THE ABOLITION OF PUNISHMENT: IS A NON-PUNITIVE CRIMINAL JUSTICE SYSTEM ETHICALLY JUSTIFIED?

- Przemysław Zawadzki -

Abstract: Punishment involves the intentional infliction of harm and suffering. Both of the most prominent families of justifications of punishment – retributivism and consequentialism – face several moral concerns that are hard to overcome. Moreover, the effectiveness of current criminal punishment methods in ensuring society's safety is seriously undermined by empirical research. Thus, it appears to be a moral imperative for a modern and humane society to seek alternative means of administering justice. The special issue of *Diametros* "The Abolition of Punishment: Is a Non-Punitive Criminal Justice System Ethically Justified?" was brought into life precisely to give the authors a platform for such progressive inquiries. And it is now safe to say that this platform has been put to excellent use, since Valerij Zisman, Alexander Stachurski, Giorgia Brucato, Perica Jovchevski, Sofia M. I. Jeppsson, Stephen G. Morris, Benjamin Vilhauer, John Lemos, Saul Smilansky, Elizabeth Shaw, Mirko Farina, Andrea Lavazza and Sergei Levin have presented such thought-provoking texts that they are bound to set the stage for debate in the years to come. This article is an introduction to this special issue and to the authors' papers.

Keywords: social and political philosophy, criminal justice, punishment, abolitionism, retributivism, desert, free will, moral responsibility, criminal quarantine, restorative justice

Published online: 19 April 2024

Punishment involves the intentional infliction of harm and suffering. The state typically defends its citizens against such mistreatment. Thus, it is widely accepted that legal punishment requires both moral and political justification. Philosophers have therefore studied "the justifications of punishment" to explain why the state can intentionally harm its citizens when they commit criminal offences. Historically, the justifications of punishment have tended to fall into one of the two main camps. Consequentialists justify punishment instrumentally, in virtue of some valuable end, such as crime reduction, that punishment supposedly allows achieving. Retributivists justify punishment on the grounds that it is an intrinsically appropriate, because deserved, response to crime. However, both of the most prominent justifications of punishment face a number of moral concerns, and the effectiveness of current criminal punishment methods in ensuring society's safety and reducing crime is seriously undermined by empirical research.

Przemysław Zawadzki Jagiellonian University Faculty of Philosophy, Institute of Philosophy Grodzka 52, 31-044 Krakow, Poland E-mail: przemyslaw.zawadzki@uj.edu.pl In this situation, it seems to be a moral imperative for a modern and humane society to seek alternative means of administering justice. This special issue was brought into being precisely to give the authors a platform for such inquiries. And, it can now be said, they have made excellent use of it, as leading philosophers in the field, alongside a younger generation of rising stars, have developed such thought-provoking texts that they are bound to set the stage for debate in the coming years.

As might be expected given the importance and controversial nature of the topic, perspectives proposed in this special issue of Diametros - both in terms of philosophical and ethical starting points as well as pragmatic and political conclusions - varied. However, if I were to hazard any generalization - doing this unholy misdeed in full knowledge (and cheeky enjoyment) of the potential agitation it may inspire in some of the authors who published their nuanced texts here because of my crude (but still fair, I maintain) simplification – it would be as follows: It is the end of the paradigm of thinking about criminal punishment as a tool of social policy aimed at justice; there is also less and less faith in traditional justifications for punishment appealing to philosophical notion of desert-grounding free will; some, however, for fear of the revolutionary consequences of realizing this state of affairs at the level of society as a whole, nevertheless prefer the bonds of tradition, which, though unsatisfactory in the light of reason, provides the opium for the people, whose soothing effect allows us to keep intact how we think of ourselves as moral agents. Everyone, however, is on the lookout for change when it comes to criminal justice policies, some more revolutionary than others, but being stuck in the old patterns of thinking is unanimously perceived as leading at best to becoming like the society in Aldous Huxley's Brave New World, consuming soma instead of facing the not-so-pretty injustices of reality.

