argument—that *prison’s* supposed side effects (such as deterioration in mental health) genuinely are side effects.<sup>95</sup> Giving the example of restlessness, they claim that restricting someone’s freedom by sending the offender to prison does not constitute negative mental effects, but instead causes these unintended side effects.<sup>96</sup> However, the logic underlying their claim that medicines’ supposed side effects are actually intended seems to imply that prison’s side effects are intended as well.<sup>97</sup> Birks and Buyx’s account stresses how *knowledge* of the relationship between states of affairs can transform a side effect into an intended effect.<sup>98</sup> They do not explain exactly why it is obvious to most people that blowing someone up constitutes killing them, whereas it is not obvious to most people—at least initially—that medicines cannot have known side effects.<sup>99</sup> Seemingly implicit in their view is an explanation that while everyone knows that one cannot survive being blown to bits, lay

---

94. *See id.* at 138-39.

95. *See id.* at 134, 140.

96. *Id.* at 140.

97. *See id.*

98. *See id.*

99. *See id.* at 134-38.

people do not initially know about the mechanisms through which drugs have their effects.<sup>100</sup> Once people know more about the brain and fully accept propositions like “you cannot reduce testosterone to a score of twenty without diminishing a valuable sexual desire,” they will realize that administering a drug that reduces an individual’s testosterone to twenty constitutes diminishing the valuable sexual desire.<sup>101</sup>

The difficulty with using the argument above to try to distinguish neurointerventions from prison is that the same logic implies that the more we know about the neurological and psychological effects of prison, the more those effects will seem intended. For example, consider the following findings:

In neuropsychology, the functions that are at the base of self-regulation are called executive functions, such as working memory, sustained attention, impulse control, planning and cognitive flexibility. The prefrontal cortex of the brain is crucial for self-regulation and executive functions.... The prefrontal cortex and, thus, self-regulation and executive functions are sensitive to environmental influences.... A recent study ... found a decline in executive functions after only 3 months of imprisonment. More specifically, the participants, tested within 1 week after arrival, and retested after 3 months of imprisonment, performed worse on neuropsychological tasks measuring attention and impulse control. It is hypothesized that this decline in executive functions is due to the impoverished prison environment.<sup>102</sup>

The more brain science progresses and the more studies of this kind are conducted, the clearer it may become that we cannot keep people—or at least people with certain mental conditions or prior brain states—in prison without some deterioration in brain functions and in the mental capacities involved in self-regulation.<sup>103</sup> Following Birks and Buyx’s logic, this makes imprisonment and the deterioration in mental capacity and brain function look more like a constitutive relationship than a causal one, which would preclude

---

100. *See id.*

101. *See id.* at 139-40.

102. Lighthart et al., *supra* note 15, at 292-93 (footnotes omitted).

103. *See id.*

the state from arguing that the adverse impact of prison on mental integrity is merely an unintended side effect.<sup>104</sup> The logic of Birks and Buyx's own argument undermines their conclusion that we can use the side effect/intention distinction to show that prison's adverse impact on mental integrity is easier to justify than neurointervention's adverse impact on mental integrity.<sup>105</sup>

James Lim makes a similar point from a more sociological perspective, arguing:

[O]ur self-worth, identity, rationality, and capacities such as autonomy ... depend on our status as social beings. Some have even argued that our capacities do not just depend on our relationships, but that they are "*constitutively* social." As Taylor argues, "Something only offends my dignity because it upsets or challenges the way I present, project or express myself in this public space." If this is true, then incarceration, which constitutively deprives individuals of particularly valuable social interactions, and designates for them a certain social status as untrustworthy and unworthy of participating in society, also constitutes a harm to an individual's self-conception. Arguably, it is equivalent to interferences on our mental capacities.<sup>106</sup>

It is submitted that this logic goes some way toward showing that the adverse impact of prison on mental integrity is *more like* the adverse impact of neurointerventions on mental integrity than had traditionally been assumed.<sup>107</sup> However, it is possible to resist the conclusion that the mental-integrity-undermining impact of prison is *intended*. Just as it was argued above that administering a drug that reduces testosterone *causes*, but does not constitute, the diminishment of a valuable desire,<sup>108</sup> sending someone to prison *causes*, but does not constitute, impairment of brain functions. Knowing a lot about the strength of the causal connection between an action and an effect does not, by itself, make the action intended. If it did, then well-informed doctors would intend their patients to suffer the known adverse effects of medication and there could be no

---

104. See Birks & Buyx, *supra* note 2, at 139-41.

105. See *id.*

106. Lim, *supra* note 16, at 169 (citations omitted).

107. See *id.*

108. See *supra* notes 84-87 and accompanying text.

known but unintended side effects.<sup>109</sup> It is plausible that doctors would find this counterintuitive, which suggests that the intuition that medicine can have known, unintended side effects does not just rest on lay people's ignorance of drugs' causal mechanisms. So, even if a judge became very familiar with the adverse impact of prison on offenders' mental capacity and brain functions, this would not show that the judge *intends* offenders' mental capacity and brain functions to deteriorate when the judge sentences the offender to prison.

