



# 25

## Against Legal Punishment

Nathan Hanna

### 1 Introduction

Here is a preliminary statement of my thesis: legal punishment is morally wrong because it is too morally risky. I will revise the preliminary statement later in response to objections, but for now it gives a good sense of the position that I will defend on the morality of punishment.<sup>1</sup> In this introductory section, I will briefly explain how my argument differs from similar ones in the philosophical literature on punishment. Then, in the rest of the chapter, I will explain why punishment is morally risky, argue that it is too morally risky, and discuss objections.

I am not the only one who thinks that punishment is morally risky or that it is wrong. The former view is increasingly popular and the latter—called *abolitionism*—has a few defenders. Those who share my worry about moral risk are rarely abolitionists, though. They typically argue that we should reform how we punish or reject certain justifications for punishment (e.g., Pereboom 2001, 161; Vilhauer 2009; Gross 2012, 9–14; Tomlin 2013, 2014; Huemer 2018, 15–16; Kolber 2018; Caruso 2020; Caruso and Pereboom

---

<sup>1</sup> For ease of exposition, I will use *punishment* to mean *legal punishment* throughout.

---

N. Hanna (✉)  
Drexel University, Philadelphia, PA, USA  
e-mail: [nth34@drexel.edu](mailto:nth34@drexel.edu)

2020; Jeppsson 2021).<sup>2</sup> And most abolitionists do not appeal to moral risk. The few who do focus on just a single source of it (e.g., Roebuck and Wood 2011). By contrast, I will argue that punishment is morally risky in a variety of ways that combine to make a good case for abolitionism. Most abolitionists try to show that extant justifications for punishment all fail or that a necessary condition for its permissibility cannot be satisfied (e.g., Sayre-McCord 2001; Golash 2005; Boonin 2008; Zimmerman 2011). These strategies have important limitations. The first does not show that better justifications cannot be developed. And the second invariably relies on controversial claims about why the relevant condition cannot be satisfied. By contrast, I will argue that there cannot be a successful justification for punishment. And I will do it by appealing to fairly uncontroversial claims about our epistemic fallibility, the badness of wrongful punishment, and what the necessary conditions for punishment's moral permissibility might be.

## 2 Punishment Is Morally Risky

In this section, I will argue that punishment is morally risky in a variety of ways. That is, I will argue that the moral risk involved in punishment has multiple sources. I will not say anything about *how* morally risky punishment is until the next section.

We can start to get a sense of punishment's moral risks if we think about what the necessary conditions for its moral permissibility are. Here are some plausible candidate conditions:

- The punishee [P] must have broken the law.
- P must have acted freely.
- P must have acted wrongly.
- P must have acted culpably.
- P must be liable to punishment.

With the possible exception of the liability condition, which I will clarify later, I expect that readers will have a good intuitive grasp of what these conditions mean.<sup>3</sup> The conditions are also fairly uncontroversial and

---

<sup>2</sup>Tomlin (2014, 444–46) and Kolber (2018, 491) do anticipate arguments like mine, though.

<sup>3</sup>The key concepts in conditions 2–5 typically have both legal and nonlegal senses. I am using the nonlegal senses. So, to take an example, P may not have acted freely in my sense even if the law counts her as having acted freely. Ditto for conditions 3–5. It is uncontroversial that there are nonlegal necessary conditions for punishment's permissibility and that the law can be mistaken about whether they are satisfied.

consistent with many different theoretical commitments, including different moral theories and different justifications for punishment. Of course, some theorists will want to reject or even add some conditions. I will discuss the significance of this later. But for now, I will just assume that the list is a good enough approximation of the truth. To start: the list highlights the fact that there are multiple conditions that must be satisfied for punishment to be permissible. This is important because more conditions tend to increase the likelihood that we will be mistaken whenever we think that all of the necessary conditions are satisfied.<sup>4</sup> This risk of epistemic error makes punishment morally risky because the relevant errors can lead to wrongful punishment. Each condition on my list is associated with specific risks of error. I will illustrate by discussing each condition, starting with the ones that involve the most familiar risks of error: the lawbreaking and wrongdoing conditions.

We sometimes falsely believe that someone broke the law. This can be due to false beliefs about what they did or about what the law forbids. The causes of these errors are familiar, such as misleading evidence and ignorance of the law. Similarly, we sometimes falsely believe that certain acts are wrong. This can be due to false beliefs about what makes acts wrong in general or false beliefs about whether the things that do so are present in a given case. The causes of these errors are also familiar, for example, bias and bad moral education. These specific errors can and do lead us to punish wrongfully. Many punishments are morally risky partly because we inflict them despite the risk that we are making these errors.

That said, it is plausible to think that we do not always risk making these errors when punishing. Sometimes there is just no denying that someone wrongfully broke the law—or so I will assume. To keep my argument simple, I will focus on risks of error that we run whenever we punish. I mention the above risks for two reasons. First, their familiarity makes pointing to them the clearest way to show that punishment can be morally risky. Second, their similarities to the risks of error that I will focus on will support my claims about the latter. Both the familiar and the unfamiliar risks of error stem from our epistemic fallibility in the face of what are often complex matters of fact and value. And both generate moral risk. The free will, culpability, and liability conditions are all associated with risks of error that we run whenever we punish. I will discuss these conditions in turn.

