

THE NON-MORAL BASIS FOR ELIMINATING RETRIBUTIVISM

– Stephen G. Morris –

Abstract: While increasing numbers of philosophers have argued for eliminating the retributivist elements of criminal justice systems, their arguments often fall short due to internal inconsistency. Some of the best known of these arguments – such as those provided by Derk Pereboom and Gregg Caruso – rely on the claim that there are moral grounds for rejecting retributivism. In defending this claim, these philosophers typically provide arguments seeking to undermine the type of agent responsibility that they believe is needed to justify retributivism. This is usually followed by an assertion that since the excessive suffering caused by retributivist punishment cannot be justified, it is therefore immoral to preserve remnants of it in our criminal justice systems. This paper – which opposes the moral case against retributivism provided by Pereboom and Caruso, and favors non-moral reasons that recommend rejecting a retributivist approach – is divided into three parts. In part one, I discuss why Pereboom and Caruso are correct in thinking that free will skepticism undercuts the moral justification for retributivism. Part two sets out to explain why their moral case against retributivism fails insofar as it runs afoul of the folk intuitions that are called upon to defend Pereboom’s account of free will skepticism. In part three, I provide a non-moral case against the retributivism that is founded on considerations of self-interest and empathy.

Keywords: free will, skepticism, retributivism, criminal justice, empathy, self-interest, anti-realism, Pereboom

Submitted: 13 January 2023

Accepted: 31 May 2023

Published online: 28 December 2023

Introduction

While many philosophers have expressed concern about the implications of denying that human agents can exercise free will, free will skeptics have frequently brushed aside such concerns, with some even going so far as claiming that the widespread rejection of free will could have important social benefits. One of the areas that many of them believe would likely benefit from the widespread abandonment of free will is criminal justice. Some of the most interesting contemporary philosophical work on this subject is being carried out

Stephen G. Morris
College of Staten Island (CUNY)
Department of Philosophy
Bldg. 2N Room 232
2800 Victory Blvd, Staten Island, NY 10314
Email: Stephen.Morris@csi.cuny.edu

by two prominent free will skeptics, Derk Pereboom and Gregg Caruso, who argue that acknowledging the impossibility of free will among human beings provides the basis for developing systems of criminal justice that are more rational and just. On the basis of their skeptical views, these philosophers advocate for a criminal justice system that eschews retributivist elements in favor of one in which criminal behavior is approached in a manner that is analogous to that towards carriers of communicable diseases, which they call the *Public Health-Quarantine Model* (hereafter the *PHQM*). In a similar fashion to how curtailing the freedoms of disease carriers via quarantine is viewed as acceptable only when doing so is necessary for protecting the general populace, defenders of the PHQM argue that curtailing the freedoms of criminals is only justifiable on similar grounds. Skeptics like Pereboom and Caruso argue that this is the best approach to criminal punishment since criminals are no more in control of the factors that produce their behavior than disease carriers are for their afflictions. Furthermore, advocates of the PHQM affirm that the kind of emphasis on prevention we see in healthcare should steer our approach to dealing with criminal behavior. While I agree with Pereboom and Caruso regarding both the claim that free will skepticism (henceforth, simply *skepticism*) is true, and that the PHQM offers a superior approach to criminal justice than what we find in many retributivist-oriented societies, I deny the philosophical basis they provide for adopting the PHQM. That is, I reject their view that the denial of free will provides a *moral* basis for rejecting retributivist-based systems of criminal justice in favor of the PHQM. After explaining why their moral case for adopting the PHQM is inconsistent with the defense of skepticism given by Pereboom, I offer what I believe is a superior non-moral defense of the PHQM that appeals only to self-interest and empathy.

1. What Pereboom and Caruso Get Right: Retributivism Lacks Moral Justification

As I mentioned, I find the arguments for skepticism favored by Pereboom and Caruso to be convincing.¹ However, since this paper aims primarily to explore whether Pereboom and Caruso's skeptical stance — or, more accurately, their basis for it — is consistent with the case they make for eliminating retributivism, I will not be offering a defense of it. Instead, I will simply assume that the skeptics are correct in denying free will to human agents. A main question, then, is what specific implications does skepticism have for criminal justice? Like Pereboom and Caruso, I agree that one clear implication is that retributivism can have no moral justification in a world devoid of free will.² By "retributivism" I mean the justification for treatment whereby one is deemed a proper target of either reward or punishment on the sole basis of having performed either a morally good or morally bad decision/action, respectively. The administration of such rewards or punishments from this perspective is seen as an intrinsic good, and any justifications for treatment that are geared to bringing about positive outcomes in the future (such as less

¹ See, for instance, Pereboom (2001); Caruso (2012).

² In saying that *X* "lacks moral justification," I am only asserting that we lack moral reasons for accepting *X*. This is not meant to imply the stronger claim that we have moral reasons for rejecting *X*. Hence, saying that *X* lacks moral justification (i.e., is morally unjustified) is consistent with *X* being morally neutral.

bad behavior on the part of the punished individual) are irrelevant. Retributivist justice, furthermore, implies that the extent of the deserved reward or punishment corresponds to the relative goodness or badness of the action or decision in question. Thus, an act deemed to be extremely bad (such as murder) would warrant a harsher punishment than an act deemed to be slightly bad (such as cheating on one's taxes). Though much more could be said regarding the nature of retributivism, the foregoing discussion should be enough to inform the reader as to the basic meaning of the term and the general role that it plays in people's attitudes about reward and punishment.