Valerij Zisman (2024) will certainly interest those who have always wanted to know but were afraid to ask about the revolutionary movement of criminal punishment abolitionism. Zisman aims at dispelling the fears that have grown up around this agenda, arguing that, contrary to common assumptions, the abolition of punishment within the criminal justice system would not result in an ideological void or leave us bereft of alternatives. On the contrary, there are already promising alternatives that can replace criminal punishment: compulsory victim restitution, restorative justice, and Derk Pereboom and Gregg Caruso's public health-quarantine model (hereafter: PHQM). Moreover, as the author argues, these currently most influential versions of abolitionism are not unified only by a critical assessment of how criminal punishment fulfills the functions and upholds the values that our responses to crime should fulfill or uphold, but also converge on the same alternative to criminal law: compensation as a type of sanction and restorative justice as a procedural approach. Despite some tensions between the theoretical commitments of the three non-punitive alternatives mentioned above, Zisman sees the prospect of making them compatible with each other, and thus creating a unified abolitionist approach that could ensure what ought to be secured by the criminal justice system. The author concludes that although abolitionists have not yet offered a comprehensive account which fully demonstrates how all the essential functions and values of a full-fledged criminal justice system can be realized by a non-punitive alternative, neither has such account been offered in any satisfactory way by advocates of the criminal justice system that resorts to punishment. This diagnosis shows that the state of the game is open, and the status quo is not at all ahead of the game.

Alexander Stachurski (2024) offers a rich and lively analysis of a real-world example of a non-state justice system (Sistema Comunitario de Seguridad, Justicia y Reeducación, hereafter: SCSJR) applied by some of the Afromexican and Indigenous communities in the state of Guerrero, Mexico. Although the SCSJR is not non-punitive in nature, it shares significant similarities with restorative justice principles, emphasizing reconciliation, the importance of addressing victims' harm, and the educative intention of sanctions. Stachurski describes each stage of the SCSJR process that takes place when a person from the community resorts to anti-social behavior - a category applicable to both petty and a grave crime. The SCSJR process consists of the investigation stage, the judgment stage, and the reeducation stage. Although Stachurski interprets the SCSJR as an example of the maximalist restorative justice system, the author points out how at times it departs from restorative principles. For instance - being developed in an area where state law is almost unenforced, yet ideal for drug cultivation and trafficking - the SCSJR allows for the prosecution of the possession of drugs. Since drugs possession does not directly harm any party, its sanctioning is questionable from the perspective of restorative justice. Another significant difference is that the SCSJR is to some extent based on a hierarchical structure, while restorative justice should strive for a process between equals. Stachurski is not shy when it comes to the flaws of the SCSIR; the author mentions, for example, the problem of a lack of clear guidelines for challenging decisions, which undermines the system's ability to ensure the legitimacy of its results, as well as the lack of adequate support for those without relational ties and familial support in the community. Stachurski also identifies limits of the SCSJR; firstly, the system may not be translatable to individualistic societies as it presupposes interdependencies and values of cohesive communities; secondly, it faces some limitations, particularly salient in regions plagued by organized crime, in dealing with extreme crimes like murder or sexual violence, which may be impossible to overcome without the intervention - not necessarily the War on Drugsstyle intervention - of a powerful and functioning state. Despite these shortcomings, as the author points out, the reduction in crime rates under the SCSJR shown by studies is admirable - all the more so in the context of the unfavorable geographic location and systemic difficulties faced by communities that have implemented the SCSJR.

Giorgia Brucato and Perica Jovchevski (2024) are also among a group of philosophers who embrace the prospect of imminent changes in the criminal justice system, seeing no good justification for punitive practices in their progressive vision of society as a fair system of cooperation among free and equal individuals. They argue against fair-play retributivists, who claim that punishment is morally justified as a matter of justice. The authors begin their argument by assuming two premises of fair-play retributivism: the vision of society and the social purpose of criminal justice, which is the duty to respond to criminal acts to uphold the functioning of the cooperative venture. Along the way, however, they depart from fair-play retributivism, arguing that what constitutes crime is not taking unfair advantages, but rather inflicting unfair disadvantages on the relevant stakeholders: victims, community, and the state. Each of these stakeholders is disadvantaged by the crime in a different way: the victims by direct harm on them,

the community by betrayal of trust, and the state by undermining its authority. This conceptualization of crime allows the authors – in contrast to fair-play retributivists, as they argue – to recognize the seriousness of the offenses and adjust responses to them on that basis, thereby avoiding "proportionality objection" and to adequately account for the impersonal "crime" of disobeying the law. Moreover, the authors do not make do at criticizing retributivism, but advocate for their own non-punitive approach, defending it against commonly raised charges that without punishment criminals cannot be adequately held accountable for their deeds and victims cannot be given justice. The arguments presented by the authors will be particularly attractive to those who endorse the premises of the fair-play model of political obligations.

