



Neuroenhancement, the Criminal Justice System, and the Problem of Alienation

Jukka Varelius 

Received: 19 June 2019 / Accepted: 17 December 2019
© The Author(s) 2019

Abstract It has been suggested that neuroenhancements could be used to improve the abilities of criminal justice authorities. Judges could be made more able to make adequately informed and unbiased decisions, for example. Yet, while such a prospect appears appealing, the views of neuroenhanced criminal justice authorities could also be alien to the unenhanced public. This could compromise the legitimacy and functioning of the criminal justice system. In this article, I assess possible solutions to this problem. I maintain that none of them qualifies as a satisfactory general solution to it, a solution that could reasonably be taken to solve the problem or to suffice for dealing with it in at least most cases. Yet I also suggest that, depending on contingent empirical circumstances, the responses – either singly or together – can sometimes amount to a sufficient answer to it.

Keywords Alienation · Criminal justice system · Democracy · Neuroenhancement · Public opinion

Introduction

Neuroscientific knowledge is nowadays often considered important for developing the criminal justice system. For example, neuroscientific results have been seen to speak for replacing retributive punishment practices with more

lenient consequentialist ones (see e.g. [1]). It has also been proposed that neuroscientific technology can help in assessing defendants' culpability and the reliability of testimonies and that novel neurointerventions could be employed in rehabilitating offenders (see e.g. [2–5]).¹ Moreover, it has been suggested that neuroenhancements could be used to improve the abilities of criminal justice authorities. Judges could be made more able to make adequately informed and unbiased decisions, for instance (see e.g. [8–10]).

Several authors have, however, been concerned about such uses of neuroscientific knowledge. Neuroscientific results may not support consequentialist punishment practices after all,² neuroscience-based methods of culpability assessment and “mindreading” may remain overly imprecise, and, for a yet further instance, neurointerventions can violate offenders' rights and have disproportionate side effects (see e.g. [4, 11–14]). The debate on questions such as these continues (see e.g. [15]). One additional difficulty relates to a possible discrepancy between the views of

¹ Neurointerventions are already employed in the rehabilitation of, for instance, mentally disordered offenders (see also e.g. [6]). First mentions of uses of chemical means to reduce the sex-drive of sexual offenders reportedly date back to 1940s (see e.g. [7]).

² Although a retributivist theory of punishment can include some consequentialist considerations, retributivists may well find the consequentialist conclusions derived from neuroscientific results implausible. As theories of punishment are normative theories and assuming that normative conclusions cannot be drawn from purely factual starting points, retributivists can also have reasonable grounds for rejecting the support neuroscientific facts are seen to provide for consequentialism. Hence, advocates of retributivist views on punishment might not see the adoption of consequentialist punishment practices as (positive) development in the first place.

J. Varelius (✉)
Department of Philosophy, Contemporary History, and Political Science, FIN-20014 University of Turku, Turku, Finland
e-mail: jukka.varelius@utu.fi

neuroenhanced criminal justice authorities and those of the unenhanced public: the views of the former could be alien to the latter.³ Given that the alienation could compromise the legitimacy and functioning of the criminal justice system (see below), the prospect appears worrying.

That the unenhanced public could find the views of neuroenhanced criminal justice authorities counterintuitive and difficult, or even impossible, to accept – the problem of alienation, as I henceforth call it – has been acknowledged in the literature [8].⁴ Yet so far the problem would not appear to have received any detailed attention. This article takes a step towards remedying the shortage. I begin by explicating the main starting points of its argument. Then I assess four possible answers to the problem of alienation. I maintain that none of the responses qualifies as a satisfactory general solution to the difficulty, a solution that could reasonably be taken to solve the problem or to suffice for dealing with it in at least most cases in which it arises. Yet I also suggest that depending on contingent empirical circumstances the responses – either singly or together – can sometimes amount to a sufficient answer to the problem of alienation.⁵

Main Starting Points

How neuroenhancement should be defined has been debated (see e.g. [19, 20]). Yet, for the purpose of this article, it suffices to characterize the notion as referring to the improvement of the healthy abilities of or provision of new abilities for humans by means provided by

neuroscience. In practice neuroenhancement would typically work by affecting brain and other pertinent bodily mechanisms by such means as drugs, electric devices, and ultrasound (see e.g. [21]). In terms of developing existing mental abilities, it has been proposed that neuroenhancement could improve, for instance, memory, attention, wakefulness, self-control, ability to cope with stress, empathy, and capacity to make unbiased decisions (see e.g. [10]). Humans radically improved with novel technology – so-called posthumans (see e.g. [22]) – could have abilities unimaginable for us unenhanced people.⁶

Instead of trying to assess the effectiveness of neuroenhancements (see e.g. [23]), I now just assume that the possibility that human abilities could be significantly improved with neuroscientific means is realistic enough to merit the kind of attention given to it here. Importantly from the present viewpoint, besides the judiciary, neuroenhancements could also be employed in improving the legislature, the executive, and offenders. However, for reasons of space, I here focus on the judiciary.⁷ Neuroenhanced criminal justice authorities could be able to, say, comprehend and process a significantly larger amount of information and be markedly better able to avoid the undue influence of cognitive biases (see e.g. [24]) than unenhanced ones. That could allow the authorities to make decisions that are more efficient in reaching the goals of the criminal justice system than those made by unenhanced criminal justice authorities (see also e.g. [8–10]).⁸

What is the appropriate relationship between public opinion and the policies and practices of the criminal justice system? Few accept that the former should decide the latter. Indeed, it has been argued that allowing public opinion to determine punishments, for instance,

⁶ The below considerations apply to using genetic means to improve the abilities of criminal justice authorities as well, if genetic enhancements were to be deemed useful in the context.