It is worth pointing out that the definition of retributivism I have provided is more or less the standard one that has appeared in the philosophical literature for years, and it matches up very closely with other well-known conceptions of retributivism, including Alec Walen's description in his *Stanford Encyclopedia of Philosophy* entry for "Retributive Justice."³ With this understanding of retributivism in mind, it should be apparent how eliminating all retributivist aspects from our criminal justice systems would have a profound impact on how justice is carried out. Eliminating retributivism from legal considerations would have a powerful effect on, for instance, the justification for punishment (insofar as punishment can no longer be defended on the grounds of giving a criminal "his just deserts"), the specific types of punishments levied against those found guilty of a crime (since purely punitive measures would now appear to lack moral justification), and the very principles underlying criminal punishment (since one of the main traditional aims of criminal punishment would no longer be philosophically viable). In short, rejecting retributivism as a legitimate basis for inflicting punishment on criminals would require a significant reappraisal of what the purpose of criminal justice is and how it is best carried out.

To better understand the challenge that skepticism poses to retributivism, it will help to discuss the particular type of skepticism shared by Pereboom and Caruso in some detail. A good place to start will be to explain the specific notion of free will that they favor. Though there is a virtual consensus among philosophers that free will is best understood as being a necessary condition (more specifically the *control condition*) for moral responsibility, there is much disagreement about which type of moral responsibility is most relevant to the main philosophical disputes surrounding free will. In relation to free will, moral responsibility comes in two main varieties: *forward-looking* and *backward-looking*. To say that an individual is morally responsible in the forward-looking sense is to say, in essence, that the individual is a proper target for certain kinds of judgments, attitudes, etc., insofar as they are deemed necessary for bringing about positive future outcomes (e.g., they would make the individual and perhaps others more likely to abstain from improper behavior in the future). Another way to put this is that attributions of forward-looking moral responsibility are justified on *consequentialist* grounds. On the other hand, to say that one is morally responsible in the backward-looking sense implies that certain kinds of judgments, attitudes, or treatments towards the agent are justified simply in virtue of past decisions or actions that the agent performed or failed to perform. This sense of moral responsibility — which is closely associated with retributivism — is captured by the view that agents *deserve* certain types of responses on the basis of their past actions or decisions alone.

³ Walen (2020).

Both Pereboom and Caruso believe that a specific type of backward-looking moral responsibility, which Pereboom has labeled *basic desert moral responsibility*, is most relevant to free will. Pereboom describes it as follows:

For an agent to be morally responsible for an action in this sense is for it to be hers in such a way that she would deserve to be blamed if she understood that it was morally wrong, and she would deserve to be praised if she understood that it was morally exemplary. The desert at issue here is basic in the sense that the agent would deserve to be blamed or praised just because she has performed the action, given an understanding of its moral status, and not, for example, merely by virtue of consequentialist or contractualist considerations.⁴

While there has been some debate among philosophers as to what exactly basic desert moral responsibility entails, Pereboom has made it increasingly clear that it is tightly connected to retributivism. For instance, he says that of all the justifications for punishment, retributivism “is the one that most intimately invokes the basic desert sense of moral responsibility, together with the freedom it entails.”⁵ More recently, Pereboom has said that he takes a “more decisively retributivist characterization of negative basic desert than he has done previously.”⁶ On the basis of the connection he sees between retributivism and basic desert moral responsibility (BDMR), he asserts that “Classical retributivism would be undermined if the free will skeptic is right about [the lack of justification for] basic desert, since this view aims to justify punishment solely on the grounds of basic desert, and the skeptical position contends that we lack control in action required for punishment so conceived.”⁷

Taking stock, if Pereboom and Caruso are correct in thinking that: 1) free will is properly understood as being a necessary condition for BDMR; 2) free will of this type does not exist; and 3) BDMR is a necessary condition for the moral justification of retributivism; it follows that retributivism (along with the retributivist approach to criminal justice) is without moral justification. As I already mentioned, I will simply grant for the purposes of this paper that human beings are incapable of the kind of free will that the skeptic denies (the second point above). To keep things somewhat simple, I will also assume that Pereboom and Caruso are correct in thinking that retributivism can only be morally justified if BDMR is applicable to human agents. The question at this point, therefore, is whether we should accept Pereboom and Caruso’s recommendation to adopt the basic desert sense of moral responsibility as being most appropriate for disputes about free will, such that free will is properly conceived of as being a necessary condition for BDMR.

One of Pereboom’s most notable critics during the past two decades has been Manuel Vargas, whose revisionist approach to free will has challenged Pereboom’s attempt to situate basic desert at the center of the main disputes surrounding it. Vargas’s stance on free will also opposes Pereboom’s in that Vargas uses it to defend our

⁴ Pereboom (2014): 2.

⁵ Pereboom (2013): 52.

⁶ Pereboom (2021): 12; FN 12.

⁷ Ibidem: 81-82.

backward-looking retributivist practices. His revisionist stance is based on the claim that while the folk conceptions of free will and moral responsibility contain philosophically implausible (or at least highly contentious) elements such as a commitment to libertarianism, we can retain the essence of these concepts while dispensing with their problematic elements, since these are not integral to the concepts themselves. Vargas denies that BDMR is an essential feature of the folk conception of free will, and he says that he rejects “Pereboom’s conception of ‘basic desert’ if, by that, he means a substantive account of desert that requires that desert is warranted by nothing more than the agent and the action.”⁸ Vargas believes that it is enough for him to defend a kind of non-basic desert that considers not only the agent and the action, but also other factors that may include “facts including social context and the broader normative significance of the act in a system of human practices.”⁹ He believes that accommodating this type of desert within his revised account would be enough to ground a responsibility system that would “look very familiar, accommodating both backward-looking attitudes (such as gratitude) and forward-looking attitudes and practices.”¹⁰ At its heart, Vargas’s approach to free will is pragmatic in that it aims primarily to reinforce attitudes and practices (i.e., the “responsibility system”) that are most conducive to a well-functioning society. Since he believes that retributivism in general helps serve this purpose, he favors retaining terms like “free will” and “moral responsibility” in the folk lexicon since these labels are helpful for identifying individuals who are aptly subjected to both retributive and non-retributive forms of treatment that, in turn, promote the general welfare. In light of this justification of retributivism and of our responsibility practices in general, Vargas finds it unnecessary to defend the basic desert sense of moral responsibility that he is willing to concede is part of the folk conception of free will.