---

<sup>4</sup>Note that each condition on my list can fail to be satisfied when the ones preceding it are satisfied. To illustrate, someone can break the law freely and wrongly but non-culpably, e.g., because they were non-culpably ignorant of morally relevant facts. Or someone might not be liable despite satisfying the other conditions, e.g., because they freely and culpably committed a crime that is really just a private or minor wrong that should not have been criminalized.

Consider the free will condition: P must have acted freely. Most of us believe that free will is real. But what it takes to have free will and whether we have it at all are hard questions that are subjects of longstanding philosophical debate.<sup>5</sup> This is not the place to rehearse these debates. What matters for my purposes is that they are complex and the best arguments on each side are sophisticated and hard to evaluate. Given our epistemic fallibility, we should not be completely confident in our views here. Even those of us who are confident that we have free will should admit that there is a possibility that the free will deniers are correct. (I will say more about the sense of *possibility* that I am using later.) This possibility makes punishment morally risky to at least some extent. Whenever we punish people, it is possible that we are punishing people who lack free will and so punishing wrongly.<sup>6</sup> Similar points apply to the other conditions.

Take the culpability condition: P must have acted culpably. Most of us believe that people are typically culpable for their wrongdoing. But again, what it takes to be culpable and whether we ever are culpable are hard questions that are subjects of longstanding philosophical debate. These debates are also complex and the best arguments on each side are sophisticated and hard to evaluate. Some philosophers worry that exculpating factors such as non-culpable ignorance and the absence of relevant kinds of control are ubiquitous (e.g., Nagel 1979, 35–38; Strawson 1994; Zimmerman 1997, 2002, 2011; Rosen 2003, 2004; Levy 2011).<sup>7</sup> It is possible that our belief in culpability is mistaken and that the skeptics' doubts are correct. This possibility also makes punishment morally risky: whenever we punish people, it is possible that we are punishing non-culpable people and so punishing wrongly.

Finally, consider the liability condition: P must be liable to punishment. Before getting to the moral risks here, some clarifications are in order because it may not be obvious what this condition means. I take the claim that P is liable to punishment to mean that certain facts that used to be moral reasons not to punish P either no longer hold or are no longer reasons not to punish

---

<sup>5</sup>Prominent contemporary free will deniers include Pereboom (2001) and Levy (2011). For an overview of arguments against free will, see O'Connor and Franklin (2021).

<sup>6</sup>One might object that, if we lack free will, punishment is never wrong because none of our acts are ever wrong. I will reply to this objection later.

<sup>7</sup>Levy (2011) equates these kinds of control with free will, but I think that free will is a distinct kind of control and not the only kind that might be necessary for moral responsibility. Zimmerman (2011, 144–50) argues that we can be both culpable and *inculpable* (his term) for our acts and that punishing us for our acts is permissible only if we are culpable for them and not inculpable for them. For my purposes, we do not need to worry about what inculpability is. Instead, we can take Zimmerman to be endorsing an additional necessary condition on punishment's permissibility. I will discuss the significance of possible additional conditions below. For an overview of arguments against moral responsibility, see Caruso (2021).

P. Such facts include the fact that punishing people is normally a violation of their rights and the fact that it intentionally harms them. Some theorists argue that people can forfeit their rights and that, when they do, punishment does not violate their rights. And some argue that people can deserve to be harmed or punished and that, when they do, the fact that punishment will harm them is not a reason not to punish them and may even be a reason to punish them.<sup>8</sup> For convenience, I will understand liability in terms of rights forfeiture and desert. But what I say about it will generalize to many other ways of understanding it.

Here is why the liability condition, so understood, makes punishment morally risky: As with the other conditions, whether this condition is ever satisfied turns on the answers to hard questions that are subjects of longstanding philosophical debate. These questions include questions about what rights people have, what the basis of our rights are, what people deserve, and what the normative significance of desert is. The debates over these sorts of questions are complex and the best arguments on each side are sophisticated and hard to evaluate. Some philosophers question whether people can forfeit their rights against punishment, some question whether people can deserve to be harmed or punished, and some question whether deserving to be harmed or punished can eliminate reasons not to harm or punish.<sup>9</sup> It is possible that the skeptics' doubts about these things are correct. This possibility makes punishment morally risky: whenever we punish people, it is possible that we are punishing people who are not liable to punishment and so punishing wrongly.

I started this section by listing some fairly uncontroversial candidate conditions for punishment's permissibility. I have argued that each of them makes punishment morally risky. But when I listed the conditions, I said that some theorists will want to reject some of them or even add to them. I will conclude this section by briefly discussing the significance of these reactions.

First, consider the position of those who want to reject some of the conditions on my list. I am willing to grant for argument's sake that such critics might be right and that any given condition on my list might not be a genuine condition, even approximately. But this is not a problem for me because I do not need to insist that these conditions are genuine to establish my claims about moral risk. The mere possibility that they are genuine and never satisfied suffices for moral risk. To illustrate, it is possible that the following claim

---

<sup>8</sup> Wellman (2012) defends the first claim. Berman (2008) defends both claims. For an interestingly different take on liability, see Tadros (2011).

<sup>9</sup> Boonin (2008, 103–19) and Hanna (2012, 609–16) question the first position. Hanna (2019, 111n6) gives several examples of philosophers who question the second. Hanna (2013, 2019) and Nelkin (2019) question the third.

is true: punishing people is permissible only if they acted culpably and no one ever acts culpably. The mere possibility that this is true makes punishment morally risky to at least some extent.