⁷ This is not to say that the topics would be independent of each other. For instance, neuroenhanced criminal justice authorities could act in ways that motivate a (neuroenhanced) legislature to alter the constraints the legal framework places on judicial discretion. In some countries, the judiciary also makes common law and many jurisdictions allow the judiciary to change laws through the process of judicial review. Moreover, the solutions to the problem of alienation assessed below can also be of relevance from the viewpoint of the corresponding difficulties possibly arising in the cases of the legislature, the executive, and offenders.

⁸ As mere cognitive improvement might lead to more egoistic decisions only, I now assume that neuroenhanced criminal justice authorities would be improved in a way that also leads them to make morally better choices.

³ ‘Unenhanced public’ is here used, not derogatorily, but just to refer to those members of the public who would not be improved by neuroscientific or other corresponding means.

⁴ The problem of alienation focused on here is thus not the same as, for instance, the existentialist difficulty relating to a person’s estrangement from the world and herself or a Marxist concern focusing on the effects of a capitalist mode of production. By calling the problem now focused on the problem of alienation, I do not mean to imply that it necessarily is more important than the other problems previously referred to with the same name.

⁵ Needless to say, the problem of alienation is not the only difficulty related to neuroenhancing criminal justice authorities. The enhancement could, for instance, also make it more difficult for unenhanced people to engage in what then would be considered praiseworthy action and entail (further) unwarranted benefits to the enhanced, see e.g. [16, 17]. On the other hand, if a criminal justice authority would undergo neuroenhancement, say, because she is pressured to do so by her employer or out of fear that otherwise she is unable to compete with her colleagues, her decision to improve herself with neuroscientific means might not be autonomous (see e.g. [18]). I here abstract from difficulties such as these.

would lead to chaotic, unprincipled sentencing (see e.g. [25]).⁹ But also the other extreme, that the policies and practices of a criminal justice system should be totally independent of public opinion appears unconvincing. Indeed, it has been maintained that lack of adequate public support for the criminal justice system could lead to lower compliance, diminished confidence, less cooperation, and weaker penal censure (see e.g. [25]). Even when these kinds of practical repercussions would not ensue, the problem of alienation would arguably amount to a theoretical difficulty for the democratic legitimacy of a criminal justice system.¹⁰ Accordingly, below I assume that both of the above extremes are implausible and that the legitimacy and adequate functioning of a criminal justice system relies on the respect and acceptance of the public (see also e.g. [8]). As the argument of this article does not presuppose determining the precise degree of public support a criminal justice system should have, I now put that question aside.¹¹

Besides those focused on in the above paragraphs, the problem of alienation has three further preconditions. First, neuroenhancements could perhaps be used by most people or even by the population as a whole (see e.g. [33]). If the moral and cognitive abilities of most or all people were similar to those of neuroenhanced criminal justice authorities, the problem of alienation would be unlikely, or at least less likely, to arise. Hence, I now assume that at least a significant amount of the public would remain unenhanced.¹² Second, many criminal justice authorities might not want to use neuroenhancements and the idea that neuroenhancements would become mandatory for them appears unfeasible. However, the emergence of a significant degree of alienation does not presuppose that all or even most criminal justice

authorities would use neuroenhancements. Indeed, when circumstances are unfortunate enough, even one occurrence of the problem of alienation could be quite significant. Moreover, especially given the increasing demands of professional life, several criminal justice authorities could assumedly want to use neuroenhancements to improve their ability to meet the challenges they face in their work. Accordingly, I now presuppose that a significant number of criminal justice authorities would employ neuroenhancements.

Third, even if only the abilities of criminal justice authorities were improved with neuroscientific means, their decisions might be welcomed by the public frequently enough. However, given that the unenhanced public would remain, say, significantly less informed and more biased than neuroenhanced criminal justice authorities, I now assume that the former could often fail to understand and consequently refuse to endorse the views of the latter, even if those views were adequately warranted.¹³ Hence, I am not suggesting that the unenhanced public would not want criminal justice authorities to make better decisions. Instead, the idea is that the unenhanced public might not be able to appreciate the quality of the decisions neuroenhanced criminal justice authorities would make. As the problem of alienation has preconditions such as these, it is not bound to arise: whether it does emerge depends on contingent empirical circumstances. Yet, the public often finds the views of even unenhanced criminal justice authorities objectionable (see e.g. [34]). This suggests that just dismissing the problem of alienation as unlikely to actualize would often be overly optimistic.¹⁴ Accordingly, in practice it would be reasonable for a criminal justice system contemplating the use of neuroenhancements to consider the problem of alienation. Duly acknowledging it and the possible ways of solving it could also help in avoiding the problem of alienation, when it threatens to arise.

⁹ Yet, according to Morris Hoffmann [26, p., 1010], for example, “[j]urors are in a better position than judges not only to take the measure of a crime, but also to take the measure of its proportionate retribution.”

¹⁰ Of course, not everyone accepts that only democratic state power can be legitimate. And sometimes public support is considered relevant only to the degree that it leads to outcomes that are better in terms of some criteria defined independently of the democratic process. Yet most theories of legitimate state power would appear to give public support a significant role. For pertinent discussion, see e.g. [27–31]. The kind of practical difficulties mentioned above could also be severe even when public support is given only a minor, or no, role.

¹¹ For pertinent discussion, see e.g. [32].

¹² Neuroenhancing the entire population, or even most of its members, could be overly expensive and/or many people might refuse the offered enhancement, for instance.

¹³ In practice, the unenhanced public might object to the views of neuroenhanced criminal justice authorities in different ways and/or for different reasons. Yet, from the present viewpoint, the central thing is that the public would not accept those views. Accordingly, and for the sake of simplicity, I now speak of how the unenhanced public sees things.

¹⁴ The considerations below are relevant to addressing the opposition already faced by the views of unenhanced criminal justice authorities as well. Yet, the existing opposition is plausibly not, at least universally, as stark as it can be. And, given that they would be distributed in the above proposed way, the opposition could be expected increase with the possible use of neuroenhancements (see also the section ‘Informing the Public’ below).