Pereboom has responded to Vargas by arguing that the type of revisionism that he favors – whereby the terms “free will” and “moral responsibility” are stripped of both libertarian and basic desert elements – veers too far from the folk conception of these terms to merit acceptance. His stance is that by changing the meanings of these key terms so drastically, Vargas and other revisionists of his ilk are likely to cause confusion in the eyes of the folk (and quite possibly other philosophers) who will interpret the revisionists to be defending the old folk concepts. He explains this using the following example: “If people started saying ‘he’s morally responsible for the murder since he did it of his own free will,’ but did not mean to claim that he in the basic sense deserved blame or punishment, then it could well be misleading to use the old terminology, since an audience might well be confused about which concept these words stand for.”¹¹ Since Pereboom believes that our common practices and attitudes invoking free will are anchored by BDMR, he is against any revisionist conception of the term that eschews basic desert. It is also important to him that philosophers adopt an understanding of free will that does not infringe on their ability to show us when our ordinary attitudes are mistaken. Since he believes that the folk take the existence of free will to provide

⁸ Vargas (2009): 56; FN 14.

⁹ *Ibidem*.

¹⁰ Vargas (2007): 158.

¹¹ Pereboom (2009): 26.

the primary philosophical justification for our retributivist attitudes, he worries that revising the term “free will” along the lines that Vargas suggests would make it more difficult to convince the folk that they ought to dispense with these attitudes as well as the problematic practices (e.g., retributivist criminal punishment) that the latter give rise to. For these and other reasons, Pereboom rejects revisionist attempts to extricate BDMR from what it means to exercise free will.

Pereboom’s case against revisionists like Vargas is bolstered by empirical evidence suggesting that BDMR (along with the retributivist attitudes attached to it) figures centrally in folk attitudes about free will. The work of Azam Shariff and his associates has provided particularly robust support for the view that backward looking, retributivist impulses – as opposed to forward-looking, consequentialist ones – are primarily responsible for eliciting increased attributions of free will for perceived immoral agents among test subjects.¹² Shariff and his group set out to clarify the kind of punishment associated with the folk concept of free will by exploring whether reducing people’s beliefs in free will would make them less retributive. They assumed that if folk attributions of moral responsibility are dependent upon attributions of free will, then we should find that people tend to be less retributive in their judgments about the kinds of treatment that an individual should receive insofar as they believe that the individual in question did not act freely. An implicit assumption these researchers were operating under is that the folk conception of moral responsibility is closely connected to deservingness of retributivist treatment. Put another way, the researchers assumed that the folk are operating under a conception of basic desert-based, as opposed to consequentialist-based, moral responsibility. The results of their studies suggest that this assumption is correct. In Study 1, they found that while stronger beliefs in free will predicted retributivist punishment, they were not predictive of consequentialist punishment. In Study 2, the researchers found that the subjects who had their beliefs in free will diminished by reading a scientific argument against free will recommended prison sentences that were roughly half of those that were recommended by subjects in a control group, suggesting that reducing the free will beliefs of the test group made them significantly less retributive than subjects in the control group. These results led Shariff and his colleagues to infer a tight connection between actions perceived as immoral and an inherent desire to inflict retributivist punishment upon agents committing them. Hence, they conclude that “Humans respond to transgressions with an urge to exact punitive costs on the transgressor.”¹³ Considering the data provided by the studies of Shariff and his colleagues, the burden of proof would appear to fall upon those who would deny that the folk concept of free will features a basic desert element as a central component.¹⁴

So far so good, then, for Pereboom and Caruso’s skeptical case against retributivism. Experimental work indicates that BDMR is, in fact, the primary kind of moral

¹² Shariff et al. (2014).

¹³ *Ibidem*: 1564.

¹⁴ Shariff et al.’s conclusions gain additional support through research conducted by Everett et al. (2021), which set out to explain why political conservatives tend to believe in free will more than political liberals. Their data suggest that conservatives’ higher belief in, and attributions of, free will are fueled by their desire to justify punitive attitudes toward perceived wrongdoers. I discuss the work of Everett et al. in more detail below.

responsibility at work in the folk concept of free will. This being the case, we have – based on arguments that Pereboom has offered – reasons for resisting the recommendations that philosophers like Vargas have made to revise the folk concept of free will away from its basic desert elements. Combining this with the assumptions that a) the kind of libertarian free will that is needed to provide the metaphysical and, hence, moral justification for basic desert is impossible (a position that revisionists like Vargas are willing to concede), and b) that BDMR is necessary for the moral justification for retributivism, we arrive at the conclusion that there is no moral justification for our retributivist attitudes and practices (including the retributivist model of criminal justice). But is this conclusion, assuming it is correct, enough to justify Pereboom and Caruso’s claim that we ought, therefore, to eliminate all remnants of retributivism from our systems of criminal justice? To see why a negative answer is warranted, we will need to explore the case that Pereboom offers for dispensing with retributivism in more detail.

2. What Pereboom and Caruso Get Wrong: The Alleged Immorality of Retributivism

At the end of the previous section, I explained how Pereboom and Caruso are well-positioned to defend the claim that moral reasons in favor of retributivism are lacking given the impossibility of libertarian free will. But to say that *there are no moral reasons supporting retributivism* is not to say that *there are moral reasons for rejecting retributivism*. Retributivism could simply be morally neutral. Furthermore, it could turn out that there are significant non-moral reasons, perhaps cashed out in terms of self-interest/prudence, for preserving retributivism. It is perhaps due to this realization that Pereboom attempts to strengthen the skeptic’s case against retributivism by stating, “If we are only moderately confident about the justification of a punishment, it would be morally wrong to administer it.”¹⁵ Caruso employs what is essentially the same premise in his “Epistemic Argument” against the retributive justification of legal punishment.¹⁶ This basic premise is central to Pereboom and Caruso’s position that while the PHQM of treating criminals is moral, retributivist models of punishment are not. The plausibility of this premise, however, is undermined by an argument that Pereboom relies upon to defend skepticism.

