

# The Problem of Punishment

## 1.0 OVERVIEW

Legal punishment involves treating those who break the law in ways that it would be wrong to treat those who do not. Even if we assume that those who break the law are responsible for their actions and that the laws they break are just and reasonable, this practice raises a moral problem. How can the fact that a person has broken a just and reasonable law render it morally permissible for the state to treat him in ways that would otherwise be impermissible? How can the line between those who break such laws and those who do not be morally relevant in the way that the practice of punishment requires it to be? This is the problem of punishment.

The problem of punishment has generated a large and increasingly sophisticated literature with a wide variety of attempted solutions. This book contributes to that literature in two ways. First, it offers a comprehensive, up-to-date introduction to the contemporary literature by providing a detailed account of the nature of punishment and of the problem it poses, followed by a survey of the many solutions to the problem in the current literature. Second, it provides a critical evaluation of these solutions both as a means of introducing the reader to the various debates that these solutions have generated and as a way of defending a particular thesis about the problem that stands in stark contrast to the position taken in the vast majority of the literature on the subject. This is the thesis that there is no solution to the problem of punishment and that it is morally impermissible for the state to punish people for breaking the law.

The claim that it is morally impermissible for the state to punish people for breaking the law is likely to strike most people as implausible, if not absurd. While debates persist about precisely which forms of behavior a government may justly and reasonably prohibit, there is widespread agreement that if it is appropriate for a state to prohibit a particular form of behavior, then it is permissible for the state to punish

those who engage in it. The thesis defended by this book in response to the literature on the problem of punishment will likely be met with a good deal of resistance.

Much of this resistance is likely to be based on the belief, or at least the assumption, that there is a satisfactory defense of the moral permissibility of punishing people for breaking the law. Philosophers and legal theorists have typically sought to justify this practice either by appealing to the consequentialist claim that the presumed benefits of punishment are sufficient to render it morally permissible or by relying on the retributivist claim that punishment is justified because it is a fitting response to wrongdoing, regardless of its consequences. Others have attempted to justify legal punishment on a variety of additional grounds, such as the claim that such punishment is a form of moral education, social expression, or collective self-defense. Still others have appealed to some combination of these views. If resistance to this book's thesis rests on the belief that one or more of these attempts to establish the moral permissibility of legal punishment is successful, the only way to try to overcome this resistance is to try to demonstrate that all of these attempts are unsuccessful. That is the task of the central chapters of this book. In Chapter 2, I explain and argue against a variety of consequentialist attempts to justify the claim that legal punishment is morally permissible. In Chapter 3, I explain and argue against a variety of attempts to justify this claim along retributivist lines. And in Chapter 4, I explain and argue against a variety of further attempts to support the moral permissibility of punishment that do not fall readily under either of these two headings or that attempt to fall simultaneously under both.

A second likely source of resistance to this book's thesis is the belief that there is no way for us to do without punishment. Punishment, on this understanding, is necessary, either as a condition for the existence of a social order at all or as a condition for the kind of social order that makes possible just relationships among its members. On either version of this appeal to necessity, the practice of punishing people for breaking the law is said to be necessary, and if a practice is necessary, then an argument against its permissibility may seem pointless at best, incoherent at worst. I respond to this second source of resistance to this book's thesis in the final chapter by presenting and defending a single counterexample to the claim made by the appeal to necessity. This is the proposal, most widely associated with Randy Barnett's provocative article "Restitution: A New Paradigm of Criminal Justice" (1977), that we do without punishment by embracing a system of compulsory victim restitution. Following Barnett, I refer to this proposal as the "theory of pure restitution," and I argue that the theory is good enough to warrant rejecting the appeal to necessity. I do not insist that compulsory victim restitution is the only acceptable alternative to punishment or even that it is the best alternative. But I do

argue that it is an acceptable alternative. If it is an acceptable alternative, then punishment is not necessary. And if punishment is not necessary, then the appeal to necessity fails to undermine this book's central claim that it is morally impermissible for the state to punish people for breaking the law.