Having said this, it is questionable whether the intention/side effect distinction has the moral weight in this context that many have assigned to it.<sup>110</sup> This Article will consider two alternative reasons for this. Firstly, intended effects and side effects may lie on a spectrum, and the closer the mental-integrity-undermining effects of a punishment or neurointervention fall toward the intended end of the spectrum, the harder the punishment or neurointervention may be to justify.<sup>111</sup> Knowledge that one's action always has a particular effect, or is virtually certain to have that effect,<sup>112</sup> may be one

---

109. See Birks & Buyx, *supra* note 2, at 135.

110. See, e.g., *id.* at 134, 137.

111. The idea that intentions and side effects lie on a spectrum has been put forward by William J. Fitzpatrick, *The Intend/Foresee Distinction and the Problem of "Closeness,"* 128 PHIL. STUD. 585, 606-07 (2006); Timothy Chappell, *Two Distinctions that Do Make a Difference: The Action/Omission Distinction and the Principle of Double Effect,* 77 PHIL. 211, 214-15 (2002). However, they do not analyze the spectrum in the way outlined in the accompanying text.

112. See generally Victor Tadros, *The Homicide Ladder,* 69 MOD. L. REV. 601, 604 (2006) (providing a general definition of a virtually certain effect in the criminal context). Knowledge that an effect is virtually certain to occur is sometimes referred to as "oblique intention" in criminal law theory, and it is sometimes suggested that this is a type of intention. A. P. Simester, *Moral Certainty and the Boundaries of Intention,* 16 OXFORD J. LEGAL STUD. 445, 452-53 (1996). In English law, in the leading murder case of *R. v. Woollin,* the court held that where the accused foresaw that the effect (death or serious bodily harm) was virtually certain to occur, the jury was entitled to find intention. [1999] 1 AC 82 (HL). In the Scottish murder case of *Petto v. HM Advocate,* per Lord Justice Clerk, foresight with virtual certainty was "equiparated with" intention. [2011] HCJAC [13]. There are convincing counterexamples to the idea that knowledge with virtual certainty is the same as intention. For instance, a father who drops his baby from a burning skyscraper because it is the only possibility of saving the baby's life does not intend to kill the baby even though the baby's death is virtually certain to occur. JONATHAN HERRING, *CRIMINAL LAW* 168 (8th ed. 2018). This counterexample also suggests that knowledge with virtual certainty *by itself* cannot push an effect toward the intended end of the spectrum, which is one reason why I have suggested that the position in the spectrum may be determined by a range of factors. The last two factors listed in the text, which I have suggested are indicative of near-intention, do not apply to the father.

factor among others that helps to place the effect closer to the intended end of the spectrum, though it is still unintended, strictly speaking. Another factor that may help to place the effect closer to the intended end of the spectrum could include whether it occurs at the same time as the intended effect.<sup>113</sup> In the case of incarceration, the intended effect of confinement in the under-stimulating environment of prison occurs simultaneously with the deterioration of the offender's mental capacity and brain functions, which is unintended but perhaps more similar to an intended effect than effects that are more remote in time.<sup>114</sup> The fact that the explorer exploding occurred simultaneously with his death may be one factor leading us to conclude that both of those things are intended. Another factor that may place the effect toward the intended end of the spectrum is the presence of alternative courses of action—which would not have caused the effect (or not to the same extent)—that the actor refused to take.<sup>115</sup> A final factor may be the actor's failure to take available steps to mitigate the adverse impact of the effect.<sup>116</sup>

A second, alternative, and probably preferable line of argument is that in some situations the actor is not entitled to rely on the fact that the adverse effect of his action was an unintended side effect and this fact does not significantly diminish the wrongfulness of the action. The last two factors mentioned in the previous paragraph are particularly relevant here.<sup>117</sup> If a doctor, at a patient's request, gives the patient a pain-relief drug in order to ease the patient's suffering, but the doctor foresees that the drug will hasten death, then it is often argued that death is an unintended side effect and the doctor can rely on this fact when denying that he is guilty of any

---

113. See generally Mikkel C. Vinding et al., *The Time Between Intention and Action Affects the Experience of Action*, 9 FRONTIERS HUM. NEUROSCIENCE 1 (2015) (hypothesizing that the delay between the decision to act and the action renders the consequences less intentional).

114. See Lighthart et al., *supra* note 15, at 288, 292-93.

115. See Chappell, *supra* note 111, at 214-25.

116. See generally FOOT, *supra* note 75, at 23 (considering the morality of a trolley driver's attempt to mitigate the adverse impact of the inevitable crash by diverting from a track with five potential victims to a track with one potential victim). This factor is problematic because it seems unlikely that something someone does *after* the intentional action that causes the effect could alter the extent to which the effect was intended. This is one reason why the second, alternative explanation given in the text seems preferable.