In the previous section, I explained how Pereboom’s attempt to defend skepticism against the revisionist challenge posed by Vargas relies on the claim that by seeking to extricate the basic desert sense of moral responsibility from the folk concept of free will, Vargas ends up with a notion of free will that strays too far from how the folk generally understand it. What’s of note here is that Pereboom believes the importance of preserving folk concepts in philosophical discussions extends beyond the free will debate. As he puts it, “More generally, when deciding how to revise, we need to retain concepts that facilitate our thinking that some of our attitudes and beliefs are mistaken.”¹⁷ The main point I wish to make here is that the same arguments that Pereboom offers against revising the folk concept of free will along the lines that Vargas suggests can be applied with just as much force against revising the folk concept of morality in the manner he

¹⁵ Pereboom (2021): 68.

¹⁶ Caruso (2018): 16.

¹⁷ Pereboom (2007): 203.

recommends given how the account of morality that he offers, much like the account of free will that Vargas favors, seeks to eliminate any connection that morality has to basic desert in the eyes of the folk. A serious question facing Pereboom, then, is whether he can consistently defend his version of morality given the kinds of arguments he offers against free will in general and Vargas-style revisionism in particular. In what follows, I will assume that we ought to accept the reasons Pereboom gives for rejecting Vargas-style revisionism, thereby supporting the kind of skepticism he and Caruso accept. Under this assumption, I will argue that consistency demands that we reject the account of morality that Pereboom and Caruso adopt. More specifically, my claim is that if Pereboom is correct in thinking that an appeal to folk intuitions provides reason for rejecting free will in any form, a similar appeal to folk intuitions recommends rejecting morality in any form given that skepticism is true. If this is correct, it would undermine the moral case that Pereboom and Caruso give for rejecting the retributivist model of punishment.

In his 2021 book, *Wrongdoing and the Moral Emotions*, Pereboom sets out to provide a detailed look at the type of moral outlook that he believes is consistent with skepticism. This account of morality is notable for being much more comprehensive than the one he provided some twenty years earlier in *Living Without Free Will*.¹⁸ It is also notable for how much more revisionist it is in terms of how the moral concepts he believes are open to the skeptic differ dramatically from how the folk understand them. To illustrate, let us begin with the notion of moral responsibility that Pereboom adopts. On page 36, he states that his conception of moral responsibility — as well as related concepts such as blame/praise and punishment/reward — is, for all intents and purposes, forward-looking. In an earlier passage, he is explicit about the revisionist features of this conception:

This proposal presupposes that there be distinct senses of moral responsibility in the actual [folk] practice, some to be eliminated, others retained. Reasons for maintaining that some are best eliminated are twofold. The first is that they can successfully be criticized on moral grounds, and the second that they can be challenged because they require the sort of control in acting we do not have.¹⁹

Pereboom goes on to mention that he is most concerned with offering a defensible account of moral responsibility that eliminates any remnant of basic desert. Hence, he says, “I propose a conception of blame as moral protest, a stance apt to engage and at times confront the wrongdoer with reasons to forbear, and conceived as forward-looking, not as backward-looking impositions of pain and harm a wrongdoer basically deserves.”²⁰ He acknowledges how this differs from the folk notion of moral responsibility when he says, “Retribution is an element of our ordinary human practice of holding each other, and ourselves, morally responsible.”²¹ As I discussed earlier, the idea that basic desert and retributivism are integral to the folk concepts of free will and moral responsibility has been given strong support by the studies conducted by Shariff and his colleagues.²²

¹⁸ Pereboom (2001).

¹⁹ Pereboom (2021): 28.

²⁰ Ibidem: 27.

²¹ Ibidem: 1.

²² Shariff et al. (2014).

Shariff et al. also discuss how their research suggests that the folk concept of moral blame shares these backward-looking features when they mention how “the mediational effect of perceived blameworthiness made a strong case for the role of moral responsibility in the effect of diminished free will belief on retribution.”²³

Another facet of Pereboom’s account of morality that appears significantly revisionist concerns the notion of *moral wrongness* that he employs. Most importantly for present purposes, he suggests that one can be morally wrong for an action without one being morally responsible for performing that action “in any sense essentially connected with culpability or blameworthiness.”²⁴ He uses an example of a psychopath who is unmoved by empathy as one who would be open to this type of judgment. Consistent with his effort to revise the folk understanding of moral responsibility, Pereboom also believes that moral wrongness should be extricated from any basic desert elements that it carries in the eyes of the folk. And yet recent research indicates that judgments of moral wrongness among the folk are unsurprisingly imbued with the same basic desert elements that we find in the folk concept of moral responsibility. In 2021, psychologist J.A.C. Everett and his colleagues set out to explain why political conservatives tend to believe in free will more than political liberals. They sought to test their hypothesis that it was conservatives’ greater tendency to moralize than liberals – i.e., to give moral weight to a larger set of actions and behaviors than liberals – that accounts for this phenomenon. They found that it was conservatives’ tendency to make more moral judgments (particularly judgments about *moral wrongness*) – mediated by a desire to hold agents blameworthy – that explained this difference among liberals and conservatives rather than political views or metaphysical beliefs about human autonomy. Based on their experiments, Everett et al. conclude that “People endorse the idea of free will in order to justify their desire to blame others for moral wrongdoing.”²⁵

Combining the research of Shariff et al. and Everett et al. suggests that folk moral judgments – including those about moral rightness and wrongness – are inseparable from free will, BDMR, and retributivist-based blame. Shariff and his colleagues found that folk judgments about free will are tied primarily to backward-looking moral responsibility and retributivist, as opposed to consequentialist, punishment. Everett and his associates helped further clarify the relationship between folk moral judgments, free will, and backward-looking moral responsibility by finding that judgments about moral wrongness were positively correlated to judgments about free will and a desire to blame immoral agents. It is important to clarify here that my point is not that there aren’t any forward-looking elements in the folk concept of morality, but merely that it contains fundamental backward-looking elements as well. Similarly, while Pereboom does not deny that the folk concept of free will shares compatibilist components, he asserts that it contains fundamental libertarian elements.²⁶ Because of this, he finds it wrongheaded

²³ Ibidem: 1567.