Finally, it seems likely that at least part of the resistance to this book's thesis lies in the failure to recognize that punishing people for breaking the law requires moral justification in the first place. The practice of punishment, after all, is ubiquitous. Ubiquitous practices are rarely called into question. Addressing this important concern is the goal of this introductory chapter. In Section 1.1, I explain why a critical assessment of punishment must begin with a definition of legal punishment (1.1.1), briefly present some criteria for adjudicating between rival definitions (1.1.2), and then present and defend a definition of legal punishment that does best by these criteria (1.1.3–1.1.7). I conclude this section by showing that, by this definition, punishment is importantly different from compulsory victim restitution (1.1.8). In Section 1.2, I explain why punishment, so understood, requires moral justification and poses a genuine moral problem (1.2.1), respond to two arguments against this claim (1.2.2–1.2.3), and conclude by identifying and explaining two tests that any solution to the problem of punishment must pass to be considered successful (1.2.4). This analysis, in turn, sets the stage for the remainder of the book, in which I argue that no solution to the problem of punishment passes both of these tests and that we should abolish our practice of punishing people for breaking the law.

## 1.1 WHAT PUNISHMENT IS

### 1.1.1 The Need for a Definition

When we talk about the moral permissibility of legal punishment, what, precisely, do we mean? A general answer to this question is easy: we mean such practices as the state's imposition of monetary fines, forced incarceration, bodily suffering, and – in extreme cases – death. A more specific answer is more difficult. Simply illustrating punishment,<sup>1</sup> even by appealing to clear paradigmatic examples, is not the same as defining it.

But is a more specific answer necessary for our purposes? It is tempting to suppose that it is not. As long as we all know what counts as examples of punishment, it might be said, we can move directly to the task of arguing about whether or not it is morally defensible. Indeed, one book on punishment begins by declining to offer a definition of the term for

<sup>1</sup> Unless otherwise noted, when I say “punishment” in this book, I mean “legal punishment.”

precisely this reason: “one does not require a definition of ‘punishment’ in order to recognize clear cases of punishment’s being imposed and to distinguish such cases from those in which individuals are treated in ways that, although similar to punishment in certain respects, are nevertheless something else entirely” [Montague (1995: 1)]. An “understanding” of punishment is certainly needed, Montague concedes, but one can understand punishment well enough without defining it.

While the reluctance to begin a discussion of punishment by developing a clear, specific definition is understandable, however, it is ultimately misguided. For a fully satisfactory inquiry into the moral permissibility of punishment, it is not enough to point to examples and say either that they are cases of punishment or that they are cases of something else. One must also be able to identify the properties that make them something else. If one cannot do this, then one cannot fully determine what, precisely, makes the permissibility of punishment problematic. More importantly, if one cannot do this, then one cannot satisfactorily determine whether or not a purported justification of punishment succeeds in justifying punishment or only in justifying something very much like it. Indeed, as we will see in Chapter 4, Montague’s own attempt to defend punishment on grounds of social self-defense fails in part for precisely this reason.<sup>2</sup> Even if the argument from self-defense succeeds, I will argue, the practice that the argument would justify lacks two of the necessary characteristics that any satisfactory definition of punishment must include. Montague’s failure to define punishment at the beginning of his book results in his failure to see that what he is defending at the end of his book is not exactly punishment.

Finally, and perhaps most importantly for the purposes of this book, we cannot fully disentangle the importantly related practices of punishment and compulsory victim restitution without understanding what makes some cases cases of punishment and others cases of something else. Such disentanglement is crucial to the project of this book: it is necessary to see precisely why rejecting the claim that punishment is morally permissible does not entail rejecting the claim that compulsory victim restitution is morally permissible. For all of these reasons, then, we must begin our investigation by clarifying what makes some forms of treatment cases of punishment and others cases of something else. And it is difficult to see how to do this without a definition.

### 1.1.2 The Criteria for a Definition

So, we want a definition of punishment. But we do not want just any definition. We want a good one. What would constitute a good definition of punishment? First, it must be *accurate*. It must provide us with a set of

<sup>2</sup> See Section 4.4.5.

necessary and sufficient conditions that clearly demarcates cases of punishment from cases of something else. The results produced by this demarcation must cohere sufficiently well with what we mean by punishment when we argue about it and must do so over a sufficiently wide range of cases. If it is clear that responding to an offender's behavior by fining him, beating him, or executing him do count as punishments, for example, and that responding to his offense by writing him a check, throwing him a parade, or giving him a free meal do not, then an adequate definition of punishment must account for these judgments.<sup>3</sup> If it is unclear or indeterminate whether or not responses such as voter disenfranchisement, supervised probation, public shaming, or certain forms of taxation should count as forms of punishment, then a good definition should help us to make sense of these facts as well.