Next, consider the position of those who want to add to my list. Potential additions include: punishing P is the only way to achieve goods like deterrence and is the only way to express adequate disapproval of wrongdoing. For my purposes, there is nothing special about these conditions. Everything that I have said about the ones on my list applies to these new ones. It is possible that these are—or at least that they approximate—genuine conditions and that the theorists who think that they are never satisfied are correct.<sup>10</sup> My point here is not just to identify more necessary conditions that are sources of moral risk, though. My point is more general: no matter how confident we are in our views about what it takes for punishment to be permissible, it is always possible that we have overlooked some necessary conditions that may not be satisfiable. This adds to punishment's moral risks.<sup>11</sup>

I have argued that punishment is morally risky in a variety of ways. But I have not yet said anything about how morally risky it is. For all that I have said, the risks might be minimal. In the next section, I will argue that they are not and that punishment is morally wrong because it is too morally risky.

### 3 Punishment Is Too Morally Risky

Here is my main argument. Call it the *Risk Argument*.

1. Punishment is permissible only when we have shown the following beyond reasonable doubt: all of the conditions that are necessary for its permissibility are satisfied.
2. We can never do that.
3. So, punishment is wrong.

Before I defend the premises, I should emphasize that this argument is about a *conjunction* of claims, specifically the conjunction of all of the conditions (whatever they may be) that are necessary for punishment's permissibility. The argument says that this conjunction—*not* its individual conjuncts—must be demonstrated beyond reasonable doubt and cannot be demonstrated beyond

<sup>10</sup> For arguments that punishment is not the only way to do such things, see, e.g., Sayre-McCord (2001, 514–16), Golash (2005, 22–48, 153–72), Boonin (2008, 264–67), and Hanna (2008, 2014).

<sup>11</sup> Another upshot is that particular justifications for punishment are often committed to conditions that would make punishment morally risky. This makes such justifications vulnerable to distinctive moral risk-based objections. For an argument that this is true of retributive justifications, see Kolber (2018).



reasonable doubt. The argument does not say that we must demonstrate beyond reasonable doubt that someone acted freely (or that someone acted culpably, or that someone is liable to punishment) and that we cannot demonstrate this beyond reasonable doubt.<sup>12</sup> The argument is consistent with the view that these claims often cannot be reasonably doubted. The argument exploits the fact that a conjunction of claims can sometimes be reasonably doubted even if its individual conjuncts cannot be. This is true because the probability that a conjunction is true can be much lower than the probability of each conjunct. I will illustrate.

Suppose that I can reasonably doubt a claim only if I judge that it is less than 99 percent likely to be true. And suppose that I judge six claims to each be 99 percent likely to be true. Even though I cannot reasonably doubt them individually, I can still reasonably doubt their conjunction because I am rationally committed to judging it to be 94.1 percent likely to be true.<sup>13</sup> This illustration is artificial, partly because the numbers are made up and partly because reasonable doubt is arguably a qualitative rather than quantitative concept. But none of this affects my point, which is that a conjunction of claims can sometimes be reasonably doubted even if its individual conjuncts cannot be. Premise 2, which I will defend below, says that the conjunction of all the conditions that are necessary for punishment's permissibility can always be reasonably doubted.

Now I will defend the Risk Argument's premises. Here is my argument for 1:

4. Punishment is permissible only when we have shown the following beyond reasonable doubt: the punishee broke the law.
5. If premise 4 is true, then punishment is permissible only when we have shown the following beyond reasonable doubt: all of the conditions that are necessary for its permissibility are satisfied.
6. So, punishment is permissible only when we have shown the following beyond reasonable doubt: all of the conditions that are necessary for its permissibility are satisfied.

Premise 4 is widely accepted. Here is a paraphrase of the standard rationale for it (Tomlin 2013, 48–52; Tomlin 2014, 434–35; Huemer 2018, 16):<sup>14</sup>

<sup>12</sup>In this respect, the argument differs from those of Vilhauer (2009), Caruso (2020), and Caruso and Pereboom (2020), among others. They argue that the claim that we have free will can be reasonably doubted and that this undermines retributive justifications for punishment specifically.

<sup>13</sup>Assuming that the probabilities for the six claims are independent of each other, the calculation is straightforward:  $0.99^6 = 0.941$  (Kolber 2018, 490).

<sup>14</sup>For arguments against the rationale, see Laudan (2006, 2011, 2012). For a critique of Laudan's arguments, see Gardiner (2017).

Punishing people who have not broken the law is wrong, and our beliefs about whether people broke the law are fallible. This makes punishment morally risky. And this risk is a comparatively serious one: other things equal, wrongful punishment is much worse than wrongful non-punishment. 4 is true because using the reasonable doubt standard is necessary to appropriately balance these risks.

I will assume that this rationale for premise 4 is correct. Premise 5 is true because there is no relevant difference between the claim that the punishee broke the law and the claim that all of the conditions that are necessary for punishment's permissibility are satisfied. As others have observed, the rationale for premise 4 generalizes. It applies to any claim that is like the lawbreaking condition in relevant respects. That is, it applies to claims that state necessary conditions on punishment's permissibility and about which we are fallible.<sup>15</sup> The conjunction of all of the necessary conditions is like the lawbreaking condition in these respects, so the same rationale applies to it.