117. See *supra* notes 115-16 and accompanying text.



crime.<sup>118</sup> However, now imagine that the doctor also had an alternative, effective pain-relief drug that did not hasten death and the doctor intentionally refused to use it. In the latter case, even if hastening death were still an unintended side effect of the doctor's chosen drug, then the doctor should not be able to rely on this fact, as this fact would not significantly diminish the wrongfulness of administering the drug.<sup>119</sup> Or imagine that only one pain-relief drug is available that would have the unintended side effect of hastening death if an antidote were not also administered, and the doctor decides when administering the death-causing drug not to administer the antidote. Again, the doctor should not be able to rely on the fact that the patient's death was an unintended side effect of the drug the doctor administered, as this would not diminish the wrongfulness of his action.<sup>120</sup>

These considerations cast some doubt on whether the intention/side effect distinction has the moral weight in this context that some have previously assumed.<sup>121</sup> Even if the integrity-undermining effects of incarceration are unintended (or less intended) than the integrity-undermining effects of neurointerventions, it does not follow that imposing mandatory incarceration on offenders is compatible with the right to mental integrity. As suggested in the next Part, there may sometimes be alternative policies to mandatory incarceration which have less of an adverse impact on mental integrity.<sup>122</sup> If the state refuses to adopt such policies, then the claim that undermining mental integrity is a mere side effect of imprisonment does not necessarily weaken or remove the mental-integrity-based objection against mandatory incarceration.

---

118. See, e.g., Henry Palmer, *Dr. Adams' Trial for Murder*, in *THE CRIMINAL LAW REVIEW* 365, 375-76 (1957).

119. *Cf. id.* This is not meant to exclude the possibility that there might be other considerations in such a case that meant that the doctor's action was not wrong.

120. *Cf. id.* One might object that, in this case, administering the drug is not wrongful—the omission to mitigate the effect of the drug is wrongful. In response, it seems plausible that if he intended at the time of administering the pain-relief drug not to administer the antidote, then administering the death-causing drug was part of his wrongful course of conduct. The course of conduct of this doctor (Dr. A) is somewhat more wrongful than the conduct of Dr. B, who simply fails to give the antidote to a patient who has already been given the death-causing drug by Dr. C.

121. See, e.g., Birks & Buyx, *supra* note 2, at 134, 137.

122. See, e.g., *infra* Part III.

### III. PRACTICAL IMPLICATIONS

This Part will discuss some of the practical implications of the arguments I have set out in the previous Part. If we regard the adverse effects of prison and neurointerventions as roughly morally equivalent where their effects are similarly severe, one way to respond is to be equally willing to impose either and allow the state to impose whichever is most effective at preventing crime. In contrast, this Article will favor a second, human-rights-based approach, which aims to respect the offender's right to mental integrity.

In some situations, there may be alternative policies to mandatory incarceration that would have less of an adverse impact on mental integrity. Where forms of incarceration and neurointerventions are available that are similarly effective and humane, offenders should be offered the choice between these options. This policy could potentially reduce the extent to which both incarceration and neurointerventions infringe the right to mental integrity. As explained in more detail below, there are at least two possible ways in which an intervention can infringe the right to mental integrity: (1) by undermining a valuable aspect of the offender's psychological makeup, or (2) by expressing disrespect for the offender's valuable mental capacities.<sup>123</sup>

Offering offenders the choice between the neurointervention and incarceration could—at least in some situations—avoid or minimize both types of infringement. Regarding valuable aspects of the offender's psychological make-up, different interventions/punishments may affect the offender's psychology in different ways, and each method may differ in the extent to which it adversely affects various aspects of the offender's psychology.<sup>124</sup> The question of which intervention affects mental integrity to the greatest extent may depend on which conception of mental integrity is adopted. Reasonable

---

123. See, e.g., Lisa Forsberg, *Anti-Libidinal Interventions and Human Rights*, 21 HUM. RTS. L. REV. 384, 402 (2021).

124. See generally Danielle Wallace & Xia Wang, *Does In-Prison Physical and Mental Health Impact Recidivism?*, 11 SS M-POPULATION HEALTH 1 (2020) (discussing the mental impact of various prison conditions on different populations).



people may disagree, for example, about whether the social conception of self-respect that Lim described,<sup>125</sup> which can be affected by incarceration, is more valuable to them than aspects of their psychology that might be affected by neurointerventions.<sup>126</sup> In such a situation, it is undesirable for the state to attempt to decide on the offender's behalf which aspect of—or conception of—the offender's mental integrity is most valuable.<sup>127</sup> Rather, offenders should be offered the choice between the different interventions, as this takes account of what is *valuable to the particular offender*.<sup>128</sup> For example, the offender might highly value his legitimate sexual desires that might be weakened through the side effects of neurointerventions; or the offender might value the ability to reform himself by engaging in practical deliberation about why he should not act on his preexisting illegitimate desires rather than having those desires directly weakened by neurointerventions.

Giving offenders some choice between the sentencing options could reduce the likelihood that the aspects of their psychology that offenders themselves most value will be impacted. Secondly, giving the offender the choice between different humane sentencing options can show respect for the offender's valuable capacities by recognizing that the offenders have the valuable mental capacities that enable them to make informed choices about such issues.<sup>129</sup> Furthermore, it is possible to mitigate some of the adverse effects of incarceration on mental integrity.<sup>130</sup> For example, Sjouis Ligthart and others point out that enriching the prison environment by

---

125. See *supra* note 106 and accompanying text.

126. See Lim, *supra* note 16.

127. See Birks & Buyx, *supra* note 2, at 134.

128. See, e.g., Geraldine Sealey, *Some Sex Offenders Opt for Castration*, ABC NEWS (Jan. 7, 2006, 10:08 AM), <https://abcnews.go.com/US/story?id=93947&page=1> [<https://perma.cc/6XNS-AJ8V>].