Finally, below I use as an example the replacement of retributive punishment practices with more lenient consequentialist ones. As proposed, although controversially, neuroscientific results have been seen to speak for such a replacement [1] (cf. e.g. [13]), and it has also been proposed that neuroenhancing judges might make them less retributive (see e.g. [8, p., 336]). Moreover, such a replacement would typically amount to a significant change that might well conflict with the views of the unenhanced public. Yet the argument of this article does not presuppose that neuroenhancing criminal justice authorities would necessarily lead to that particular kind of change. Any significant alteration in the criminal justice system that the unenhanced public would find alien would give rise to the problem of alienation. The replacement of retributive punishment practices with more lenient consequentialist ones is just one potentially useful example.

Responses to the Problem of Alienation

As explained, the problem of alienation focused on here arises (when it does) from a dissonance between public opinion and the views of neuroenhanced criminal justice authorities. Accordingly, it could be taken that an adequate response to the difficulty can be found by bringing the conflicting positions sufficiently close to each other. In principle, there are two main ways of trying to reconcile them. First, the dissenting public could be persuaded to adopt the views of the neuroenhanced criminal justice authorities. Second, neuroenhanced criminal justice authorities could formulate their views so as to avoid excessive discrepancy with public opinion. Of the four possible responses to the problem of alienation considered below, the first three are based on the former idea whereas the fourth relates to the latter. These responses are suggested by previous philosophical discussion on the problem of alienation and on difficulties reminiscent of it arising in other contexts (see below).

Informing the Public

In a democracy, the power of governmental institutions such as the criminal justice system ultimately derives from the public. Yet democracies also typically consider it important to protect individuals against the excesses of majoritarian politics. Accordingly, individuals are commonly seen as having rights, entitlements with special normative force that provide reasons to treat their

holders or permit their holders to act in certain ways even if doing otherwise would serve a social goal or accord better with public opinion. In Ronald Dworkin's famous view, for instance, "[r]ights are best understood as trumps over some background justification for political decisions that states a goal for the community as a whole" [35 p., 153] (see also e.g. [36 p., 33] and [37, pp. 249–250]).¹⁵ Although some striking concrete examples to the contrary still exist, nowadays it is also quite commonly accepted that offending does not deprive one of one's rights.¹⁶ Indeed, such international declarations as the United Nations' *Standard Minimum Rules for the Treatment of Prisoners* [40] and *The United Nations Rules for the Treatment of Women Prisoners and Non-Custodial Measures for Women Offenders* [41] emphasize the protection of offenders' rights. A central responsibility in the protection of offenders' rights too apparently falls to the criminal justice system.

Moreover, while public opinion frequently conflicts with decisions made by criminal justice authorities (see e.g. [34]), the discrepancy does not seem that intractable. For instance, on the basis of an empirical study on public opinion on sentencing decisions, Leclerc, Niang, and Duval [42, p. 488, emphases in original] conclude that.

... there is a significant difference between the *general opinion* that citizens express on an abstract question about the work of the courts and the *specific opinion* that they give when asked to look at the work of the courts in a specific case.

In view of the study the authors conducted, the discrepancy between public opinion and court decisions is likely to diminish when the public is sufficiently informed about the reasons behind the decisions. Although they are not expressed in the context of neuroenhancement, these considerations could be taken

¹⁵ The position that rights have special normative force need not be taken to entail that a right can never be overridden, outweighed, or excluded by competing considerations. For pertinent discussion, see e.g. [38]. For a practical example, see e.g. [39].

¹⁶ To clarify, this is compatible with thinking that offending can, say, weaken, waive, or activate an exception clause in some of one's rights. Accordingly, the above stated view allows for punitive measures that involve restricting some rights of offenders in certain ways. Paying due attention to offenders' rights is important here insofar as the unenhanced public would, say, want offenders to have significantly harder punishments than neuroenhanced criminal justice officials would issue, punishments that could be taken to violate the rights of offenders.

to be relevant in connection with the problem of alienation as well. Accordingly, Chandler and Dodek [8, p. 337] write as follows:

... it is at least possible that the public, educated about the processes of moral reasoning, might come to regard unenhanced decisions as inferior and be willing to rely on enhanced judges even if they generate decisions that the unenhanced public would not have reached.

Assuming that the criminal justice system would duly protect the rights it is meant to protect and adequately inform the public about what it is doing and why, the unenhanced public might indeed consequently become favorable to the views of neuroenhanced criminal justice authorities even if it initially found those views alien. Revealing the reasons why neuroenhanced criminal justice authorities, say, arrived at a punishment of a particular degree of severity could make the sentence acceptable to the unenhanced public that first considered it overly lenient.

However, decisions made by neuroenhanced criminal justice authorities could also differ from the views of the unenhanced public to a significantly greater extent and markedly more frequently than choices made by unenhanced criminal justice authorities. This worry would apparently be the more pronounced the closer we got to posthuman criminal justice authorities. Yet the decisions of even more moderately enhanced criminal justice authorities could often be harder for the public to understand than those made by unenhanced criminal justice authorities. Hence, as such, this response does not qualify as an adequate general solution to the problem of alienation. It might now be taken that the public need not be expected to fully understand the reasons behind the choices neuroenhanced criminal justice authorities would make. After all, it is commonly accepted, for instance, that a patient can autonomously decide to undergo a treatment even if she does not fully comprehend the mechanisms by which it would alleviate her condition (see e.g. [43]).