Sofia M. I. Jeppsson (2024) deals another blow to retributivism. Jeppsson makes a strong case that even in an idealized - let alone flawed, real-world - version of the retributivist criminal justice system, respect for offenders is unlikely to be provided. To this end, the author challenges Strawsonian idea that retributivism can function as a continuation of ordinary interpersonal practices in which people reciprocally place demands on each other and hold each other responsible, thereby taking each other seriously as moral agents who express mutual respect. The retributivist criminal justice system cannot function this way because, in a nutshell, there is hardly any reciprocity between the accused and the court. More specifically, Jeppsson discusses three differentiating features between the responsibility dynamics found in everyday relationships lacking significant power disparities and interactions within the retributivist criminal court: Firstly, accountability within the court operates unilaterally. Secondly, offenders are unable to opt out of the court proceedings. And thirdly, the court does not care about morality, but adheres to the law when making decision to impose punishment (so the only moral justifications for the defendant's behavior that can be taken into account by the court are the ones already recognized by the law). According to Jeppsson's examination, even in an idealized retributivism - envisioned within a hypothetical society devoid of prejudice - there remains little room left for respect for offenders; upon scrutinizing real-world instances of retributivism, it becomes evident that a respectful attitude towards offenders is spectacularly absent.

Stephen G. Morris (2024) – in contrast to Jeppsson, who provides a respect-based and thus morality-invoking argument for rejecting a retributivist criminal justice system – offers a non-moral case against retributivism grounded in self-interest and empathy considerations. Morris agrees with skeptics, such as Pereboom and Caruso, that the denial of free will and basic desert moral responsibility is the best justified view, and only if it were proven false could retributivism be morally justified. At the same time, however, the author believes that such skepticism commits to a global rejection of morality (a view, which is worth mentioning to expose the dialectical nature of the current special issue, contradicted by another of the authors, Benjamin Vilhauer); and for this reason, Morris maintains, Pereboom and Caruso's moral argument against retributivism fails. More specifically, Pereboom and Caruso's attempt to rebut retributivism by appealing to its moral wrongness is argued to be at odds with Pereboom's response to free will revisionist challenge issued by Manuel Vargas and Daniel Dennett. Pereboom claims that free will revisionism should be rejected on the grounds that it deviates too far from folk moral intuitions (i.e., predominantly backward-looking and libertarian); but this

kind of criticism, Morris contends, commits Pereboom to endorse a notion of morality that is consistent with folk morality. Instead, Pereboom defends forward-looking revision of morality stripped of both libertarian and basic desert elements (i.e., a revision according to which, e.g., moral wrongness is possible without moral responsibility or blameworthiness). To be consistent, Morris argues, Pereboom cannot be selective in relying on concepts that resonate with folk beliefs. This leads to a dilemma for Caruso and Pereboom: either rejecting any kind of morality-appealing argument, including one that purports to show that retributivism is immoral, or providing a justification for revising the folk concept of free will, which would, however, undermine Pereboom's offered reason for why skepticism is preferable to free will revisionism. Despite his own skepticism about the moral case against retributivism, Morris believes it is untenable for reasons beyond morality, such as the desire to live in a society that is best tailored for individuals to flourish (a desire grounded both in their self-interested concern with security, demonstrably best provided in non-punitive systems, and in their empathy for other members of society who under such systems are less likely to suffer either from the crime itself or from an unjust attribution of culpability and punishment).