Second, a good definition of punishment must be *illuminating*. A definition may be accurate, successfully discriminating between cases of punishment and cases of something else, but if it does so only because it contains various stipulations that are thrown in solely to produce the desired results and have no further independent motivation, then the definition will be unacceptably *ad hoc*. When we appeal to it in asking whether or not a particular act counts as an act of punishment, such a definition will give us the correct answer, but it will do nothing to demonstrate why the answer is correct. In part, we want a good definition to get at the essence of the thing defined, to tell us not just that a given subject belongs in a certain class with certain other subjects, but in virtue of what fact or set of facts this is so.

Finally, a good definition of punishment must be *neutral* on the question of whether or not punishment is morally permissible. A definition is unacceptable if it begs the question one way or the other, with respect to either the merits of punishment in general or the merits of any kind of justification of punishment in particular. If, for example, one attempted to discriminate between punishment and mere private vengeance by saying that punishment is "authorized" while private vengeance is not, and if part of what one meant by an act's being authorized was that it was legitimate, then the resulting definition of punishment would unacceptably beg the question in favor of the claim that punishment is morally permissible. If one defined punishment so that part of what made an act a punishment is that it was justified because of its effects

<sup>3</sup> The claim that a good definition must accurately capture actual usage of the term 'punishment' does not mean that in order to use such a definition, a defender of the permissibility of punishment must defend the permissibility of all forms of punishment. It means that a defender of punishment must acknowledge that capital punishment and corporal punishment are forms of punishment, for example, but it does not mean that he or she must insist that they are morally permissible.

on society, or that it was not justified in this way, then the result would fail to be neutral with respect to the various competing solutions to the problem of punishment. In short, we want a definition of legal punishment that respects and reflects both our beliefs about what counts as punishment and our puzzlement over what, if anything, renders it morally permissible for the state to punish people.

### 1.1.3 Harm

A definition that satisfies these requirements can be obtained by testing various conditions against our intuitive reactions to clear, paradigmatic instances of legal punishment. As already noted, such cases include monetary fines, forced incarceration, bodily suffering, and, in extreme cases, death. So, we should begin by asking what these various practices have in common.

Perhaps the most obvious quality that these practices have in common is that they are all in some way bad for the person on whom they are inflicted.<sup>4</sup> This point is often expressed by saying that punishment necessarily involves “pain,” but this way of putting things is unsatisfactory.<sup>5</sup> A murderer, for example, could be executed painlessly, and this would clearly be bad for him even if he does not experience pain. The same problem arises if punishment is defined, as it sometimes is, in terms of subjecting people to experiences that are “unpleasant.” Other writers have attempted to capture the sense in which punishment involves something negative for the person on the receiving end by saying that punishment involves an “evil,” but this runs the risk of defining punishment as something that is, at least in itself, a wrong; and this, in turn, would violate the requirement of neutrality by begging the question against those retributivists who maintain that the treatment that punishment inflicts on an offender is not merely allowable but a positive good. Finally, some writers have defined the negative effect of punishment on the person who is punished in terms of the language of rights. Punishment, on this account, involves depriving someone of what would otherwise be a right. If one holds the view that losing a right is always bad for someone, then putting things in terms of rights poses no real difficulties for an analysis of punishment as something that is bad for someone. But if, as seems plausible to me, there can be cases in which a person loses a

<sup>4</sup> That punishment involves treatment that is, by some measure, of negative value for its recipient is accepted by virtually every philosopher who has written on the subject, including such historical figures as Plato, Aristotle, Aquinas, Hobbes, Locke, Kant, and Hegel, as well as more recent writers such as Flew, Benn, Hart, McCloskey, Honderich, and Primoratz [for citations, see Adler (1991: 285–6)].

<sup>5</sup> See, e.g., Newman, who insists that “Punishment must, above all else, be painful” (1983: 6), and Corlett (2001: 68).

right but is not made worse off by this loss, then such cases would seem to provide a good reason not to link punishment to rights by definition. A woman who is physically incapable of becoming pregnant, for example, might still have a legal right to an abortion, and if depriving her of that right would in no way be bad for her, it is difficult to see how it could count as punishing her. It therefore seems more sensible to say that acts of punishment all, in some way, make the person who is punished worse off than she would otherwise be. If an offender received a monetary prize for her offense, or a paid vacation, a relaxing massage or life-extending therapy, for example, we would not be inclined to say that she had been punished for her transgression. And so, a natural starting point in generating a definition of punishment is to say that punishment *harms* the person who is punished, where harming someone means making her worse off in some way, which includes inflicting something bad on her or depriving her of something good. I will refer to this as the “harm requirement.”