129. See generally Farah Focquaert et al., *Offering Neurointerventions to Offenders with Cognitive-Emotional Impairments: Ethical and Criminal Justice Aspects*, in *NEUROINTERVENTIONS AND THE LAW* (Nicole A. Vincent et al. eds. 2020) (arguing that, under certain circumstances, granting offenders the choice of neurointervention increases the likelihood of successful therapy outcomes). If the sentencing options involve excessive hardship or are wrongful in some other way, then offering the choice per se will not necessarily make the situation morally better or more respectful. See *id.* For example, giving offenders the choice between waterboarding or thumbscrews would not make these into respectful forms of punishment.

130. See Ligthart et al., *supra* note 15, at 296-97.

providing more opportunities for exercise and education can, to some extent, counteract the damaging neurological effects of understimulation.<sup>131</sup> These considerations suggest that the state has an obligation to research ways of minimizing the adverse effects of both prison and neurointerventions on mental integrity and mitigate these adverse effects where possible.<sup>132</sup>

In light of the above philosophical discussion, the remainder of this Article will make some suggestions about what the most ethically defensible interpretation and development of certain aspects of the legal right against degrading treatment<sup>133</sup> would be.<sup>134</sup> This Article will focus on the right against degrading treatment because this legal right includes aspects of the moral right to mental integrity that are particularly relevant in the context of criminal rehabilitation.<sup>135</sup> Specifically, according to comments in recent ECHR case law discussed below, Article 3 includes respect for, and protection of, the mental capacities required for criminals to reform themselves, to be rehabilitated, and to reintegrate into society.<sup>136</sup>

---

131. *See id.*

132. The idea of giving offenders the choice between neurointerventions and prison has been defended before, but based on rationales that differ in important respects from the arguments presented in this Article. *See, e.g.,* Focquaert et al., *supra* note 129; Farah Focquaert & Adrian Raine, *Ethics of Community-Based Sanctions*, in *ENCYCLOPEDIA OF COMMUNITY CORRECTIONS* 144, 144-48 (Shannon M. Barton-Bellessa ed., 2012); Elizabeth Shaw, *Cognitive Enhancement and Criminal Behavior*, in *COGNITIVE ENHANCEMENT: AN INTERDISCIPLINARY PERSPECTIVE* 265, 267-69 (Elisabeth Hildt & Andreas G. Franke eds., 2013).

133. This Article will focus on the “degrading treatment” aspect of Article 3, because the suffering caused by under-stimulating prison conditions or certain mandatory neurointerventions may not cause the level of suffering required for “torture” or “inhuman treatment.” *See Aydin v. Turkey*, 1997-VI Eur. Ct. H.R. 1866; JOHN COOPER, *CRUELTY: AN ANALYSIS OF ARTICLE 3* (2003).

134. Aspects of the moral right to mental integrity are relevant to several legal rights that are protected by the European Convention on Human Rights (ECHR); for example, Article 8 protects the right to privacy, Article 9 protects the right to freedom of thought, conscience, and religion, and Article 3 prohibits torture and inhuman or degrading treatment and punishment. *See* FREDE CASTBERG, *THE EUROPEAN CONVENTION ON HUMAN RIGHTS* 83, 138, 146 (Torkel Opsahl & Thomas Ouchterlony eds., 1974). On the connection between these rights and mental integrity, see, for example, PAUL TIEDEMANN, *PHILOSOPHICAL FOUNDATIONS OF HUMAN RIGHTS* 139-55 (2020); Sjors L.T.J. Ligthart, *Coercive Neuroimaging, Criminal Law, and Privacy: A European Perspective*, 6 *J. L. & BIOSCIENCES* 289, 293-309 (2019); Bublitz & Merkel, *supra* note 42, at 51-60.

135. *See* CASTBERG, *supra* note 134, at 83.

136. *See id.*

The right against degrading treatment is closely connected with the idea of respect for human dignity,<sup>137</sup> which echoes the concern in the literature on mental integrity to show respect for rational agency.<sup>138</sup> However, the ECHR case law, unlike the literature on the moral right to mental integrity, focuses primarily on experiential suffering.<sup>139</sup> Degrading treatment or punishment has been described as treatment which arouses “feelings of fear, anguish and inferiority” capable of “humiliating” those subjected to it and of “breaking their physical and moral resistance.”<sup>140</sup> It is unclear from the case law the extent to which experiential suffering is an *essential* component of Article 3, or whether disrespect for human dignity in the absence of experiential suffering could amount to degrading treatment. This Article proposes that the legal interpretation of degrading treatment should include disrespect for rational agency—even in the absence of experiential suffering—as this would provide better legal protection for the moral right to mental integrity.<sup>141</sup>

It is also worth noting that Article 3 has the potential to provide especially effective support for the right to mental integrity because Article 3 is an absolute right.<sup>142</sup> Once an absolute right is engaged, states cannot justify restricting it in order to protect the public interest or the rights of others,<sup>143</sup> nor can signatories to the ECHR

---

137. See, e.g., *Svinarenko and Slyadnev v. Russia* 2014-V Eur. Ct. H.R. 181; *Murray v. Netherlands*, App. No. 10511/10 (Apr. 26, 2016), <http://hudoc.echr.coe.int/eng?i=001-162614> [<https://perma.cc/H8WN-TYUK>].