Yet relying on a neuroenhanced criminal justice authority appears importantly different from choosing to undergo a treatment one does not fully comprehend. For trust in medical experts typically is, or at least arguably should be, based on a track record of adequate previous success. Hence, that a patient can autonomously choose to have a treatment she does not fully understand would

appear to presuppose that she has sufficient reason to think that the health care providers treating her can be relied on to provide adequate care for her. In the absence of such evidence, a patient's decision to undergo a treatment she does not comprehend does not appear reasonable.¹⁷ As long as no corresponding track record exists in the case of neuroenhanced criminal justice authorities, it would not appear to be reasonable for the public to endorse the views of neuroenhanced criminal justice authorities it finds hard or even impossible to comprehend. And in the absence of such track record, the public could also be quite unwilling to endorse the views. Therefore, I now turn to another possible response to the problem of alienation.

Challenging the Moral Views of the Public

The above considerations notwithstanding, informing the public about the grounds of the views of neuroenhanced criminal justice authorities could still be considered important to dealing with the problem of alienation. To refer to an analogy from a reminiscent context, in the debate on using artificial intelligence to improve morality it has been proposed that the moral judgments made by intelligent machines could function as a kind of Socratic assistant for the moral development of humans. In this view, by challenging the moral views of unenhanced humans, intelligent machines could lead individuals towards improved moral thinking (see e.g. [44, 45]). A similar idea could be taken to apply in this context too. In the resulting view, even if information about the grounds of the views of neuroenhanced criminal justice authorities would not as such suffice to persuade the dissenting public, the information could make the public question its opinions on the matters at hand. This questioning could ultimately lead the public to endorse improved views, ones that would accord with the decisions made by neuroenhanced criminal justice authorities.

Providing information about the grounds of the views of neuroenhanced criminal justice authorities could lead at least some members of the unenhanced public to question their related opinions. Sometimes those members of the public might also consequently adopt positions that better accord with the views of neuroenhanced

¹⁷ Yet a patient could assumedly sometimes autonomously decide to take the risk of undergoing such a treatment as, say, a measure of last resort.

criminal justice authorities. Yet it is quite unclear whether this would happen sufficiently often to avoid or to significantly alleviate the problem of alienation. In the debate on artificial intelligence and moral improvement, several authors have been concerned that the employment of intelligent machines could have severely detrimental effects on human moral agency. Whitby [46, p., 559] points out that many human societies have practiced a degree of moral authoritarianism and proposes that the majority of people might not feel the need to form moral judgments for themselves, if artificial intelligence made choices for them.¹⁸ Vallor [47, p., 112] maintains that “there are several areas of technological advancement in which the risks of moral deskilling are not presently balanced by forces driving moral upskilling or reskilling.”¹⁹ And Danaher [48, p., 133], for a yet further example, argues that “the rise of the robots will push the agency-like qualities to the background by reducing the need and desire for their expression.”²⁰

Although neuroenhancement would not turn criminal justice authorities into machines,²¹ the concern expressed in the above quotations applies in the present context as well. Hence, instead of instigating unenhanced people to develop their opinions, neuroenhancing criminal justice authorities could have severely detrimental effects on the moral agency of the unenhanced public. The result of the use of neuroenhancements could be that the unenhanced public starts to fatalistically submit to the official decisions. This would not be likely to give rise to such practical problems as lower compliance to or diminished cooperation with the criminal justice system. Yet, a criminal justice system in which the public more or less mindlessly submits to alien decisions made by a neuroenhanced minority would arguably lack adequate democratic legitimacy. Indeed, the detrimental effects that neuroenhancing criminal justice authorities could have on the moral agency of the unenhanced public

might even deepen the problem of alienation. For the losses in moral agency could lead the unenhanced public to endorse, say, less informed and more biased opinions than it had in the first place. Such views could apparently be even more discordant with the views of neuroenhanced criminal justice authorities than the opinions the public initially had. The possibility of such untoward results makes this response unsatisfactory as a general answer to the problem of alienation.

Better Results Will Persuade the Public

A criminal justice system typically aims to deter people from committing crimes, to incapacitate offenders, to provide retribution to victims of crime, and to help offenders to return to a law-abiding life (see e.g. [50]). With the exception of retribution, criticisms of criminal justice policies and practices usually rather relate to the means by which the goals are strived after – such as mass incarceration – than to the acceptability of the goals themselves (see e.g. [51–54]). Accordingly, it would seem that even neuroenhanced criminal justice authorities would strive after goals such as those just listed above. Given that the goals are quite concrete, it could be taken that the problem of alienation is best dealt with by just using neuroenhancements and letting the public see that neuroenhanced criminal justice authorities are more efficient in achieving pertinent goals than unenhanced ones. Statistics and studies demonstrating that a neuroenhanced criminal justice system really achieves better results would function like the track record supporting public trust in the medical profession. Hence, in the view at hand, the problem of alienation would dissolve in the face of concrete, practical evidence of the effectiveness of neuroenhancing criminal justice authorities.

However, given the now assumed relative rarity of neuroenhanced members of the public, offenders would most probably be unenhanced. Consequently, the decisions made by neuroenhanced criminal justice authorities would assumedly often be as alien to offenders as they would be to the public in general. If the choices made by enhanced criminal justice authorities would be difficult to offenders to comprehend, the decisions would also be unlikely to be more efficient in promoting such goals as rehabilitation than decisions made by unenhanced criminal justice authorities. Yes, neurointerventions could also be used to improve offenders and perhaps in a way that would increase their

¹⁸ Yet Whitby [46, p., 559] is also optimistic that some humans would still be willing to engage in moral reflection.

¹⁹ Moral deskilling refers to devaluation of the knowledge and skills involved in moral agency, whereas moral reskilling and moral upskilling denote the revaluation and increased valuation of the knowledge and skills, respectively.

²⁰ By ‘agency-like qualities’ Danaher refers to qualities such as understanding, intelligence, rationality, and the power to act in the world [48, p., 132].