²⁴ Pereboom (2021): 62.

²⁵ Everett et al. (2021): 479.

²⁶ Compatibilists hold free will to be compatible with the truth of causal determinism. This view is in contrast to incompatibilism, which holds that free will is not possible if causal determinism is true. Causal determinism is, roughly speaking, the view that all events are necessitated by a combination of prior events and the laws of nature. Libertarians are incompatibilists who believe that humans can sometimes act freely. Hence, they reject causal determinism.

to attempt to revise the folk concept of free will in a way that eliminates these libertarian aspects. My argument against Pereboom's revised conception of morality is that if we adopt this line of reasoning, then we should find his attempt to revise the folk concept of morality equally wrongheaded. This is because two of the key aspects of Pereboom's moral perspective (i.e., that moral responsibility and moral blame lack any connection to basic desert or retributivism, and that one can be morally wrong without being morally responsible or morally blameworthy) are strongly at odds with folk moral intuitions.

Speaking to the issue of how to define terms that are central to important philosophical disputes (such as those relating to free will), Pereboom says, "Here is one principle for settling on definitions in contexts such as this one: select those that best distinguish contrasting positions that divide parties in the debate."²⁷ He goes on to say that since virtually everyone in the free will debate believes that determinism is compatible with forward-looking – but not backward-looking – moral responsibility, it makes sense to adopt an understanding of free will that is connected to backward-looking moral responsibility. I agree. But now notice that this logic also supports adopting a conception of morality that is based on the backward-looking, rather than the forward-looking, notion of moral responsibility when it comes to disputes between moral realists and moral anti-realists. This is because virtually all moral anti-realists are willing to allow the propriety of subjecting individuals to forward-looking praise/blame and reward/punishment. And given Pereboom's stance that neither backward-looking moral responsibility nor the type of morality that is based upon it are applicable to human agents, the preceding principle he recommends for settling on important definitions suggests that he ought to side with the moral anti-realists in rejecting morality altogether. This conclusion is bolstered by considering how the conception of forward-looking morality that he favors appears to violate another of the principles he accepts insofar as it deviates dramatically from how the folk conceive of morality.

In this section I have argued that the primary means by which Pereboom and Caruso have attempted to refute the retributivist basis of criminal punishment – namely, by arguing that it is morally wrong – is undermined by how the specific account of morality that Pereboom provides (and which Caruso also adopts) is inconsistent with other positions Pereboom has taken. Chief among these are the arguments he has offered in defending skepticism against the revisionist challenge presented by Vargas. In the end, it appears that both Pereboom and Caruso face the horns of a dilemma when it comes to their case against retributivism, one which stems from their reliance upon folk concepts to motivate their skepticism. On the one hand, they could continue arguing for the importance of retaining the folk understanding of key terms like free will and moral responsibility with their essential backward-looking elements. Since consistency appears to require that they also accept a notion of morality that shares these same backward-looking features – features that they do not believe exist – they would seem committed to rejecting any type of moral argument whatsoever, including that which holds that retributivism (and, therefore, the retributivist model of criminal punishment) is immoral. On the other hand, they could argue for revising folk moral

²⁷ Pereboom (2021): 11.

concepts in a way that eliminates their essential backward-looking attributes. Insofar as doing so would also likely provide the justification for revising the folk concept of free will, however, it would undermine the reasons that Pereboom has offered for why skepticism is preferable to Vargas-style revisionism. And if we are without a compelling reason to accept skepticism over alternative views of free will, Pereboom and Caruso's claim that retributivism ought to be rejected loses its force.

3. A Non-Moral Case Against Retributivism

In what preceded, I discussed how the moral case that Pereboom and Caruso attempt to levy against the retributivist model of criminal justice — and against retributivism in general — is at odds with Pereboom's response to Vargas's revisionist challenge. In this section, I describe what I take to be the best way for skeptics to challenge the retributivist model of criminal justice. On the one hand, I agree with Pereboom and Caruso's account of skepticism, partly because I believe Pereboom is correct in insisting that we ought to adopt notions of free will and moral responsibility that resonate with the folk. Given this reliance upon folk concepts, however, I believe that skepticism about free will commits us to rejecting the existence of morality altogether insofar as human agents and their actions are concerned. Hence, my rejection of the moral argument that Pereboom and Caruso provide against retributivism. Nonetheless, I believe we have substantial non-moral reasons, in the forms of self-interest and empathy, for rejecting the retributivist model of criminal punishment.