Benjamin Vilhauer (2024) presents a sophisticated defense of the fairness of a policy of imprisoning violent offenders for free will and moral responsibility skeptics and deniers. Vilhauer's social contract justification of punishment – or incapacitation, since the permitted responses to crime on this account seem devoid of intentional infliction of harm and suffering - draws on a Kantian and Rawlsian heritage, making it deontological at its core and thus providing a different foundation for the justification of imprisonment than the usual reasons put forth by skeptics with their consequentialist underpinning. According to Vilhauer's account, a fair criminal justice system can be determined if we situate ourselves in the Rawlsian original position deliberation, in which we assume that we would be the target of a criminal justice system when "the veil of ignorance" is lifted. From this position, we are then supposed to consent to a certain kind of response to our conducts from a criminal justice system, which has a standard of conviction based on reasonable doubt. Vilhauer proposes that if we were better off living in prison than in the state of nature (a rather grim, Hobbesian vision of a situation in which the state apparatus does not function), then as rational deliberators we should accept a criminal justice system that has us imprisoned when we commit a violent crime; for relying solely on lenient measures, such as fines, would not do in the long run. On the other hand, no rational deliberator would prefer death over life in the state of nature, so we would not consent to the capital punishment. At the same time, knowing human psychology, we would not opt for "funishment" in our envisioned state prisons, but rather conditions unpleasant enough to be sufficiently deterrent not to encourage crime. But this implies general deterrence, which inevitably amounts to using punished as a means to an end. Vilhauer notices that his deontological social contract justification reveals principles of imprisonment similar to those advocated by Pereboom and Caruso in their PHQM, but also offers the resources to provide a stronger response to "the mere means objection"; the consent of the punished required on the author's social contract account proves that they are not used as a *mere* means, but as a means, and furthermore, they end up using the citizens thusly protected from crime as a means to ensure the best conditions in prison. So, in a sense the punished are not just "a resource," but participate in a reciprocal relationship. Finally, Vilhauer offers a way to avoid moral nihilism for skeptics and deniers, entering into discussion with Morris by showing how the "ought implies can" principle can be preserved through an epistemic interpretation of "can" and an appeal to counterfactual rational consent, which is at the core of the social contract thinking.

John Lemos (2024) compares two conceptions of the criminal justice system consistent with free will and basic desert moral responsibility skepticism and denialism: the already mentioned Pereboom and Caruso's PHQM and Michael Corrado's corrections model. Lemos argues that Corrado's claim about his corrections model being preferable to rational and informed persons over the PHQM is not easy to sustain on closer scrutiny because: a) the PHQM provides better protection against crime, and the cost for those who may be caught up in the system are lower in comparison to the corrections model, (b) Pereboom and Caruso can incorporate measures into their PHQM that minimize the risk of state abuse, and (c) Corrado's model is not immune to government misuse. Lemos makes four points regarding (a): (i) Pereboom and Caruso's model allows for pre-emptive detention, thus providing additional protection against crime, (ii) Corrado's model, presuming maximum sentencing guidelines, authorizes a greater number of unreformed and still dangerous offenders to be released back into society, (iii) the PHQM releases from incapacitation only if state psychologists determine that a detainee is no longer dangerous, (iv) unlike in Corrado's model, (i)-(iii) are met in the PHQM without subjecting offenders to harsh conditions. When it comes to (b), Lemos points out that since in Pereboom and Caruso's model the determination of who poses a threat to the public could be carried out by medical health professionals, not agents of the state, such a separation of powers creates structural protection against system abuses by the state. Regarding (c), the author notes that even under the corrections model, judges, prosecutors and juries can still be subject to corruption, incompetence, laziness and biases that lead to conviction of innocent people and excessive criminal sentences; in such unfortunate cases, the PHQM provides an opportunity to free many wrongly convicted individuals on the basis that they cease to be dangerous, which would often happen sooner than under the recommendations of the corrections model. In conclusions, Lemos places these findings in the context of his own views. The author believes that although given the cost/benefit calculations the corrections model cannot be shown to be preferrable for rational and informed persons, Corrado's concerns about the PHQM are legitimate. With that in view, Lemos suggests that it may be worth reconsidering the merits of believing in the existence of free will and basic desert moral responsibility, or at least living and acting as if people possess them, to provide principial protections for human dignity within the context of the criminal justice. In this way, the author proposes to open the door to endorsing a belief in desert-grounding free will on purely pragmatic grounds.