138. See, e.g., *Bublitz & Merkel*, *supra* note 42, at 53-55.

139. See, e.g., *Sviarenko*, 2014 Eur. Ct. H.R. at 183; *Murray*, 2016 Eur. Ct. H.R. at ¶ 68-69.

140. See *V v. United Kingdom*, 1999-1X Eur. Ct. H.R. 111 ¶ 71; *Peers v. Greece*, 2001-III Eur. Ct. H.R. 275 ¶¶ 74-75.

141. See *infra* notes 150-53 and accompanying text.

142. See Michael K. Addo & Nicholas Grief, *Does Article 3 of the European Convention on Human Rights Enshrine Absolute Rights?*, 9 EUR. J. INT'L L. 510, 523 (1998) (“[I]t is now generally accepted that Article 3 ... guarantees absolute rights.”).

143. See *id.* at 513. However, it is debatable whether the ECHR’s interpretation of Article 3 is really consistent with this notion of an absolute right or whether ECHR has taken into account public interest considerations after all. See, e.g., Natasa Mavronicola, *Crime, Punishment and Article 3 ECHR: Puzzles and Prospects of Applying an Absolute Right in a Penal Context*, 5 HUM. RTS. L. REV. 721, 733-39 (2015); Mark D. Kielsingard & John Khatini Vinod, *Trending Toward Precaution at What Cost? Reconciling Positive and Negative Human Rights Obligations in the Use of Neuro Intervention for Sex Offenders*, 18 CONN. PUB. INT. L.J. 393, 400-38 (2019).

suspend, derogate, or depart from Article 3, even in times of war or national emergencies.<sup>144</sup>

In a series of cases on Article 3—in the context of life sentences—the European Court of Human Rights (ECtHR) has increasingly emphasized the importance of rehabilitating offenders.<sup>145</sup> Rehabilitation is defined as facilitating offenders’ reintegration into society with the aim of, for example, social protection and crime prevention.<sup>146</sup> The case law on the implications of Article 3 for criminal rehabilitation is ambiguous and still developing.<sup>147</sup> States are under a positive obligation to facilitate offenders’ rehabilitation.<sup>148</sup>

The ECtHR has also commented on the State’s negative obligation to not undermine the capacities required for rehabilitation.<sup>149</sup> For example, in the context of solitary confinement, the ECtHR noted that lack of “appropriate mental and physical stimulation are likely, in the long term, to have damaging effects, resulting in a deterioration of mental faculties and social abilities.”<sup>150</sup> Some comments suggest that the capacity for positive change or reform are particularly worthy of protection and central to our common humanity.<sup>151</sup> In the case of *Vinter and Others v. United Kingdom*, Judge Power-Forde, in her concurring judgment, stated: “Those who commit the most abhorrent and egregious of acts and who inflict untold suffering upon others, nevertheless retain their fundamental humanity and carry within themselves the capacity to change.”<sup>152</sup> The value of the offender’s capacity for reform through exercising

144. CASTBERG, *supra* note 134, at 165.

145. *See, e.g.*, *Vinter and Others v. United Kingdom*, 2013-III Eur. Ct. H.R. 317 (2013); *Khoroshenko v. Russia*, 2015-IV Eur. Ct. H.R. 329; *Hutchinson v. United Kingdom*, App. No. 57592/08 (Jan. 17, 2017), <http://hudoc.echr.coe.int/eng?i=001-170347> [<https://perma.cc/P4H2-DN3N>].

146. *See Murray v. Netherlands*, App. No. 10511/10 (Apr. 26, 2016), <http://hudoc.echr.coe.int/eng?i=001-162614> [<https://perma.cc/H8WN-TYUK>].

147. *See, e.g.*, *Vinter*, 2013-III Eur. Ct. H.R., at 319-21.

148. *See Murray*, App. No. 10511/10, ¶104.

149. *See id.*

150. *See Piechowicz v. Poland*, App. No. 20071/07, ¶ 173 (Apr. 17, 2012), <http://hudoc.echr.coe.int/eng?i=001-110499> [<https://perma.cc/R3S3-HL9P>].

151. *See, e.g.*, *Why Promote Prison Reform?*, UNITED NATIONS OFF. ON DRUGS & CRIME, <https://www.unodc.org/unodc/en/justice-and-prison-reform/prison-reform-and-alternatives-to-imprisonment.html> [<https://perma.cc/Z7HY-6UG8>].