### 1.1.3.1 *The Beneficial Consequences Objection*

A critic of the harm requirement might object that this requirement neglects the beneficial long-term consequences that punishment can have for the person who is punished. Adler, for example, who rejects the claim that harmfulness is an essential property of punishment, appeals to what he calls the “conscientious punishee,” the offender “who wants to submit to punishment, who believes that she can achieve reconciliation, atonement, expiation, renewed innocence, greater moral knowledge, or some other good by undergoing the punishment” (1991: 91). Indeed, as we will see in Section 4.3, a number of writers have claimed not only that punishment ultimately benefits the offender who is punished, but that the moral permissibility of punishment is grounded in this very fact. A definition of punishment that incorporates the harm requirement would therefore seem to beg the question against such a position, ruling out the possibility that punishment might be justified as ultimately good for the person punished by definitional fiat. This, in turn, would violate the neutrality requirement established earlier, rendering the definition unacceptable.

This objection to the harm requirement is understandable, but it is also mistaken. The harm requirement maintains that for a certain treatment to count as a punishment, it must harm the recipient. But it is neutral on the further question of whether or not being subject to such a harm might produce beneficial consequences in the future, including beneficial consequences that are great enough to outweigh (and perhaps even to justify) the immediate harmful ones. Consider, for example, a child who is spanked as a (nonlegal) punishment for having hit another child. The parent who punishes a child in this way may believe that spanking will



make him understand more fully why what he did was wrong, and that this, in turn, will contribute to the child's moral development in various important ways. If this is so, then spanking the child now will ultimately benefit him in the future. But all of this is perfectly consistent with the harm requirement. Indeed, it presupposes it. For if spanking the child does benefit him in this way, then this will be so precisely because it involves inflicting a harmful treatment on the child as a means of demonstrating to him how it feels to be on the receiving end of such harmful treatment. If the spanking were not harmful to the child (if, for example, it felt just like being pleasingly caressed), then it would not have the desired educative effect of showing what it is like to be a victim of wrongful treatment in the first place. So, considerations of the possible long-term benefits of punishment provide no reason to reject the harm requirement. If anything, they provide further reason to accept it.<sup>6</sup>

#### 1.1.3.2 *The Masochist Objection*

A second objection to the harm requirement is that it is subject to refutation by counterexample. Most people, for example, strongly dislike being physically beaten. But some people, apparently, do not. Most people would find incarceration highly unpleasant. But some people, perhaps, would not, and others, depending on their circumstances, might find it preferable to the available alternatives. And so, it might be urged, we can say at most that punishment involves treatments that are *typically* harmful or that are considered undesirable by *most* people, but we cannot say that this is so of punishment in every instance. You and I might strongly prefer not to be whipped, for example, and so this punishment would be harmful to us, but a masochist might enjoy a beating; and, if he did, it would remain a form of corporal punishment nonetheless. Since such cases apparently involve acts that are acts of punishment but that do not harm their recipients, they seem to demonstrate that the harm requirement is not accurate over an important (even if somewhat limited) range of cases.<sup>7</sup>

The objection that appeals to cases such as the masochist rests on two claims: that in such cases the treatment in question does not harm the recipient and that it counts as punishment nonetheless. A defender of the

<sup>6</sup> It is possible, of course, that a proponent of the objection might insist that the benefits of submitting to punishment are immediate rather than delayed. But if the recipient of a given treatment is benefited at the moment that the treatment begins, it is not clear what reason we would have for considering it to be a punishment in the first place. If a pleasant caress on the child's back benefits him immediately and also somehow teaches him that it is wrong to hit other children, for example, then it may well serve the same purpose as a spanking for educative purposes, but it would clearly fail to count as punishment and so would again fail to provide a counterexample to the harm requirement.

<sup>7</sup> This problem is raised by, e.g., Kasachkoff (1973: 364–5) and Snook (1983: 131).



harm requirement might reject the objection's first claim and argue that even if the masochist enjoys being beaten, a beating is still something that is objectively harmful to him. Similarly, even if a homeless or insecure person prefers the security of prison to the unpredictability of life on the outside, one could argue that the restriction on his freedom of movement is objectively a grave harm to him even if he doesn't particularly mind it.