Before moving on, I should pause to address a potential misunderstanding. I have argued that punishment is permissible only when we have shown beyond reasonable doubt that all of the conditions that are necessary for its permissibility are satisfied. This appeal to reasonable doubt is liable to be misunderstood. The concept of reasonable doubt is most popularly associated with a legal rule of evidence that is applied by juries or judges in criminal trials. Because of this, some readers might think that I am arguing that punishment is permissible only if we have shown the above during a criminal trial. And these readers might complain that it is unreasonable to expect juries and judges to grapple with all the difficult philosophical issues on which, I have argued, punishment's permissibility depends.

In response, I am not sure that it would be unreasonable to expect juries and judges—or anyone with the power to decide whether someone will be punished—to grapple with these issues before deciding whether someone will be punished. But I do not have to insist on this. When I talk about reasonable doubt, I am not talking about a legal rule of evidence. I am talking about *an epistemic standard* that can be applied in any context. So understood, my claim that punishment's permissibility requires showing certain things beyond reasonable doubt does not entail that this must be done by juries or judges

---

<sup>15</sup> Huemer (2016, 16) applies the reasoning to the wrongdoing condition. Tomlin applies it to the claim that some conduct is worthy of punishment (2013, 45, 52) and to the claim that a given punishment is not disproportionately harmful (2014, 432, 445).



specifically. One can accept my claim and think that much of this work would have to be done by others, such as legislators or academics.<sup>16</sup>

Moving on, here is my argument for premise 2 of the Risk Argument:

7. Many of the conditions that are or might be necessary for punishment's permissibility are such that they are satisfied only if certain controversial philosophical claims are true.
8. Premise 2 of the Risk Argument is true if the following can be reasonably doubted: for every such condition, either the condition is not actually necessary or the controversial philosophical claims associated with it are true.
9. That can be reasonably doubted.
10. So, premise 2 of the Risk Argument is true; we can never show the following beyond reasonable doubt: all of the conditions that are necessary for punishment's permissibility are satisfied.

My discussion in the previous section shows that premise 7 is true. And premise 8 is obviously true. So, I will focus on defending premise 9. It is true because there are just too many conditions of the sort mentioned in premises 7 and 8 to preclude reasonable doubt about the claim that *all* of the conditions that are necessary for punishment's permissibility are satisfied. Consider just the ones from the last section. I discussed three at some length: the free will, culpability, and liability conditions. And I mentioned two more in passing: the good consequences and expression conditions. Given the complexity of the issues here and our fallibility, it seems reasonable to doubt the claim that each condition is either not genuine or that the controversial philosophical claims associated with it are true. Moreover, given the possibility that we may have overlooked some relevantly similar conditions, it seems even more reasonable to doubt the claim that all of the necessary conditions are satisfied. In short, rejecting 9 requires endorsing an overly optimistic view about our epistemic situation.

## 4 Objections

In this section, I will discuss a variety of objections. Along the way, I will revise my thesis in a couple of important ways.

---

<sup>16</sup> Compare Tomlin (2013), who argues that conduct should not be criminalized in the first place unless the conduct has been shown beyond reasonable doubt to be worthy of punishment. I have not put my arguments in terms of criminalization because, unlike Tomlin, I do not think that punishment is essential to the criminal law (Tomlin 2013, 45n1).

## 4.1 The Terrible Consequences Objection

Objection: The Risk Argument must have gone wrong somewhere because its conclusion is obviously absurd. Not punishing would have terrible consequences. At worst, it would result in social collapse and anarchy. At best, it would result in massive injustice because most if not all criminals would be able to get away with their crimes.

Reply: This objection assumes that punishment is necessary to maintain social order and hold criminals to account. But that is not obviously true. Though I do not have the space to go into the details here, abolitionists and their sympathizers have plausibly argued that there are non-punitive responses to crime that can do these things. To take just two examples, David Boonin (2008, 213–75) defends what he calls the theory of pure restitution, and Geoffrey Sayre-McCord (2001) defends what he calls legal reparations. Both argue that forcing criminals to compensate their victims is a genuine and viable alternative to punishment. Importantly, they also both argue that especially dangerous criminals can be non-punitively incapacitated or even confined to protect others (Sayre-McCord 2001, 508–9; Boonin 2008, 231–35).<sup>17</sup>

Advocates of the terrible consequences objection typically find it plausible because they mistakenly equate punishment with any coercive or harmful response to crime. It is plausible to think that refraining from such responses would have terrible consequences. But abolitionists are not committed to refraining from such responses.

## 4.2 The Overgeneralization Objection

Objection: The preceding reply fails. The arguments of this chapter generalize to every harmful act or practice, including the abolitionist alternatives to punishment just mentioned. This is because punishment's harmfulness is what makes it especially morally risky. Any harmful act or practice will be comparably morally risky.

Reply: The claim that harm is what makes punishment especially morally risky is false. To take just one set of counterexamples, civil courts often inflict harm without punishing. And this harm is sometimes comparable to the punitive harms inflicted by criminal courts. Yet the standard view seems to be

---

<sup>17</sup> Boonin's work on restitution significantly expands on the work of others. For references, see Boonin (2008, 216). Golash (2005, 22–48, 153–72) also discusses a variety of non-punitive responses to crime. For responses to the objection that abolitionist alternatives are actually punishments, see Sayre-McCord (2001, 506–7), Boonin (2008, 233–35), and Hanna (2022).

that the moral risk involved in harmful applications of the civil law are less serious—hence, the lower standards of evidence applied in civil courts. I will not defend this view or an account of what makes punishment’s moral risks especially serious. But it is plausible to think that these risks stem from a variety of features. Among them: punishment is harmful, it is intended to be harmful, and it stigmatizes the punishee in a distinctive and especially serious way.<sup>18</sup> The abolitionist alternatives mentioned above do not have all these features. So, the Risk Argument does not obviously generalize to them.