In discussing the main difference that he sees between his position on criminal punishment and that of revisionists Vargas and Daniel Dennett, Pereboom describes Vargas's and Dennett's defense of backward-looking attitudes and practices (including retributivism) as being "two-tiered."²⁸ By this he means that while both philosophers favor preserving these backward-looking elements of our society, including the retributivist model of criminal justice, their reasons for doing so are ultimately forward-looking. That is, while each is dubious about the existence of moral properties that can render human agents morally responsible in the basic desert sense, they nonetheless favor retaining certain backward-looking attitudes and practices because they believe that our society is better off with them insofar as they encourage people to avoid engaging in criminal activity, etc. Thus, at the basis of their arguments for preserving these backward-looking features is an empirical claim about the alleged benefits that these features bestow upon society. Instead of challenging Vargas and Dennett's defense of backward-looking attitudes and practices on the grounds that they are immoral, as Pereboom does, I believe the skeptic is better off simply challenging the claim that we are better off with them than without them. This would, if successful, undercut one of the major justifications for the retributivist model of criminal punishment without having to resort to dubious moral arguments. In what follows, I point to evidence suggesting that our society would be better off without the kinds of retributivist attitudes and practices that Vargas and Dennett defend, particularly with regard to how they impact criminal justice.

²⁸ Ibidem: 29.

To begin with, it should be apparent to anyone familiar with the criminal justice system in the U.S. how that system surpasses those of other industrialized democratic countries in terms of the role that retributivism plays in shaping it. This phenomenon has been identified by numerous criminal justice experts. For instance, a comparative study of the criminal justice systems in the U.S., Germany, and the Netherlands conducted by Subramanian and Shames observed that while the German and Dutch systems are based on the central tenant of “resocialization and rehabilitation,” the U.S. system is based on “incapacitation and retribution.”²⁹ In his 2014 book *Inferno*, law professor Robert Ferguson details how the natural human punitive instinct has come to dominate U.S. criminal policy in the form of its retributivist emphasis.³⁰ If an emphasis on retributivist attitudes and practices within a society was beneficial, as Vargas and Dennett contend, it stands to reason that the benefits should manifest themselves in that aspect of society that such attitudes and practices would appear to have the most direct impact upon; namely, the criminal justice system. This would suggest that a highly retributivist system, like the one we find in the U.S., should have more effective outcomes than less retributivist systems we find in other wealthy democracies like Canada and Western Europe. However, this is not what we find. Under the reasonable assumption that criteria such as less crime and cost-efficiency are proper yardsticks by which to assess the relative merits of a criminal justice system, there is little doubt that the U.S. model does not measure up to these other systems. To begin with, the U.S. system falls well short in terms of crime prevention according to a variety of measures. Holland observes that while the U.S. holds roughly five percent of the world’s population, it accounts for roughly twenty-five percent of the world’s prison population.³¹ The Vera Institute of Justice found that the U.S. incarceration rate is over eight times that of either Germany or the Netherlands.³² In terms of the monetary costs of maintaining the U.S.’s enormous prison population, journalist Casey Kuhn observes how “The U.S. spends \$81 billion a year on mass incarceration, according to the Bureau of Justice Statistics, and that figure might be an underestimate.”³³ In 2017, the Prison Policy Initiative estimated the actual cost on state and federal governments and impacted families is roughly \$182 billion.³⁴ The monetary inefficiency of the U.S. criminal justice system is apparent in an international analysis published by Graham Farrell and Ken Clark showing that the U.S. surpasses almost every other wealthy country in expenses per capita on policing and court costs, and spends well more than any other wealthy country per capita (over \$90) annually on prison expenditures.³⁵ With regard to recidivism, the U.S. criminal justice system also lags behind its less retributivist counterparts. As a comparison, while the U.S. has a recidivism rate of roughly 66% for all prisoners within three years of their

²⁹ Subramanian, Shames (2013): 7.

³⁰ More information about the retributive basis of the U.S. criminal system can be found in Caruso (2021).

³¹ Holland (2013).

³² Subramanian, Shames (2013).

³³ Kuhn (2021).

³⁴ Wagner, Rabuy (2017).

³⁵ Farrell, Clark (2004).

initial release (one of the highest among developed democracies), Norway's recidivism rate is roughly 20%.³⁶

The fact that the highly retributivist criminal justice system found in the U.S. produces outcomes that are far worse than what we find in other less retributivist-oriented systems lends support to the claim that societies are better off to the extent that retributivist attitudes and practices are minimized or eliminated altogether. Nonetheless, one might object that this disparity in outcomes may have less to do with the U.S.'s emphasis on retributivism than with other significant differences that we find between them and other nations (e.g., less robust social welfare programs, higher economic inequality, etc.). The claim here is that the wide variety of socio-political differences we find between the U.S. and other industrialized democracies undermines the confidence with which we can attribute the U.S.'s criminal justice shortcomings to its retributivist emphasis. To counter this line of reasoning, it will be instructive to compare retributivist and non-retributivist approaches to criminal justice within particular countries. Consider, for instance, a study by Charles M. Borduin and his colleagues indicating that a non-retributivist approach involving multisystemic interventions for U.S. juvenile offenders reduced recidivism rates to just above 20 percent in comparison to the baseline rate of about 70 percent that existed around that time for juvenile offenders who were exposed to the typical retributivist-based incarceration methods.³⁷ Additional research indicates that a non-retributivist approach involving behavioral-cognitive therapy in the U.S. resulted in a 25 percent decrease in recidivism.³⁸ That non-retributivist approaches to criminal justice yield better outcomes than more retributivist forms of punishment such as incarceration is apparent outside the U.S. as well. In recent years, Norway has received much attention for both the humane treatment of its prisoners as well as the positive results this approach has generated. And yet this emphasis on prisoner welfare was not always a hallmark of the Norwegian criminal justice system. In a BBC News article, the governor of Norway's Halden Prison remarks about how the Norwegian Correctional Service underwent a series of reforms aimed at veering away from a corrections system founded on "revenge" (i.e., retributivism) to one focused primarily on rehabilitation. The result? The recidivism rate in Norway dropped from around 60-70 percent (what we currently find in the U.S.) to its current 20 or so percent.³⁹ The fact that an emphasis on retributivism appears to render a criminal justice system less effective – whether we are comparing more retributivist to less retributivist countries or comparing how greater and lesser retributivist policies have fared within a given country – lends considerable support to the benefits of minimizing the influence that backward-looking and specifically retributivist sentiments have within a society.