But even if the objection's first claim can be sustained in a significant range of cases, the second should be rejected outright. For if we concede that the masochist is not harmed by being whipped or that the homeless person is not harmed by being imprisoned, then we have two good independent reasons to conclude that he is not punished either. And if he is not punished, of course, then even if he is not being harmed, he cannot serve as a counterexample to the claim that punishment requires harm.

The first reason to believe that these attempted counterexamples fail in this way arises because there is a conceptual symmetry between punishment and reward. What is true of punishment in one direction, that is, must be true of reward in the other. Yet, in the case of reward, it should be clear that a person has not been rewarded for doing a good deed if the treatment that she receives in response does not in fact end up benefiting her. Suppose, for example, that I give you a piece of candy because you did me a favor last week, but the candy causes a severe allergic reaction. We might say that I tried to reward you for your good deed or that I intended to reward you, but we would not say that you had, in fact, been rewarded. And we would not say this precisely because you had not been benefited. Since it seems reasonable to presume that reward and punishment are symmetrical in this respect, this provides support for the claim that the offender who is not actually harmed by the treatment he or she receives is not actually punished by it.

The second reason to believe that without real harm there is no real punishment arises from cases in which we believe that no harm is done because of some particular fact about the treatment itself. When a stay in a minimum-security prison for white-collar criminals seems to resemble nothing more than an all-expenses-paid vacation at a comfortable resort, for example, people do not consider the offender to have been punished and they complain about his being treated so leniently for precisely this reason.<sup>8</sup> Our intuitive response to punishments that seem clearly non-harmful and to attempts to reward that clearly do not benefit both vindicate the claim that the harm requirement is a core component of our concept of punishment. And so, the apparent counterexamples to the

<sup>8</sup> When it was reported that the son of former vice presidential nominee Geraldine Ferraro was serving his sentence for a drug conviction in a \$1,500-a-month luxury apartment, for example, the public outcry over the case prompted the governor of Vermont to discontinue the house arrest option for drug offenders [Tunick (1992: 3)].

harm requirement, in which it seems that a person is punished but is not harmed, in the end do not undermine the harm requirement but once again reinforce it.

I think that these considerations suffice to defend the harm requirement from what might be called the “masochist objection,” but there is one more concern that might be raised at this point. For if we agree that the masochist who is not harmed by his whipping is not punished by it either, it can seem that we must therefore conclude that whipping is not a form of punishment after all. And that result can seem sufficiently counterintuitive to force us back to the conclusion that the masochist really is being punished and that punishment therefore really does not require harm. This worry about my rejection of the masochist objection is understandable, but it is ultimately misguided. The reason is that there is a crucial difference between saying that a particular person has been subjected to a form of treatment that is a form of punishment and saying that this person has, in fact, been punished. And even if it is possible that some people are not harmed by being subjected to forms of treatments that are uncontroversially characterized as forms of punishments, this does not mean that we must say that such people are actually punished by such treatments.

Since this response to the objection may at first seem puzzling, an analogy may be of use. Consider a doctor who administers a sedative to a patient. An essential property of a sedative is that it makes people sleepy. But just as there are some people who may be delighted by some forms of punishment, there may be some people who are stimulated by some forms of sedatives. If the doctor gives such a drug to such a patient, then what she gives the patient might still be properly characterized as a sedative because of its general properties, but this does not mean that in giving the sedative to this particular patient she actually sedates the patient. Similarly, if the state inflicts a form of corporal punishment on someone who is not harmed by it, then while it may be proper to continue to refer to this treatment as a form of punishment (since it is a form of treatment that does, in general, harm people), this does not mean that in administering it to this particular offender the state will in fact be punishing him. It will, at most, be attempting to punish him.<sup>9</sup>

<sup>9</sup> And in at least some cases, it will not even be clear that it should be considered an attempt at punishment, let alone a successful attempt. After his lawyer reached a plea bargain agreement with Oklahoma City prosecutors for a thirty-year prison sentence for two charges of shooting with intent to kill and one weapons violation, for example, Eric James Torpy insisted that he would rather get thirty three years to match the uniform number of his basketball hero, Larry Bird. The judge in the case was quoted as saying that “We accommodated his request and he was just as happy as he could be” (“Man Asks for More Jail Time to Honor Bird” 2005). Although three extra years in prison would generally be considered a form of (additional) punishment, it is difficult to believe that the judge