To be clear, I am not saying that harm can never make an act especially morally risky, only that this is not necessarily the case. Harm comes in degrees. It can be slight or severe. Acts that inflict severe harm may be especially morally risky because of the severity of the harm. I think that considerations of moral risk generate a presumption against harming and that the strength of this presumption gets stronger as the harm becomes more severe. I am inclined to think that some harms are so severe that the presumption against inflicting them is for all practical purposes insurmountable, regardless of whether the harms are punitive or non-punitive. In short, I think that considerations of moral risk exert strong downward pressure on the severity of harm that we can permissibly inflict (Tomlin 2014). To repeat, though: there are good reasons to think that harmful acts are not necessarily especially morally risky simply because they are harmful.

### 4.3 The Countervailing Moral Risks Objection

Objection: The arguments of this chapter entail a contradiction. They assume that morally risky acts are wrong if their permissibility can be reasonably doubted. But not punishing can also be morally risky and its permissibility can also be reasonably doubted. So, the arguments of this chapter entail that punishing and not punishing are sometimes both wrong.

Reply: My arguments do not assume that morally risky acts are wrong if their permissibility can be reasonably doubted, just that this is true of *punishment*. Punishment is special in ways that require the use of the reasonable doubt standard. This view is widely accepted, partly because of something captured in my paraphrase of the standard rationale for premise 4: wrongful

---

<sup>18</sup> On my view, punishment stigmatizes the punishee in such a way largely because it is intended to harm the punishee, and this sends a highly stigmatizing message about the punishee’s moral status. Abolitionists typically think that the intent to harm is morally significant. They argue that it makes punishment especially hard to morally justify and that non-punitive alternatives to punishment that do not intend harm are easier to morally justify, other things equal. See, e.g., Sayre-McCord (2001, 506–7), Boonin (2008, 15–16, 28–9, 234), Zimmerman (2011, 159–65), and Hanna (2021).

punishment is much worse than wrongful non-punishment, other things equal. That is to say, there is a significant asymmetry in terms of the seriousness of the moral risks here. So, my arguments do not entail a contradiction. At most, they entail that an act is wrong if it has a similar moral risk profile and if its permissibility turns on a comparably complex set of philosophical issues about which we are fallible. But that is a plausible result (Guerrero 2007, 92–94).

#### 4.4 The Counterexamples Objection

Objection: There are counterexamples to the arguments of this paper. That is, there are acts or practices that are obviously permissible, and the arguments of this paper generalize in a way that entails that these acts or practices are impermissible.

Reply: I do not have the space to consider every such alleged counterexample here. But I will outline my strategy for dealing with them. In response to an alleged counterexample like this, I would say one of two things: (1) My arguments do not generalize to the act or practice. This is because (a) the act or practice does not have a similar moral risk profile or (b) its permissibility does not turn on a comparably complex set of philosophical issues about which we are fallible. (2) My arguments do generalize in the alleged way and do entail that the act or practice is impermissible, but that is the correct result (Guerrero 2007, 92–94).

#### 4.5 The My-Favorite-Theory Objection

Objection: The arguments of this paper rely on highly controversial claims. These claims are highly controversial because they are inconsistent with certain moral theories. For example, premise 4 seems inconsistent with certain versions of consequentialism.<sup>19</sup>

Reply: I have appealed to what I take to be relatively uncontroversial and widespread commonsense moral intuitions. I grant that some moral theories are inconsistent with some of these intuitions. But that is a problem for the theories, not the intuitions. Or at least, I will take it to be a problem for the theories until I hear arguments for them that are more plausible than the

---

<sup>19</sup> Compare Kolber (2018, 520–22) and Laudan (2006, 2011, 2012). Kolber argues that moral risk is a serious problem for retributivists, but not as serious of a problem for consequentialists. And Laudan challenges, along partly consequentialist lines, many of the ways that the reasonable doubt standard is used.

intuitions to which I have appealed. I do not think that good enough arguments have been given in defense of any moral theory to make any of them a reliable basis for moral reasoning (Huemer 2010, 430–31).

#### 4.6 The Magical Doubts Objection

Objection: Premises 1 and 4 are false. They basically say that doubting the morality of an act can make it wrong. But there are good objections to that view (Weatherson 2014; Harman 2015).

Reply: This objection misunderstands the premises. It takes them to be saying that a certain subjective psychological state can make an act wrong, namely doubt about the act's permissibility. But that is not what the premises say. Basically, they say that punishment is wrong if its permissibility *can be reasonably doubted*. Punishment's permissibility can be reasonably doubted only if there are objective facts that make such doubt reasonable. My arguments should be taken to be saying that there are such objective facts and that they are what make punishment wrong—not any subjective psychological states that they might justify. Those objective facts include the fact that punishment's permissibility depends on a set of complex philosophical issues about which we are fallible and the fact that wrongful punishment is far worse than wrongful non-punishment.

#### 4.7 The Low Stakes Objection

Objection: Premise 4 says that punishment is permissible only when we have shown beyond reasonable doubt that the punishee broke the law. But this is false, at least when the punishment and the crime are minor. For example, a small fine seems like a permissible punishment for a minor traffic violation. And it seems permissible even if we have not shown beyond reasonable doubt that the punishee broke the law.