To summarize the case I have presented thus far against preserving retributivist attitudes and practices (particularly as they influence criminal justice systems), I have argued that in addition to reasons for rejecting retributivism that are based on the denial that human agents can be morally responsible in the basic desert sense – which, as Per-

³⁶ Bennechi (2021).

³⁷ Borduin et al. (1995).

³⁸ Landenberger, Lipsey (2005).

³⁹ BBC News (2019).

eboom points out, is needed to justify the primary moral argument that retributivists rely upon (i.e., that agents truly *deserve* harsh treatment under certain circumstances) – the skeptic can now point to empirical evidence that calls the effectiveness of the retributivist model of criminal justice into question. Assuming that the retributivist is unable to appeal either to basic desert or to beneficial outcomes to support retributivism, it appears that she is left without any compelling reasons to preserve their preferred model of punishment.

Let me clarify here that in stating that retributivism tends to have a negative impact on society via its influence on criminal justice, I am not thereby providing a *moral* argument against it. What I have provided is a reason for rejecting retributivism that stems from my self-interested desire to live in a society that is best suited for allowing me to flourish. This itself seems a sufficient reason for why I and others should advocate for the rejection of retributivism, with no additional moral argumentation being necessary. Furthermore, my interest in living in a retributivist-free society is made even stronger by the *empathy* I feel towards members of my community – both criminals and non-criminals alike. Given that less retributivist criminal justice systems seem more effective in reducing crime, I find them appealing insofar as it is less likely that other members of my community will fall prey to crime under them. In addition, my skepticism about free will compels me to empathize with the plight of criminals. Such empathy is appropriate insofar as I share the perspective taken by Pereboom and Caruso comparing criminals to carriers of dangerous communicable diseases via the PHQM. From this perspective, criminals are ultimately no more in control of their damaging behavior than carriers of serious diseases are of their damaging illnesses. This being the case, I am drawn to a system of dealing with criminals that does not seek to harm them any more than does a public health system that seeks to quarantine disease carriers in order to serve the public good.

This notwithstanding, one might still insist that my case against retributivism constitutes a moral argument insofar as the self-interested (including empathetic) reasons I have for rejecting retributivism fall within a broad definition of consequentialist ethics. In denying this, I point to the reasons already mentioned for why key philosophical terms should not stray too far from their meanings in the folk lexicon. It was on the basis of these reasons, which Pereboom and Caruso both accept when defending their account of skepticism, that I argued that it would be a mistake for either of them to assert the existence of moral truths. Likewise, it would be a mistake to view my rejection of retributivism as having a moral basis since I deny several of the fundamental components that characterize folk morality, such as BDMR and backward-looking moral blame. If anything, attempting to pigeonhole my case against retributivism into a moral framework would be even more unwarranted than accepting the account of morality put forth by Pereboom since I, unlike him, deny that human agents can be morally responsible, morally blameworthy, or morally wrong in any senses whatsoever. The fact that my ontological commitments reject the existence of such properties speaks against all attempts to interpret any implications deriving from them in a moral light.

Before concluding, it is worth considering a possible objection to my view. Based on the previous data I mentioned suggesting that more retributivist criminal justice systems have worse outcomes than less retributivist ones, I maintain that we have good reason to believe that eliminating the retributivist elements from a criminal justice system

would likely be beneficial given that self-interest and empathy alone are guiding our decisions. In response, one may argue that preserving some elements of retributivism can be defended on the grounds that eliminating all retributivist attitudes and practices from society (assuming this were even possible) is likely to result in net negative outcomes, such as an increase in criminal activities. Whether this would occur is obviously open to debate. That said, the burden would appear to be on my opponent to provide a compelling case for why retaining some elements of retributivist punishment would bring about better results. The reason is that from a purely self-interested/empathic perspective, I (and presumably others who share this perspective) do not want to see people suffer unless it is necessary. Given this, justifying the suffering of a convicted criminal requires the same kind of hard empirical evidence that is needed to justify restricting the freedoms of individuals who contracted a dangerous communicable disease through no ultimate responsibility of their own. Furthermore, it would not be enough to establish that retributive punishment has *some benefits* over a system with no retributivist elements whatsoever. My opponent would have to provide a compelling case that retaining some retributivist elements in a criminal justice system, for instance, would have *greater net* benefits than eliminating them completely. And there are good reasons to think that any such case by my opponent is unlikely. The fact that other developed nations have placed much less of an emphasis on retribution without crime rates rising to anything close to what we see in the U.S. suggests that there are more effective ways of eliminating crime than relying on the belief that criminals are proper targets of retributivist punishment. Furthermore, it is likely that retributivist attitudes in the U.S. account for much of its criminal justice system's failings (e.g., high rates of recidivism), which stem from a desire to punish criminals for its own sake, and which are opposed to more effective measures such as rehabilitation. Finally, while one may point out that retributivist elements can be found in all known criminal justice systems today (even among the least retributivist systems such as the one in Norway), such elements are more likely due to their being remnants of less enlightened times or to the natural retributivist bent of the human psyche than to their tendency to promote social well-being.