Saul Smilansky (2024) has written extensively on the exact question of what lies behind this door. Smilansky captures his views in his well-known account of illusionism, according to which, although we lack free will and basic desert moral responsibility, we should do a great deal to maintain and nurture the illusion that we have them, because they are essential for a meaningful life; after all, for at least several hundred years, they have constituted foundations of morality, respect, justice, self-understanding, and human

relations. Giving up this illusion, as proposed by skeptics and denialists, is a dangerous gamble, since it introduces a change that would go very deep (possibly against basic features of human nature) and into uncharted territory: even such candidates for living in harmony with what would entail rejecting faith in desert-grounding free will as Protestants with their belief in predestination have not fully internalized the non-existence of free will, as evidenced by their belief that success based on action is to certify that they have been chosen for salvation, or the "honor cultures" with their complacency in punishing innocent people based on their family or clan affiliation, still do not seem to deny personal control and responsibility. Riskier still, denialism threatens to reject something that constitutes, according to the author, the greatest form of historical progress, i.e., judging and treating people based on such individual control and responsibility, thus threatening civilization with a return to something akin to an honor culture in its collective form of punishment or a straightforward collapse to barbarism. But the skeptical cut does not end there; it goes even deeper under the skin, since it leads to a moderate form of nihilism: people can of course enjoy hedonistic experiences and even be motivated to act in socially beneficial ways, but deep down their efforts, projects and work are meaningless because life can no longer be founded on agency-based values, and thereby credit, authentic pride and self-respect cannot be earned. The consequentialist mindset that denialists are left with also has dangerous social implications: it can bypass respect for persons to achieve good social ends at the expense of individuals' rights, lower the standards for prosecution and conviction, or even provide justification for the "funishment" of the innocent. For all these reasons, Smilansky concludes, we should follow the consensus direction (i.e., changing social settings from which crime emerges), give a chance to "fundamental dualism" or "compatibility-pluralism" (i.e., a mixed account, which combines compatibilism with denialism), and stay in our "bubble" safeguarded by illusion, since the good life needs to be lived within its borders.

Elizabeth Shaw (2024), similarly to other authors in this special issue, also makes Pereboom and Caruso's PHQM the focus of her analysis. Shaw argues that even being non-punitive, the PHQM should meet a high level of credibility, akin to the "beyond reasonable doubt" standard, since it allows for the intentional imposition of measures that are still coercive, severe, inescapably stigmatic and interfering with basic rights. In other words, as the author maintains, the harms and hardships that would be imposed on offenders in the name of the PHQM are not of a different kind than those permitted on the punitive criminal justice systems, which would allow to justifiably hold the PHQM to a significantly lower standard of credibility. Moreover, Shaw argues (in opposition to Caruso), that the PHQM cannot be exempted from the high epistemic standard based on allegedly involving merely foreseen, but unintended harm. The author points out that since Caruso claims that the PHQM involves "eliminative harm" and is grounded in the right to self-defense, it should follow that when people cause eliminative harm and act in self-defense, they only cause the foreseen, unintended harm. This is problematic, however, because Caruso relies on Victor Tadros's conception of "eliminative harm," which involves intentionality. To illustrate this, consider two cases presented by Shaw: (i) The defender hides her face behind a metal shield to protect herself from an attacker who tries to headbutt her. She may foresee but not intend that the attacker will ram his head against the shield and be harmed. (ii) The only way to prevent attacker from detonating a bomb and killing many people is to shoot the attacker before they can detonate. By trying their best to fire a lethal shot, the defender surely intends to kill the attacker. The PHQM is, according to Shaw, fairly often more like the defender in the second case - i.e., the PHQM often resorts to intended harm. Shaw also believes that Caruso cannot save the PHQM claim to be subject to a lower epistemic standard than punitive alternatives on the grounds that incapacitating a person constitutes intentional harm only if there is an intention to cause suffering. Imagine that Brian incapacitates Charlie to get his money, but does not intend Charlie to suffer. Now, Brian cannot deny that he intends to deprive Charlie of his liberty, and since depriving a person of their liberty constitutes a harm, Brian cannot claim that he did not intentionally harm Charlie. Thus, restricting "intentional harm" to "intentional suffering" does not seem to work. Moreover, the PHQM's special status cannot be justified by appealing to the doctrine of double effect, which states that intended harm is more difficult to justify than merely foreseen harm, not that intended suffering is more difficult to justify than other types of intended harm. Although Shaw criticizes the PHQM on points related to the epistemic argument, the author does not intend to undermine the whole model, but to encourage Pereboom and Caruso to creatively develop the PHQM by including the requirement of meeting a high standard of credibility.