Reply: Assuming that these really are cases of punishment, I am not so sure that it is permissible to inflict them without having shown beyond reasonable doubt that the punishee broke the law.<sup>20</sup> But I will not defend that view here. Instead, I will revise my thesis in a way that does not require me to take a stand on this issue. Here is my revised thesis: punishment is wrong because it is too morally risky, at least when its permissibility requires showing beyond

---

<sup>20</sup> Some philosophers seem willing to say that these are not punishments, though, e.g., Feinberg (1965, 398).

reasonable doubt that the punishee broke the law. This is still an important thesis that, if true, requires radical criminal justice reform. Punishment's advocates should find little comfort in the fact—if it is a fact—that my arguments do not apply to the sorts of cases to which the objection appeals.

#### 4.8 The Epistemic Possibility Objection

Objection: The argument for premises 7 and 9 rely on certain claims about what is possible, for example, that we have no free will and that no one is ever culpable. But there is no obviously good sense of possibility on which it is uncontroversial to say that these claims are possibly true. The most natural way to understand the possibility claims is in terms of epistemic possibility, which is typically defined as follows: a proposition *p* is epistemically possible for a subject *S* if *p* could be true for all that *S* knows. But it would be highly controversial to say that the above claims are possible in this sense because that would entail, among other things, that we do not know that we have free will or that people are sometimes culpable.

Reply: I do not want to say that the above claims are possible in this sense. For one thing, the objection is right that this would be highly controversial. For another thing, I happen to believe that some people, including myself, know that people have free will and that people are sometimes culpable. But none of this means that there is no good sense of possibility on which the above claims are possibly true. Here is one such sense of possibility: the claims are true *for all that we know for certain* (Chalmers 2011, 60).

#### 4.9 The No Necessary Conditions Objection

Objection: The arguments of this chapter assume that there are necessary conditions for punishment's permissibility. But there are not—there are only conditions that are usually necessary for its permissibility. For any allegedly necessary condition, we can imagine a case where punishment is permissible even though the condition is not satisfied. Showing this is trivial for the conditions discussed in this chapter. If, for example, not punishing someone who satisfies none of these conditions will have catastrophically bad consequences, it would be permissible to punish the person (Wellman 2012, 375n7).

Reply: At best, all that this shows is that the conditions that I have discussed must be qualified in certain ways. For purposes of illustration, let us focus on just one of them: the culpability condition. To accommodate the above considerations, the condition can be restated as follows: *P* must be culpable or there must be an outweighing factor present that makes punishing



the non-culpable permissible. Other conditions can be modified in similar ways. And everything that I have said about the conditions as I originally stated them applies to their modified versions.

#### 4.10 The Moral Fetishism Objection

Objection: The rationale for premise 4 is implausible because the claim that it can be wrong to risk doing wrong is false. It treats the fact that an act might be wrong as a reason not to perform the act. But the fact that an act might be wrong is not a reason not to perform the act. This is because wrongness does not in itself matter morally. What matters morally are the facts that can make acts wrong. Concerns about moral risk fetishize wrongness (Weatherson 2014).

Reply: Nothing that I have said commits me to the claim that wrongness itself matters morally. What matters is the risk of doing things like punishing people who did not act freely, who did not act culpably, and who are not liable to punishment. More broadly, what matters are things like not treating people in deeply disrespectful ways and not intentionally harming them without sufficient reason (Sepielli 2016, 2959–60; MacAskill et al. 2020, 27). Anything that I have said about the risk of doing wrong is just a convenient way of talking about the risk of doing these sorts of things.

#### 4.11 The Incompatibility Objection

Objection: The Risk Argument must have gone wrong somewhere because it is inconsistent with certain reasonable beliefs.<sup>21</sup> To see this, consider the following summation of the argument:

Punishment is wrong because its permissibility can be reasonably doubted.

And consider the following belief:

Punishment might not be wrong.

Anyone who accepts the former claim cannot reasonably believe the latter. But believing the latter is obviously reasonable. Considerations of moral risk

---

<sup>21</sup> The following objection is adapted from Weatherson (2014, 146), who deploys a similar objection against the following principle: if an agent has a choice between two options, and one might be wrong, while the other is definitely permissible, then it is wrong to choose the first option. For present purposes, I am agnostic about this principle. But I suspect that my reply to the above objection can be adapted to defend the principle from Weatherson's objection.

might give us some reason to doubt punishment's permissibility, but they cannot show that there is simply no possibility that it is permissible.

Reply: There are two different ways that punishment might be wrong: for reasons related to moral risk or for reasons unrelated to moral risk. For argument's sake, I grant that the belief that punishment might not be wrong in the second way is reasonable. But that belief is consistent with my arguments. The belief that punishment might not be wrong in the first way is inconsistent with my arguments. But it is also not obviously reasonable to believe, at least for people who understand my arguments. Insisting otherwise just begs the question.

#### 4.12 The Subjectivism Objection

Objection: Premises 1 and 4 are false or at least highly controversial. They entail that wrongness is subjective in the sense that it depends on our evidence. And there are good objections to that view (Zimmerman 2008; Graham 2010).