²¹ Yet enhancement employing, say, brain-computer interfaces could turn the enhanced individuals into “cyborgs” (see e.g. [49]).

ability to understand choices made by neuroenhanced criminal justice authorities. Yet, although this could alleviate the problem of alienation somewhat, the views of neuroenhanced criminal justice authorities could nevertheless still remain alien to unenhanced offenders and the unenhanced public in general.

And even if offenders were able to adequately understand the decisions made by neuroenhanced criminal justice authorities, the latter would still be unlikely to, say, come up with rehabilitation methods that would quickly turn offenders into law-abiding citizens. Even if some fast conversions to the straight and narrow occurred, convincing the skeptical public that larger scale improvement is forthcoming could require gathering data about a significant number of cases over a lengthy period. Accordingly, even if neuroenhanced criminal justice authorities could in principle be more efficient in reaching the goals of the criminal justice system, collecting adequate evidence of this could sometimes take a significant amount of time. Moreover, if, say, a harsh retributive sentiment is deeply felt among it, convincing the public about the benefits of lenient punishments might turn out to be quite difficult even within a longer period. Consider, for example, the debate on physician-assisted dying. Even where there is significant amount of evidence against the view that legalizing physician-assisted dying leads to devastating consequences (see e.g. [55, 56]), as deeply felt as questions of suicide and euthanasia are, the data has failed to persuade opponents of the procedures. Instead, requests for further studies have been frequent enough to make some authors suggest that the search for additional data, data that would suffice to convince opponents of physician-assisted dying, may be interminable (see e.g. [57]).

Gathering adequate data on the improved effectiveness of neuroenhanced criminal justice authorities might not be as demanding as proving that legalizing physician-assisted dying does not have devastating practical consequences.²² Yet, for the reasons presented above, the acquisition of a sufficient amount of information to persuade the dissenting public could take much time in the former case too. Consider that the discrepancy between public opinion and the views of neuroenhanced criminal justice authorities is extensive. In such a situation in particular, promises to the effect

that adequate evidence of the increased effectiveness of neuroenhanced criminal justice authorities is forthcoming might not suffice to allay public dissatisfaction. Then the public's fear that no adequate evidence will actually ensue could, at worst, deprive the criminal justice system of the support it would need to be able to reach the improved results. Accordingly, also the view that the problem of alienation would dissolve in the face of concrete, practical evidence appears less than fully convincing as a general solution to the difficulty.

Accommodating the Public Opinion

Perhaps, then, the problem of alienation would be best dealt with by just accommodating the fact that the unenhanced public can see things quite differently than neuroenhanced criminal justice authorities? From this perspective, instead of trying to make unenhanced individuals receptive to views that could be even radically different from their opinion, neuroenhanced criminal justice authorities should formulate their views so that they would not conflict with how the unenhanced public sees things. If, for instance, sentencing offenders in a particular way were foreseen to endanger adequate public support for the criminal justice system, neuroenhanced criminal justice authorities should calibrate punishment practices to the public opinion so that the unfortunate result would be avoided. Given that the perspectives of the unenhanced public would be duly accounted for, there should then be no danger of excessive discrepancy between the views of neuroenhanced criminal justice authorities and public opinion.²³

Perhaps neuroenhanced criminal justice authorities would be able to adequately predict public reactions to their choosing one policy instead of another? That could allow them to avoid the problem of alienation. Yet the victory thereby achieved might also be a hollow one. For instance, according to Alper, Durose, and Markman

²² This, of course, is not to say that there cannot be jurisdictions within which legalization of physician-assisted dying would have intolerable consequences.

²³ Besides by a remark by Chandler and Dodek [8 p., 337], this possible response to the problem of alienation is suggested by the discussion on the so-called desire theories of prudential value. According to the type of theory, what is good and bad for an individual is to be defined in terms of what she desires. Since a person's desires can be based on, say, misinformation, often only adequately idealized desires have been deemed relevant to a desire theory (cf. however [58]). Yet, as what a person's idealized self wants can be alien to her actual self, it has also been proposed that what is central in a desire theory of prudential value is what a person's idealized self would want for the actual person (e.g. [59]). In connection with artificial intelligence, a similar idea has been discussed under the label of "coherent extrapolated volition" [60].

[61], 83% of state prisoners released in 2005 across 30 states of the United States were arrested at least once during the nine years following their release (see also e.g. [62]). Kelly [63] states that “[t]he great irony of the past fifty years of U.S. criminal justice policy is that we could not have intentionally designed and built a better recidivism machine than the one we have.” And Davis, for a yet further example, maintains that the American criminal justice system “remains in a state of crisis” [64 p., 1] (see also e.g. [15]).²⁴ Perhaps the just quoted views are somewhat hyperbolic. The criminal justice systems of some other countries might also be in a better shape. In any case, significant improvement of a criminal justice system would assumedly often presuppose more than mere minor adjustment.

However, given that significant discrepancy between public opinion and the views of neuroenhanced criminal justice authorities would compromise adequate public support for the criminal justice system, radical revisions could often be risky. As other than minor adjustment could result in loss of adequate public respect, measures meant to substantially improve a criminal justice system could ultimately make it even worse. Therefore, adjusting the decisions of neuroenhanced criminal justice authorities to predicted public reactions would appear to allow for a rather watered-down criminal justice reform at best. Even though small improvement can be better than none, such revision would seem to be unlikely to significantly alleviate the severe problems faced by many criminal justice systems worldwide. Making several incremental revisions in a row might sometimes ultimately allow for a more substantial change. Yet even smaller changes could be difficult for the unenhanced public to accept when they came one after another. Accordingly, successfully implementing such a piecemeal approach could require an excessive amount of time. Even if the promise of a brighter future would not be without value here, making such a promise appears an unsatisfactory solution to the pressing difficulties of today (see also the previous section). Finally,

consider that the pertinent views of the unenhanced public were quite uninformed and biased. In such a case, trying to avoid the problem of alienation by formulating punishment practices to accord with those views could result in punishments that, although acceptable to the public, would be all in all morally and legally unjustifiable. Hence, also accommodating the public opinion appears inadequate as a general solution to the problem of alienation.