152. *See Vinter*, 2013-III Eur. Ct. H.R., at 358 (Power-Forde, J., concurring).

rational agency is also a theme discussed in the literature on neurointerventions and the moral right to mental integrity.<sup>153</sup>

Ligthart and others persuasively argued that, in light of neuroscientific findings on the negative neuropsychological effects of under-stimulation, the ECtHR should recognize a wider range of restrictive prison conditions as being in breach of Article 3 because under-stimulation can undermine the mental capacities required for reform.<sup>154</sup> The scope of Article 3's prohibitions should be expanded to include both restrictive prison conditions, as Ligthart and others proposed, and the undermining of mental capacities—even in the absence of mental suffering.<sup>155</sup>

The ECtHR has not applied the comments about the value of the offender's "capacity to change" to mandatory neurointerventions.<sup>156</sup> The ECtHR's discussion of the compulsory administration of brain-altering interventions to offenders has focused on the concept of "medical necessity" rather than penological goals.<sup>157</sup> It is submitted that the court should recognize the value of the capacity to reform in the context of compulsory brain interventions as well.<sup>158</sup>

Domestic laws permit these interventions to be given to rational offenders who retain the capacity to give or withhold consent.<sup>159</sup> Therefore, the recipients still presumably "carry within themselves the capacity to change."<sup>160</sup> The assertion that a particular treatment is a medical necessity cannot deprive the requirement to respect the

---

153. See, e.g., Douglas, *supra* note 1, at 112-16; Shaw, *supra* note 1, at 2, 12-18.

154. See Ligthart et al., *supra* note 15, at 299.

155. See *id.* at 299-300.

156. See Forsberg, *supra* note 123, at 391-86.

157. See *id.* at 394.

158. See generally *id.* at 394-96 ("Article 3 does not allow for any type of proportionality of necessity analysis that weighs how necessary a (degrading/inhuman) punishment is to achieve a public aim" like rehabilitation; therefore, capacity to ECtHR's current approach does not permit for consideration of an individual inmate's capacity to reform.)

159. See, e.g., Mental Health Act 1983, c. 20 § 58 (U.K.). For example, in England and Wales, criminal courts can sentence mentally disordered offenders to detention in hospital, where they may be given compulsory interventions, even if they have the capacity to refuse consent. *Id.* § 36. Section 63 permits nonconsensual interventions except psychosurgery, hormone implants, electro-convulsive therapy, and psychiatric medications lasting more than three months. *Id.* § 63. Section 58 permits nonconsensual psychiatric medications lasting more than three months, provided that the requirements concerning approval by two medical practitioners are met. *Id.* § 58.

160. See *Vinter and Others v. United Kingdom*, 2013-III Eur. Ct. H.R., 317, 355 (Power-Forde, J., concurring).

offender's capacity for reform of its moral weight.<sup>161</sup> Lisa Forsberg has argued that the "medical necessity" requirement can easily be met.<sup>162</sup> Arguably, even "treatments," whose main goal is to protect society by preventing reoffending, and which are imposed on rational offenders, may be deemed "medically necessary."<sup>163</sup> If the main goal of the neurointervention—social protection—is the same as the goal of incarceration, and the ECtHR has deemed that the latter cannot justify degrading treatment under Article 3, then the goals of the neurointervention cannot justify such treatment either.<sup>164</sup>

Under restrictive conditions, prison can fail to respect offenders' capacity for reform and rehabilitation by actually *undermining* this capacity.<sup>165</sup> In contrast, mandatory neurointerventions would fail to respect the capacity for reform and rehabilitation by denying offenders the capacity to reform themselves through exercising their own agency.<sup>166</sup> As previously argued, directly reshaping the offender's character or moral motivations "implies that offenders are radically defective with regard to one of the most fundamental aspects of their agency."<sup>167</sup> Forcibly reshaping the individual in this way represents him as having less value than he should be accorded.<sup>168</sup> This goes beyond the message that is an essential part of punishment—"that the offender committed a wrong on a particular occasion."<sup>169</sup>

Gabriel De Marco and Thomas Douglas have replied to my argument by claiming that there may be situations, such as with repeat offenders, when the state has evidence that the offender lacks or is seriously deficient in the higher order rational capacity

---

161. *See generally id.* ("[D]eserved though [the punishment] may be ... [prisoners] ought not to be deprived entirely of [the hope to atone]. To deny them the experience of hope would be to deny a fundamental aspect of their humanity.").

162. *See* Forsberg, *supra* note 123, at 391-96.

163. *See* Dvoracek v. Slovakia, App. No. 30754/04 (July 28, 2009), <http://hudoc.echr.coe.int/eng?i=001-93768> [<https://perma.cc/U9XS-9TMT>]; Kielsingard & Vinod, *supra* note 143, at 396-407; Elizabeth Shaw, *Retributivism and the Moral Enhancement of Criminals Through Brain Interventions*, 83 ROYAL INST. PHIL. SUPPLEMENT 251, 256-70 (2018).

164. *See* Shaw, *supra* note 163, at 259-61.

165. *See* Lighthart et al., *supra* note 15, at 299.

166. *See* Shaw, *supra* note 1, at 2, 12-18.

167. Shaw, *supra* note 1, at 13.

168. *Id.*

169. *Id.*



to correct one's failings and to reform.<sup>170</sup> If it is true that the offender lacks this rational capacity, then De Marco and Douglas argue it is not a negative misrepresentation.<sup>171</sup>