Reply: First, we must distinguish the claim that wrongness depends only on our evidence from the claim that it can be affected by our evidence. Critics of subjectivism typically attack the former claim, but I am not committed to it. Maybe I am committed to a version of the latter claim, but I am not obviously committed to an unacceptable version of it. It is plausible to think that excessive moral risk can sometimes make acts wrong. Here is an example: if it is reasonable to worry that an act will kill an innocent person, it is plausible to think that this can make the act wrong even if it will not in fact kill an innocent person. Or at least, it is plausible to think that this risk can make the act wrong if one has not sufficiently investigated whether the act will do this.

Second, to the extent that subjectivism seems implausible, it seems implausible as a view about individual action. It seems more plausible as a view about collective political action, such as laws and legal practices. If, as many philosophers think, laws and legal practices are morally acceptable only if they are in some sense justifiable to everyone affected, then it is plausible to think that moral risk can make certain laws and legal practices—e.g., punishment—wrong.<sup>22</sup>

Third, even if my previous replies fail, my thesis can be revised to sidestep the objection. I discuss this revision in response to the next objection.

---

<sup>22</sup>The idea that laws and legal practices must in some sense be justifiable to everyone affected is a core tenet of public reason liberalism. For discussion, see Quong (2018).

### 4.13 The Irrelevance Objection

Objection: Premises 2 and 9 are partly motivated by the possibility that we lack free will. But that possibility does not support the Risk Argument's conclusion. If we lack free will, punishment would not be wrong because all of our acts would be unfree and unfree acts are not wrong. Appealing to the possibility that we lack free will therefore exaggerates punishment's moral risks.

Reply: I am not so sure that unfree acts cannot be wrong (Kelly 2018, 83–84). But I will set that issue aside. Instead, I will revise my thesis again to sidestep the issue. Here is my modified thesis: punishment is morally *unjustified* because it is too morally risky, at least when its justifiability requires showing beyond reasonable doubt that the punishee broke the law. To say that an act is morally unjustified is to say that it is morally wrong or overall morally bad or morally vicious.<sup>23</sup> Punishing people for their unfree acts is presumptively very morally bad even if the act of punishment is not wrong because it is not free. And to take an excessive risk of doing this is overall morally bad as well as morally vicious (e.g., morally reckless), even if the act of punishment is not wrong because it is not free.

## 5 Conclusion

Here is the definitive statement of my thesis: legal punishment is morally unjustified because it is too morally risky, at least when its moral justifiability requires showing beyond reasonable doubt that the punishee broke the law. Unlike some prominent defenses of abolitionism, I have not individually criticized every justification of punishment in the literature. Instead, I have sidestepped the lengthy debates about these justifications and have tried to argue that no such justification can possibly work because there is no way to show beyond reasonable doubt that punishment is justified. And unlike many other theorists who discuss punishment's moral risks, I have not emphasized any one source of moral risk. Instead, I have argued that punishment is morally risky in a variety of ways that combine to make a good case for abolitionism.<sup>24</sup>

---

<sup>23</sup> See Hanna (2021) for further discussion of this sense of justification and its moral significance.

<sup>24</sup> Thanks to Marcus Hedahl and Adam Kolber for extensive comments on earlier versions of this chapter. Thanks also to audiences at CU Boulder, the 2021 Rocky Mountain Ethics Congress, and the 2019 meetings of the Alabama Philosophical Society, the North American Society for Social Philosophy, and the North Carolina Philosophical Society.

## References

- Berman, Mitchell. 2008. "Punishment and Justification." *Ethics* 118, no. 2 (January): 258–90.
- Boonin, David. 2008. *The Problem of Punishment*. Cambridge: Cambridge University Press.
- Caruso, Gregg D. 2020. "Justice without Retribution: An Epistemic Argument against Retributive Criminal Punishment." *Neuroethics* 13, no. 1 (April): 13–28.
- . 2021. "Skepticism about Moral Responsibility." *Stanford Encyclopedia of Philosophy* (Summer 2021 edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/sum2021/entries/skepticism-moral-responsibility/>.
- Caruso, Gregg D., and Derk Pereboom. 2020. "A Non-Punitive Alternative to Retributive Punishment." In *The Routledge Handbook of the Philosophy and Science of Punishment*, edited by Farah Focquaert, Elizabeth Shaw, and Bruce N. Waller, 355–65. New York: Routledge.
- Chalmers, David. 2011. "The Nature of Epistemic Space." In *Epistemic Modality*, edited by Andy Egan and Brian Weatherson, 60–107. New York: Oxford University Press.
- Feinberg, Joel. 1965. "The Expressive Function of Punishment." *Monist* 49, no. 3 (July): 397–423.
- Gardiner, Georgi. 2017. "In Defence of Reasonable Doubt." *Journal of Applied Philosophy* 34, no. 2 (February): 221–41.
- Golash, Deirdre. 2005. *The Case against Punishment*. New York: New York University Press.
- Graham, Peter A. 2010. "In Defense of Objectivism about Moral Obligation." *Ethics* 121, no. 1 (October): 88–115.
- Gross, Hyman. 2012. *Crime and Punishment: A Concise Moral Critique*. New York: Oxford University Press.
- Guerrero, Alexander A. 2007. "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution." *Philosophical Studies* 136, no. 1 (October): 59–97.
- Hanna, Nathan. 2008. "Say What? A Critique of Expressive Retributivism." *Law and Philosophy* 27, no. 2 (March): 123–50.
- . 2012. "It's Only Natural: Legal Punishment and the Natural Right to Punish." *Social Theory and Practice* 38, no. 4 (October): 598–616.
- . 2013. "Two Claims about Desert." *Pacific Philosophical Quarterly* 94, no. 1 (March): 41–56.
- . 2014. "Facing the Consequences." *Criminal Law and Philosophy* 8, no. 3 (October): 589–604.
- . 2019. "Hitting Retributivism Where It Hurts." *Criminal Law and Philosophy* 13, no. 1 (March): 109–27.
- . 2021. "Why Punitive Intent Matters." *Analysis* 81, no. 3 (July): 426–35.
- . 2022. "Punitive Intent." *Philosophical Studies* 179, no. 2 (February): 655–69.