Should Neuroenhancements Be Rejected in Improving the Abilities of Criminal Justice Authorities?

To sum up, as general solutions to the problem of alienation, all four of the above considered responses appear less than satisfactory. Assuming that they exhaust the relevant options, does this mean that criminal justice authorities should not be improved with neuroscientific means? Dealing with the difficulties criminal justice systems worldwide face is arguably, if not evidently, important. After all, the question is about protecting important rights of offenders, their victims, and the public in general. That neuroscientific techniques could make criminal justice authorities better able to handle the difficulties would provide reason for using the techniques.²⁵ Moreover, as proposed above, the problem of alienation is not a necessary consequence of using neuroscientific means to improve the abilities of criminal justice authorities. In sufficiently favorable circumstances, the problem would not arise at all. Accordingly, to a significant extent, the above considerations were based on empirical speculation as regards what could happen if the abilities of criminal justice authorities were improved with neuroscientific techniques. And the focus above was on rather bleak scenarios, formulated so as to prepare for the worst.

Yet, as Chandler and Dodek [8] propose, just informing the public about the reasons behind the views of neuroenhanced criminal justice authorities could sometimes make the initially reluctant public sufficiently favorable to the views. In some cases, providing the

²⁴ While after the Second World War rehabilitation was for decades tried extensively as an approach to criminal punishment in the United States, its theoretical and practical failures led to a near-consensus among criminologists that it should be viewed as a discredited approach. Yet rehabilitative ideas did not entirely disappear from the landscape and, in part due to the expectation that neuroenhancements could be employed in rehabilitation, have recently gained increased popularity. For pertinent discussion, see e.g. [65].

²⁵ Yet, as stated (note 5), the problem of alienation is not the only difficulty related to improving the abilities of criminal justice authorities with neuroenhancements. I also do not mean to say that more traditional means of improving the capacities of criminal justice authorities – such as further education about, say, how to avoid the undue effects of biases (see e.g. [24]) – would necessarily be redundant. Yet neuroscientific techniques could have a complementary role to play even when, or if, the more traditional means were considered (more) relevant.

information could lead the dissenting public to question its conflicting opinions and to consequently adopt views compatible with those of the neuroenhanced criminal justice authorities. Neuroenhanced criminal justice authorities could sometimes be more efficient in reaching the legitimate goals of the criminal justice system, whatever they are, quickly enough to make the initially dissenting public endorse their views on that basis. And, as was also already suggested, neuroenhanced criminal justice authorities could sometimes be able to accommodate the views of the dissenting public without thereby compromising the criminal justice reform at hand. If one or more of these optimistic scenarios would actualize, the problem of alienation would dissolve. Even if the scenarios would not realize in as positive a form as just depicted, their actualization could still significantly alleviate the problem of alienation – when it arises.

So, instead of implying that the abilities of criminal justice authorities should not be improved with neuroscientific means, the above considerations rather emphasize the importance of duly acknowledging the different possible consequences that such use of neuroscientific techniques could have. Given that the problem of alienation has received meager attention so far, I take it that drawing attention to the different scenarios is not trivial. Adequately acknowledging the different scenarios would, at least, allow for making a better informed decision on whether or not to employ neuroscientific techniques to improve the abilities of criminal justice authorities.²⁶ Moreover, the above considerations support the use of neuroenhancements in developing the criminal justice system in the following sense. Given that neuroenhanced individuals would be cognitively more able than unenhanced ones, neuroenhancement could make criminal justice authorities better able to predict the consequences of resorting to decisions made by neuroenhanced criminal justice authorities in improving a criminal justice system. Hence, the problem of alienation does not imply that neuroenhancement should be rejected in developing the abilities of criminal justice authorities.

²⁶ As the precise nature of the criminal justice reform initiated by neuroenhanced criminal justice authorities and the specific empirical circumstances in which the reform would be implemented could vary from one case to another, presenting definite general recommendations as to how to deal with the problem of alienation in practice – when the problem arises – is unfortunately difficult.

Conclusion

The views of neuroenhanced criminal justice authorities could be alien to the unenhanced public. This could compromise the legitimacy and functioning of the criminal justice system. In this article, I considered possible responses to this problem – the problem of alienation, as I called it. The first three of the responses were about persuading the dissenting public to endorse the views of the neuroenhanced criminal justice authorities, whereas the fourth one focused on the authorities' formulating their views so as to avoid excessive discrepancy with public opinion. I maintained that none of the responses can reasonably be taken to solve the problem of alienation or to suffice for dealing with it in at least most cases in which it arises. Yet I also suggested that, depending on contingent empirical circumstances, the responses – either singly or together – can sometimes amount to a sufficient answer to the problem of alienation.

As proposed, the problem of alienation is not the only difficulty related to using neuroscientific knowledge to reform the criminal justice system (see e.g. [4, 11–14, 23]). Hence, the argument of this article does not imply that the usage of neuroscientific knowledge would be unproblematic even if the problem of alienation was avoided or solved. Yet discussion on other problems related to developing the criminal justice system with neuroscientific means is now left for future work.²⁷

Acknowledgments I thank Juha Räikkä and two anonymous reviewers of this journal for comments and the Academy of Finland (grant number 318954) for financial support.

Funding Information Open access funding provided by University of Turku (UTU) including Turku University Central Hospital.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and

²⁷ Needless to say, the views of neuroenhanced people may conflict with those of unenhanced individuals also outside of the context of the criminal justice system. Besides other spheres of government administration, education, and the defense forces, the examples include other branches of working life, leisure activities, and the private sphere. The considerations presented above are, *mutatis mutandis*, relevant in those areas as well. And the considerations apply also when the discrepancies between the views of different parties ultimately result from some other source than neuroenhancement.