In reply, firstly, even repeat offending does not prove that the offender lacks the capacity to reform.<sup>172</sup> By convicting the offender rather than acquitting him on the basis of insanity, the state assumed that he is a rational agent.<sup>173</sup> Denying that he has the rational capacities required to change his behavior seems to contradict this assumption.<sup>174</sup> Secondly, sending the message that offenders—or a certain class of offenders—lack this capacity would negatively misrepresent these offenders' *value* even if it were true that they lacked this capacity.<sup>175</sup> People would still be valuable and worthy of moral consideration even if they lacked this capacity. Nevertheless, it is plausible that the capacity for moral improvement has come to symbolize an agent's value and our common humanity, probably because most agents share the capacity, value it, and would be outraged if they were deprived of it.<sup>176</sup> Highlighting that a class of offenders lacks this fundamental capacity is exclusionary and represents them as inferior.<sup>177</sup> Respect can be shown by celebrating valuable capacities when agents possess them (for an analogy, consider celebrating peoples' intellectual achievements) and disrespect can be shown by emphasizing the lack of a capacity (for an analogy, consider publicly declaring that certain individuals

---

170. See Gabriel De Marco & Thomas Douglas, *The Expressivist Objection to Nonconsensual Neurocorrectives*, CRIM. L. & PHIL. (2021), <https://link.springer.com/article/10.1007%2Fs11572-021-09566-9> [<https://perma.cc/G64H-KSCJ>].

171. See *id.*

172. See PENAL REFORM INT'L, ALTERNATIVES TO THE DEATH PENALTY INFORMATION PACK 7 (2015).

173. See, e.g., T.V. Asokan, *The Insanity Defense: Related Issues*, 58 INDIAN J. PSYCHIATRY 191 (2016).

174. See *id.*

175. See generally Eric Lewis, *Crime, Fear, and the Republicans*, MARSHALL PROJECT (Sept. 15, 2015, 7:15 AM), <https://www.themarshallproject.org/2015/09/15/crime-fear-and-the-republicans> [<https://perma.cc/227A-MJ7Q>] (discussing generally how government's treatment of prisoners affects public opinion of prisoners).

176. See, e.g., *Vinter and Others v. United Kingdom*, 2013-III Eur. Ct. H.R., 317, 358 (Power-Forde, J., concurring).

177. See *id.* ("To deny them the experience of hope [of atonement] would be to deny a fundamental aspect of their humanity and, to do that, would be degrading.").

performed poorly on some measure of intelligence) regardless of whether that were true.

It might be objected that, if it were true that certain offenders could not be rehabilitated without neurointerventions, and effective rehabilitative neurointerventions were available, these offenders should be compelled to receive them, even if this were disrespectful.<sup>178</sup> For example, Sebastian Holmen argues that the respect requirement could be outweighed by the value of protecting future victims and their families from harm.<sup>179</sup> In response, as noted above, even if it were true that some noninsane offenders lacked the capacity to be rehabilitated without neurointerventions, it is unlikely that the state could prove this.<sup>180</sup> Furthermore, *forcing* offenders to receive neurointerventions rather than offering neurointerventions as an option not only implies that offenders lack the rational capacities to be rehabilitated without neurointerventions—it also implies that they lack the rational capacity to assess whether there are good reasons for accepting the option of neurointerventions.<sup>181</sup> It would be even harder for the state to prove that noninsane offenders also lacked this capacity.<sup>182</sup> No one is proposing that the state should simply let hard-to-rehabilitate offenders walk free, leaving potential victims unprotected. The importance of protecting potential victims can adequately be recognized by detaining offenders in conditions that are designed to facilitate their rehabilitation<sup>183</sup> and, where appropriate, providing offenders with the choice between neurointerventions and other forms of rehabilitation.<sup>184</sup>

My argument has some similarities with Christopher Bennett's account of "opacity respect"—the idea that it shows respect for the basic equality of all moral agents to refrain from investigating the extent to which an individual possesses certain fundamental agential capacities.<sup>185</sup> According to Bennett, provided the agent

---

178. See, e.g., Sebastian Jon Holmen, *Respect, Punishment and Mandatory Neurointerventions*, 14 *NEUROETHICS* 167 (2021).

179. *Id.*

180. See *supra* notes 173-74 and accompanying text.

181. See Shaw, *supra* note 2, at 103-05.

182. See *supra* notes 173-74 and accompanying text.

183. See Lighthart et al., *supra* note 15, at 296-97.

184. See Shaw, *supra* note 163, at 259.

185. See Christopher Bennett, *Intrusive Intervention and Opacity Respect*, in *TREATMENT FOR CRIME*, *supra* note 1, at 255, 261.



meets a certain threshold of rationality that all offenders who were convicted rather than found legally insane have met, opacity respect requires refraining from investigating variations above this threshold.<sup>186</sup> The agential capacities of an individual who has passed the minimum rationality threshold should be “opaque” in that the state should be barred from assessing the extent to which some of the individual’s rational capacities may be defective,<sup>187</sup> and the state should be barred from basing its response to the offender on any information it may have about the defectiveness of his agential capacities.<sup>188</sup>

However, there are some key differences between my argument and Bennett’s conception of opacity respect. My argument focuses specifically on the disrespect of implying that offenders lack, or are seriously defective with regard to, a particularly important aspect of rationality: the ability to reform oneself.<sup>189</sup> This disrespect can be shown by denying offenders the opportunity to reform themselves through the exercise of their rational agency and instead simply re-engineering their character or moral motivations through brain interventions.<sup>190</sup> In contrast, Bennett’s account of “opacity respect” seems to interpret respect for rationality in a much broader way—a breadth that opens his account up to the charge of overinclusiveness.<sup>191</sup> Opacity respect would be violated by investigating “the offender’s underlying psychology” as part of “any attempt to ask why the person acted as they did or how they could be prevented from acting in that way in the future” and by treating them in ways that are based on “factors other than the offender’s rational self-presentation.”<sup>192</sup> He writes:

When relating to another person as an equal, our default position is to take at face value the deliverances of the person’s thought and agency, and not to look at the “mess inside” from which it stemmed. We treat the person as an equal by dealing

---

186. *See id.*

187. *See Shaw, supra* note 163, at 256.