- Harman, Elizabeth. 2015. "The Irrelevance of Moral Uncertainty." In *Oxford Studies in Metaethics*, vol. 10, edited by Rudd Shafer-Landau, 53–79. New York: Oxford University Press.
- Huemer, Michael. 2010. "Is There a Right to Immigrate?" *Social Theory and Practice* 36, no. 3 (July): 429–61.
- . 2018. "The Duty to Disregard the Law." *Criminal Law and Philosophy* 12, no. 1 (March): 1–18.
- Jeppsson, Sofia M. I. 2021. "Retributivism, Justification, and Credence: The Epistemic Argument Revisited." *Neuroethics* 14, no. 2 (July): 177–90.
- Kelly, Erin I. 2018. *The Limits of Blame: Rethinking Punishment and Responsibility*. Cambridge, MA: Harvard University Press.
- Kolber, Adam J. 2018. "Punishment and Moral Risk." *University of Illinois Law Review* 2018, no. 2 (March): 487–532.
- Laudan, Larry. 2006. *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*. New York: Cambridge University Press.
- . 2011. "The Rules of Trial, Political Morality, and the Costs of Error: Or, Is Proof beyond a Reasonable Doubt Doing More Harm Than Good?" In *Oxford Studies in Philosophy of Law*, vol. 1, edited by Leslie Green and Brian Leiter, 195–227. New York: Oxford University Press.
- . 2012. "Is It Finally Time to Put 'Proof beyond a Reasonable Doubt' Out to Pasture?" In *The Routledge Companion to Philosophy of Law*, edited by Andrei Marmor, 317–32. New York: Routledge.
- Levy, Neil. 2011. *Hard Luck: How Luck Undermine Free Will and Moral Responsibility*. New York: Oxford University Press.
- MacAskill, William, Krister Bykvist, and Toby Ord. 2020. *Moral Uncertainty*. New York: Oxford University Press.
- Nagel, Thomas. 1979. "Moral Luck." In *Mortal Questions*, 24–38. New York: Cambridge University Press.
- Nelkin, Dana Kay. 2019. "Guilt, Grief, and the Good." *Social Philosophy and Policy* 36, no. 1 (Summer): 173–91.
- O'Connor, Timothy, and Christopher Franklin. 2021. "Free Will." *Stanford Encyclopedia of Philosophy* (Spring 2021 edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/spr2021/entries/freewill/>.
- Pereboom, Derk. 2001. *Living without Free Will*. Cambridge: Cambridge University Press.
- Quong, Jonathan. 2018. "Public Reason." *Stanford Encyclopedia of Philosophy* (Spring 2018 edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/spr2018/entries/public-reason/>.
- Roebuck, Greg, and David Wood. 2011. "A Retributive Argument against Punishment." *Criminal Law and Philosophy* 5, no. 1 (January): 73–86.
- Rosen, Gideon. 2003. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103, no. 1 (June): 61–84.

- . 2004. "Skepticism about Moral Responsibility." *Philosophical Perspectives* 18, no. 1 (December): 295–313.
- Sayre-McCord, Geoffrey. 2001. "Criminal Justice and Legal Reparations as an Alternative to Punishment." *Philosophical Issues* 11, no. 1 (October): 502–29.
- Sepielli, Andrew. 2016. "Moral Uncertainty and Fetishistic Motivation." *Philosophical Studies* 173, no. 11 (November): 2951–68.
- Strawson, Galen. 1994. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75, nos. 1/2 (August): 5–24.
- Tadros, Victor. 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*. New York: Oxford University Press.
- Tomlin, Patrick. 2013. "Extending the Golden Thread? Criminalisation and the Presumption of Innocence." *Journal of Political Philosophy* 21, no. 1 (March): 44–66.
- . 2014. "Could the Presumption of Innocence Protect the Guilty?" *Criminal Law and Philosophy* 8, no. 2 (June): 431–47.
- Vilhauer, Benjamin. 2009. "Free Will and Reasonable Doubt." *American Philosophical Quarterly* 46, no. 2 (April): 131–40.
- Weatherson, Brian. 2014. "Running Risks Morally." *Philosophical Studies* 167, no. 1 (January): 141–63.
- Wellman, Christopher Heath. 2012. "The Rights Forfeiture Theory of Punishment." *Ethics* 122, no. 2 (January): 371–93.
- Zimmerman, Michael J. 1997. "Moral Responsibility and Ignorance." *Ethics* 107, no. 3 (April): 410–26.
- . 2002. "Controlling Ignorance: A Bitter Truth," *Journal of Social Philosophy* 33, no. 3 (Fall): 483–90.
- . 2008. *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.
- . 2011. *The Immorality of Punishment*. Peterborough, ON: Broadview.