Mirko Farina, Andrea Lavazza and Sergei Levin (2024) also focus on the PHQM, similarly offering its constructive criticism. The first major problem of the PHQM that the authors point out is that it must be able to assess, with a high degree of accuracy, the risk of recidivism. However, this is mostly a thing of the future; perhaps risk-assessment methods using big data, algorithms, and advanced statistics will be useful in this regard if they can be ethically justified, but for now the risk of re-offending is not predictable. For this reason, the PHQM's ability to achieve its own goals such as minimizing future crime, reducing harm, achieving economic viability, and offering appropriate rehabilitation to offenders is for the time being undermined. The second big deal for the PHQM is establishing effective deterrence for one-time offenders, e.g., violent avengers; under the PHQM, avengers would be able to exact their violent revenge without having to face the consequences associated with such actions in other criminal justice systems. In this way, the introduction of the PHQM could even create an incentive for acting on vengeful instincts. Farina, Lavazza and Levin also present three other objections according to which the PHQM may not fare better than (mixed) retributivist systems: depriving offenders of basic rights, genetic justice, and a problematic relationship with reasons-responsiveness. Regarding the first charge, by placing offenders under quarantine to protect the public, the PHQM deprives them of fundamental parts of their freedom, i.e., freedom of movement and association; in this way, the PHQM is not much better than the alternative, and the fact that the conditions under which detainees would be held under the PHQM would be more humane than in U.S. prisons does not change the moral assessment, since there are countries with humane and effective criminal justice systems based on mixed justifications for punishment. As for the second argument, there are individuals for whom a natural lottery has been unfavorable, and their genetic profile predisposes or inclines them to antisocial behavior; thus, it seems that subjecting such people to incapacitation or preventive quarantine would constitute

a certain type of "double injustice." The third allegation relates to the broader theoretical framework of the PHQM and points out that Pereboom and Caruso endorse the ability of most people to be reasons-responsive at the descriptive level, which makes it puzzling why their approach would be preferred over classical normative compatibilism. In conclusion, the authors encourage to treat these objections not as knock-down arguments, but rather as important issues that need to be resolved if the PHQM is to be considered a full-fledged and credible alternative to criminal punishment.

Acknowledgements: I would like to thank all the authors who submitted their texts, all the reviewers who helped to improve the papers, and the Editorial Team of *Diametros* for their continous help. Thank you all for making it possible to publish this special issue.

Funding: The editorial work on this special issue was funded by the Ministry of Science and Higher Education, Poland, Grant No. RCN/SP/0580/2021/1. The research for this article was funded by the National Science Centre, Poland, Grant No. 2022/47/B/HS1/02987.

Conflict of interest: The author has no conflict of interest to declare.

License: This is an open access article under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

References

- Brucato G., Jovchevsky P. (2024), "A Fairness-Based Defense of Non-Punitive Responses to Crime," *Diametros* 21 (79): 40–55.
- Farina M., Lavazza A., Levin S. (2024), "Pushing the Boundaries of the Quarantine Model: Philosophical Concerns and Policy Implications," *Diametros* 21 (79): 146–162.
- Jeppsson S.M.I. (2024), "Retributivism and the Objective Attitude," *Diametros* 21 (79): 56–73. Lemos J. (2024), "Free Will Skepticism, Quarantine, and Corrections," *Diametros* 21 (79): 107–118.
- Morris S.G. (2024), "The Non-Moral Basis for Eliminating Retributivism," *Diametros* 21 (79): 74–90.
- Shaw E. (2024), "Expanding the Scope of the Epistemic Argument to Cover Nonpunitive Incapacitation," *Diametros* 21 (79): 132–145.
- Smilansky S. (2024), "Free Will Denialism as a Dangerous Gamble," *Diametros* 21 (79): 119–131.
- Stachursky A. (2024), "Justice Without Retribution? The Case of the System of Communal Security, Justice and Reeducation of Montaña and Costa Chica in Guerrero, Mexico," *Diametros* 21 (79): 24–39.
- Vilhauer B. (2024), "Free Will Denial, Punishment, and Original Position Deliberation," Diametros 21 (79): 91–106.
- Zisman V. (2024), "The Missing Alternative Objection to Criminal Law Abolitionism," *Diametros* 21 (79): 10–23.