4. Conclusion

In this paper I have argued that the primary way that skeptics like Pereboom and Caruso have attempted to refute retributivism – namely, by arguing that it is immoral – has been unsuccessful. While skepticism about free will does provide reason for rejecting the claim that retributivists attitudes and practices are morally right, they cannot establish the stronger claim that they are morally wrong. This is because the reliance upon folk concepts that skeptics need to defend their rejection of free will prevents them from accepting the highly revisionist account of morality that skeptics like Pereboom attempt to employ against retributivism.⁴⁰ I have argued that a more effective skeptical

⁴⁰ While I have yet to come across Pereboom labeling his account of morality as “revisionist,” this label seems applicable given how several of its key elements – including moral responsibility, moral blame, and moral wrongness – are given meanings that differ significantly from the folk understanding of these terms. In light of this, it seems more appropriate to say that he offers a revisionist account of morality *in general* than to say merely that he offers a revisionist account of moral responsibility, etc.

case against retributivism can be made by pointing out that there are strong empirical reasons for believing that retributivist attitudes and practices, insofar as they give rise to highly retributivist criminal justice systems, have a negative impact upon society. Such reasons not only undermine one of the main justifications for retributivism, but also negate the arguments of free will revisionists like Vargas and Dennett who maintain that the net benefits provided by retributivist attitudes warrant preserving the belief that human agents can act freely. By relying on the argument that retributivism goes against individual self-interest, the case for which is made even stronger for those who have empathy for both criminals and non-criminals alike, the skeptic can consistently provide reasons for rejecting both free will and our retributivist attitudes and practices, including the retributivist model of criminal justice.

Acknowledgments: I would like to thank two anonymous referees and especially Przemysław Zawadzki for providing helpful feedback during the writing of this article.

Funding: This article received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

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

- BBC News (2019), "How Norway Turns Criminals Into Good Neighbours," URL = <https://www.bbc.com/news/stories-48885846> [Accessed 4.05.2023].
- Bernechi L. (2021), "Recidivism Imprisons American Progress," *Harvard Political Review*, URL = <https://harvardpolitics.com/recidivism-american-progress/> [Accessed 15.11.2022].
- Borduin Ch.M., Mann B.J., Cone L.T., Henggeler S.W., Fucci B.R., Blaske D.M., Williams R.A. (1995), "Multisystemic Treatment of Serious Juvenile Offenders: Long-Term Prevention of Criminality and Violence," *Journal of Consulting and Clinical Psychology* 63 (4): 569–578.
- Caruso G. (2012), *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will*, Lexington Books, Plymouth.
- Caruso G. (2018), "Justice Without Retribution," *Neuroethics* 13 (1): 13–28.
- Caruso G. (2021), *Rejecting Retributivism*, Cambridge University Press, Cambridge.
- Everett J., Clark C., Meindl P., Luguri J., Earp B., Graham J., Ditto P.H., Shariff A. (2021), "Political Differences in Free Will Belief Are Associated With Differences in Moralization," *Journal of Personality and Social Psychology* 120 (2): 461–483.
- Farrell G., Clark K. (2004), "What Does the World Spend on Criminal Justice?," *HEUNI Paper Series No. 20*, European Institute for Crime Prevention and Control, Helsinki.
- Ferguson D. (2014), *Inferno*, Harvard University Press, Cambridge.

- Holland J. (2013), "Land of the Free? US Has 25 Percent of the World's Prisoners," *Moyers and Company*, URL = <https://billmoyers.com/2013/12/16/land-of-the-free-us-has-5-of-the-worlds-population-and-25-of-its-prisoners/> [Accessed 6.08.2014].
- Kuhn C. (2021), "The U.S. Spends Billions to Lock People Up, But Very Little to Help Them Once They're Released," *PBS News Hour*, URL = <https://www.pbs.org/newshour/economy/the-u-s-spends-billions-to-lock-people-up-but-very-little-to-help-them-once-theyre-released> [Accessed 11/17/22].
- Landenberger N.A., Lipsey M.W. (2005), "The Positive Effects of Cognitive-Behavioral Programs for Offenders: A Meta-Analysis of Factors Associated with Effective Treatment," *Journal of Experimental Criminology* 1 (4): 451-476.
- Pereboom D. (2001), *Living Without Free Will*, Cambridge University Press, New York.
- Pereboom D. (2007), "Response to Kane, Fischer, and Vargas," [in:] *Four Views on Free Will*, J.M. Fischer, R. Kane, D. Pereboom, M. Vargas (eds.), Blackwell Publishing, Malden, MA: 191-203.
- Pereboom D. (2009), "Hard Incompatibilism and Its Rivals," *Philosophical Studies* 144 (1): 21-33.
- Pereboom D. (2013), "Free Will Skepticism and Criminal Punishment," [in:] *The Future of Punishment*, T.A. Nadelhoffer (ed.), Oxford University Press, Oxford: 49-78.
- Pereboom D. (2014), *Free Will, Agency, and Meaning in Life*, Oxford University Press, Oxford.
- Pereboom D. (2021), *Wrongdoing and The Moral Emotions*, Oxford University Press, Oxford.
- Schmitt J., Warner K., Gupta S. (2010), "The High Budgetary Cost of Incarceration," *Center for Economic and Policy Research*: 1-15.
- Shariff A. F., Greene J. D., Karremans J. C., Luguri J., Clark C. J., Schooler J. W., Baumeister R., Vohs K. (2014), "Free will and Punishment: A Mechanistic View of Human Nature Reduces Retribution," *Psychological Science* 25 (8): 1563-1570.
- Subramanian R., Shames A. (2013), *Sentencing and Prison Practices in Germany and the Netherlands: Implications for the United States*, Vera Institute of Justice, New York.
- Vargas M. (2007), "Revisionism," [in:] *Four Views on Free Will*, J. Fischer, R. Kane, D. Pereboom, M. Vargas (eds.), Blackwell Publishing, Malden, MA: 126-165.
- Vargas M. (2009), "Revisionism About Free Will: A Statement and Defense," *Philosophical Studies* 144 (1): 45-62.
- Wagner P., Rabuy, B. (2017), "Following the Money of Mass Incarceration," *Prison Policy Initiative*, URL = <https://www.prisonpolicy.org/reports/money.html> [Accessed 23.10.2023].
- Walen, A. (2020), "Retributive Justice," *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/entries/freewill/> [Accessed 23.10.2023].