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Greene, Joshua, and Jonathan Cohen. 2004. For the law, neuroscience changes nothing and everything. *Philosophical Transactions Royal Society London B Biological Science* 359: 1775–1785.
- Birks, David, and Thomas Douglas. 2018. *Treatment for crime: Philosophical essays on neurointerventions in criminal justice*. Oxford: Oxford University Press.
- Douglas, Thomas. 2019. Nonconsensual neurocorrectives and bodily integrity: A reply to Shaw and Barn. *Neuroethics* 12: 107–118.
- Murphy, Emily R., and Henry T. Greely. 2011. What will be the limits of neuroscience-based mindreading in the law? In *Oxford handbook of neuroethics*, ed. Judy Illes and Barbara J. Sahakian, 635–653. New York: Oxford University Press.
- Pugh, Jonathan, and Hannah Maslen. 2017. 'Drugs that make you feel bad'? Remorse-based mitigation and neurointerventions. *Criminal Law and Philosophy* 11: 499–522.
- Bublitz, Jan C. 2018. 'The soul is the prison of the body': Mandatory moral enhancement, punishment, and rights against neurorehabilitation. In *Treatment for crime: Philosophical essays on neurointerventions in criminal justice*, ed. David Birks and Thomas Douglas, 289–320. Oxford: Oxford University Press.
- Miller, Robert D. 1998. Forced administration of sex-drive reducing medications to sex offenders: Treatment or punishment? *Psychology, Public Policy, and Law* 4: 175–199.
- Chandler, Jennifer A., and Adam Dodek. 2016. Cognitive enhancement in the courtroom: The ethics of pharmacological enhancement of judicial cognition. In *Cognitive enhancement: Ethical and policy implications in international perspectives*, ed. Fabrice Jotterand and Veljko Dubljevic, 329–345. New York: Oxford University Press.
- Douglas, Thomas. 2013. Moral enhancement via direct emotion modulation: A reply to John Harris. *Bioethics* 27: 160–168.
- Sandberg, Anders, Walter Sinnott-Armstrong, and Julian Savulescu. 2011. Cognitive enhancement in courts. In *Oxford handbook of neuroethics*, ed. Judy Illes and Barbara J. Sahakian, 273–284. New York: Oxford University Press.
- Freedman, David, and Simona Zaami. 2019. Neuroscience and mental state issues in forensic assessment. *International Journal of Law and Psychiatry* 65: 1–8.
- ter Meulen, Ruud, Ahmed Mohamed, and Wayne Hall. 2017. *Rethinking cognitive enhancement*. New York: Oxford University Press.
- Morse, Stephen J. 2015. Neuroscience, free will, and criminal responsibility. In *Free will and the brain: Neuroscientific, philosophical, and legal perspectives*, ed. Walter Glannon, 251–286. Cambridge: Cambridge University Press.
- Shaw, Elizabeth. 2019. The right to bodily integrity and the rehabilitation of offenders through medical interventions: A reply to Thomas Douglas. *Neuroethics* 12: 97–106.
- Buyx, Alena, and David Birks. 2018. Neuroscience and social problems: The case of neuropunishment. *Cambridge Quarterly of Healthcare Ethics* 27: 628–634.
- Archer, Alfred. 2016. Moral enhancement and those left behind. *Bioethics* 30: 500–510.
- Sparrow, Robert. 2014. Egalitarianism and moral bioenhancement. *American Journal of Bioethics* 14: 20–28.
- Appel, Jacob. 2008. When the boss turns pusher: A proposal for employee protections in the age of cosmetic neurology. *Journal of Medical Ethics* 34: 616–618.
- Raus, Kasper, Farah Focquaert, Maartje Schermer, Jona Specker, and Sigrid Sterckx. 2014. On defining moral enhancement: A clarificatory taxonomy. *Neuroethics* 7: 263–273.
- Savulescu, Julian, Anders Sandberg, and Guy Kahane. 2011. Wellbeing and the concept of enhancement. In *Enhancing human capacities*, ed. Julian Savulescu, Ruud ter Meulen, and Guy Kahane, 3–18. Chichester: Wiley-Blackwell.
- Jwa, Anita S. 2019. Regulating the use of cognitive enhancement: An analytic framework. *Neuroethics* 12: 293–309.
- Bostrom, Nick. 2009. Why I want to be a posthuman when I grow up. In *Medical enhancement and posthumanity*, ed. Bert Gordijn and Ruth Chadwick, 107–137. Dordrecht: Springer.
- Schleim, Stephan, and Boris B. Quednow. 2018. How realistic are the scientific assumptions of the neuroenhancement debate? Assessing the pharmacological optimism and neuroenhancement prevalence hypotheses. *Frontiers in Pharmacology* 9: 1–7.
- Kahneman, Daniel. 2011. *Thinking, fast and slow*. New York: Farrar, Straus & Giroux.
- Roberts, Julian V. 2014. Mitigation and aggravation at sentencing. In *Encyclopedia of criminology and criminal justice*, ed. Gerben Bruinsma and David Weisburd, 3103–3112. New York: Springer.
- Hoffman, Morris B. 2003. The case for jury sentencing. *Duke Law Journal* 52: 951–1010.
- Christiano, Thomas. 2008. *The constitution of equality*. Oxford: Oxford University Press.
- Estlund, David. 2008. *Democratic authority*. Princeton: Princeton University Press.
- Kolodny, Niko. 2014a. Rule over none I: What justifies democracy. *Philosophy and Public Affairs* 42: 195–229.
- Kolodny, Niko. 2014b. Rule over none II: Social equality and the justification of democracy. *Philosophy and Public Affairs* 42: 287–336.
- Peter, Fabienne. 2009. *Democratic legitimacy*. New York: Routledge.
- Ryberg, Jesper, and Julian V. Roberts. 2014. *Popular punishment: On the normative significance of public opinion*. New York: Oxford University Press.