188. *See Bennett, supra* note 185, at 266, 269.

189. *See supra* note 55 and accompanying text.

190. *See Shaw, supra* note 163, at 256.

191. *See Bennett, supra* note 185.

192. *Id.* at 266.

with them in the image of a rational agent presenting herself to the world on the basis of rational deliberation. We take at face value this rational self-presentation, averting our gaze from its psychological sources.<sup>193</sup>

Bennett's account may be viewed as overinclusive, in that it is hard to see why it does not rule out forms of rehabilitation, such as talking therapies, that many people find acceptable.<sup>194</sup> Bennett's response to this criticism makes some valid and important points. However, his response seems to contradict his earlier statements, and also makes it less clear what work the concept of "opacity respect" is meant to be doing. Bennett suggests that standard "talking therapies" can be acceptable because they "often engage reason and challenge the offender to think differently."<sup>195</sup> This is true, but these therapies also assume that various ways in which the individual has been and may, in the absence of therapy, go on thinking have *not* been rational.<sup>196</sup> Many therapists invite individuals to consider whether the supposed rational justifications they present for their actions are, in fact, merely rationalizations masking nonrational factors that influence their actions.<sup>197</sup> By viewing such therapies as acceptable, Bennett seems to be abandoning the idea that respect requires us to take "at face value [the individual's] rational self-presentation."<sup>198</sup> Bennett then replies that neurointerventions can be distinguished from talking therapies because the former bypass rationality, whereas the latter engage rationality.<sup>199</sup> However, this makes it unclear what work the concept of "opacity respect" is doing.

On Bennett's account, there seem to be two separate requirements: (1) the requirement to engage with—rather than bypass—rational agency, and (2) the supposed requirement to abstain from investigating the agent's underlying psychology, which is at the

---

193. *Id.* at 265 (footnote omitted).

194. *See id.* at 271-72.

195. *Id.* at 272.

196. *See id.*

197. *See, e.g.,* Michael E. Addis & Kelly M. Carpenter, *The Treatment Rationale in Cognitive Behavior Therapy: Psychological Mechanisms and Clinical Guidelines*, 7 COGNITIVE & BEHAV. PRAC. 147, 148-50 (2000).

198. *See* Bennett, *supra* note 185, at 265.

199. *See id.* at 272.

heart of his conception of opacity respect.<sup>200</sup> If, when it comes to talking therapies, the second requirement (that is, opacity) can be dispensed with if the former (that is, engaging with rational agency) has been met, why do we need to rely on opacity to explain what is wrong about neurointerventions, rather than simply appealing to the wrongfulness of bypassing rational agency?

In contrast, on my account, the value of engaging with offenders' rational agency is not separate from the *specific* capacity that the state should be barred from denying. Engaging with offenders' rational agency reaffirms their possession of the specific higher order capacity to correct their rational and moral failings through exercising their agency.<sup>201</sup> In contrast, mandatory neurointerventions deny this specific capacity through bypassing agency in order to reshape offenders' character or moral motivations.<sup>202</sup> Rehabilitating an offender partly through *consensual* neurointerventions would not necessarily fail to show respect for this capacity because taking an autonomous decision to accept help, including help via neurointerventions, can in itself be a way of exercising the higher order ability to start on the route to reform.<sup>203</sup> So, giving an offender the opportunity to exercise this higher order ability by giving him choices about different routes to rehabilitation reaffirms his capacity to reform.<sup>204</sup>

## CONCLUSION

This Article examined the role of the right to mental integrity as a constraint on state use of neurointerventions and also discussed the extent to which traditional forms of punishment, such as incarceration, interfere with the right to mental integrity. This Article focused on an influential objection against neurointerventions—the idea that neurointerventions threaten mental integrity in a way that prison does not, because even if prison infringes mental integrity, that infringement is merely foreseen, whereas the threat

---

200. *See id.* at 259-61.

201. *See supra* note 48 and accompanying text.

202. *See supra* note 48 and accompanying text.

203. *See* Douglas, *supra* note 1, at 112-16; Shaw, *supra* note 1, at 2, 12-18.

204. *See supra* notes 152-53 and accompanying text.

that neurointerventions pose to mental integrity is intended. It attempted to cast some doubt on whether the intention/side effect distinction has the *moral weight* in this context that has often been assigned to it. It argued that the intention/side effect distinction does not warrant regarding mandatory incarceration as being significantly less problematic than mandatory neurointerventions in cases where the adverse effects of prison and neurointerventions are similarly severe. This Article discussed some of the practical implications of treating neurointerventions and prison in a roughly similar way and argued that the state has strong obligations to research techniques that minimize adverse so-called side effects, or any adverse effects on mental integrity, and to offer prisoners the choice where possible between different rehabilitative sentencing options, some of which may include neurointerventions.