33. Persson, Ingmar, and Julian Savulescu. 2012. *Unfit for the future: The need for moral enhancement*. Oxford: Oxford University Press.
34. Roberts, Julian V., Loretta J. Stalans, David Indermaur, and Mike Hough. 2003. *Penal populism and public opinion: Lessons from five countries*. Oxford: Oxford University Press.
35. Dworkin, Ronald. 1984. Rights as trumps. In *Theories of rights*, ed. Jeremy Waldron, 153–167. Oxford: Oxford University Press.
36. Nozick, Robert. 1974. *Anarchy, state, and utopia*. Oxford: Wiley.
37. Raz, Joseph. 1986. *The morality of freedom*. Oxford: Oxford University Press.
38. Rumbold, Benedict. 2019. Towards a more particularist view of rights' stringency. *Res Publica* 25: 211–233.
39. Canadian charter of rights and freedoms. 1981. <https://www.justice.gc.ca/eng/csj-sjc/rfc-dlc/ccrf-ccdl/resources-resources.html#copy>. Accessed 23 Sept 2019.
40. United Nations. 1955. *Standard minimum rules for the treatment of prisoners*. <https://www.ohchr.org/en/professionalinterest/pages/treatmentofprisoners.aspx>. Accessed 24 April 2019.
41. United Nations. 2010. *The United Nations rules for the treatment of women prisoners and non-custodial measures for women offenders*. <https://www.un.org/en/ecosoc/docs/2010/es%202010-16.pdf>. Accessed 24 April 2019.
42. Leclerc, Chloe, Anta Niang, and Marie-Chloe Duval. 2018. Understanding the relationship between public opinion and experience with the criminal justice system. *International Journal of Public Opinion Research* 30: 473–492.
43. Beauchamp, Tom L., and James F. Childress. 2013. *Principles of biomedical ethics*. 7th ed. New York: Oxford University Press.
44. Giubilini, Alberto, and Julian Savulescu. 2018. The artificial moral advisor. The “ideal observer” meets artificial intelligence. *Philosophy & Technology* 31: 169–188.
45. Lara, Francisco, and Jan Deckers. 2019. Artificial intelligence as a Socratic assistant for moral enhancement. *Neuroethics*: 1–13. <https://doi.org/10.1007/s12152-019-09401-y>.
46. Whitby, Blay. 2008. Computing machinery and morality. *AI & Society* 22: 551–563.
47. Vallor, Shannon. 2015. Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character. *Philosophy & Technology* 28: 107–124.
48. Danaher, John. 2019. The rise of the robots and the crisis of moral patiency. *AI & Society* 34: 129–136.
49. Clark, Andy. 2003. *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. New York: Oxford University Press.
50. Duff, Antony, and Zachary Hoskins. 2018. Legal punishment. In *The Stanford encyclopedia of philosophy (fall 2018 edition)*, ed. Edward N. Zalta. <https://plato.stanford.edu/archives/fall2018/entries/legal-punishment/>. Accessed 17 April 2019.
51. Bam, Gulzaar. 2019. Can medical interventions serve as ‘criminal rehabilitation’? *Neuroethics* 12: 85–96.
52. Caruso, Gregg D. 2018. Justice without retribution: An epistemic argument against retributive criminal punishment. *Neuroethics*: 1–16. <https://doi.org/10.1007/s12152-018-9357-8>.
53. Pereboom, Derk. 2005. *Living without free will*. Cambridge: Cambridge University Press.
54. Pereboom, Derk. 2014. *Free will, agency, and meaning in life*. New York: Oxford University Press.
55. Potter, Jordan. 2019. The psychological slippery slope from physician-assisted death to active euthanasia: A paragon of fallacious reasoning. *Medicine, Health Care and Philosophy* 22: 239–244.
56. Rietjens, Judith A.C., Paul J. van der Maas, Bregje D. Onwuteaka-Philipsen, Johannes J. M. van Delden, and Agnes van der Heide. 2009. Two decades of research on euthanasia from the Netherlands: What have we learnt and what questions remain? *Journal of Bioethical Inquiry* 6: 271–283.
57. Parker, Malcolm. 2005. End games: Euthanasia under interminable scrutiny. *Bioethics* 19: 523–536.
58. Lin, Eden. 2019. Why subjectivists about welfare needn't idealize. *Pacific Philosophical Quarterly* 100: 2–23.
59. Railton, Peter. 1986. Facts and values. *Philosophical Topics* 14: 5–31.
60. Yudkowsky, Eliezer. 2004. *Coherent extrapolated volition*. San Francisco, CA: The Singularity Institute <https://intelligence.org/all-publications/>.
61. Alper, Mariel, Matthew R. Durose, and Joshua Markman. 2018. 2018 update on prisoner recidivism: A 9-year follow-up period (2005–2014). U.S. Department of Justice. <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=6266>. Accessed 7 May 2019.
62. Durose, Matthew R., Alexia D. Cooper, and Howard S. Snyder. 2014. Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010 - update. U.S. Department of Justice. <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=4986>. Accessed 7 May 2019.
63. Kelly, William R. 2018. Why punishment doesn't reduce crime: Evidence demonstrates why punishment does not change criminal offending. *Psychology Today*. <https://www.psychologytoday.com/us/blog/crime-and-punishment/201804/why-punishment-doesnt-reduce-crime>. Accessed 24 April 2019.
64. Davis, Angela J. 2018. The progressive prosecutor: An imperative for criminal justice reform. *Fordham Law Review* 87. <https://ir.lawnet.fordham.edu/fdro/>. Accessed 29 April 2019.
65. Matravers, Matt. 2018. The importance of context in thinking about crime-preventing neurointerventions. In *Treatment for crime: Philosophical essays on neurointerventions in criminal justice*, ed. David Birks and Thomas Douglas, 71–93. Oxford: Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.