
REHABILITATION VS. RETRIBUTION, RESTITUTION, AND DETERRENCE.

By Noah McKay

Resolved: Criminal justice ought to prioritize rehabilitation over retribution, restitution, or deterrence.

Questions about criminal justice have occupied philosophers for millennia. The philosophical problem most relevant to this year's topic is the *problem of punishment*, which is most clearly put as a question: When and how is criminal punishment justified? That is a hard question to answer – after all, punishment involves intentionally inflicting harm on a person, something which we normally consider to be wrong. So what, precisely, distinguishes punishment from other kinds of intentional harm? And, given that it is possible for punishments themselves to be immoral, what is the difference between punishments that are justified and those that are not? Countless different answers have been given to these questions, and the answers one accepts will affect one's position on the resolution.

Part I: What is Punishment?

Before examining various solutions to the problem of punishment, we ought to be clear about precisely what punishment is. Philosophers disagree on the details, but there is a consensus that punishment is hard treatment given to a person who has done something wrong *because* they have done something wrong. Punishment must be intentional: if you commit a heinous crime and shortly thereafter you happen to be crushed by a falling boulder, that is not punishment. That is just a fitting accident. For hard treatment to count as punishment, it must be administered knowingly in return for a known wrong, and the person being punished must know what it is he is being punished for. This hard treatment may come in many forms – at one extreme end there is torture, and on the other extreme end is a humane kind of confinement. Rehabilitation is compatible with punishment in this sense. Whether or not rehabilitation involves harsh penalties, in the context of the criminal justice system it almost always involves incarceration or other sanctions. Indeed, rehabilitation is considered by many to be one of the justifications of punishment: on this view, the purpose of punishing criminals is to rehabilitate them.

One last clarification is in order: this resolution is concerned only with legal punishment, that is, punishment by government authorities for violations of the law. Other kinds of punishment, like the punishment administered by a parent to a child, may be of different natures and require different kinds of justification.

Part II: Retributivism

PHILOSOPHY

The traditional and most thoroughly developed answer to the problem of punishment is that punishment is justified because criminals *deserve* to be punished in proportion to the seriousness of their crimes. This view is known as retributivism. Recall the following definition from Alec Walen, professor philosophy at Rutgers University:

Alec Walen (Alec Walen is Professor of Law & Philosophy at Rutgers University.), "Retributive Justice," in Stanford Encyclopedia of Philosophy (Stanford Metaphysics Research Lab, 2014), <https://plato.stanford.edu/entries/justice-retributive/>.

“The concept of retributive justice has been used in a variety of ways, but it is best understood as that form of justice committed to the following three principles:

1. that those who commit certain kinds of wrongful acts, paradigmatically serious crimes, morally deserve to suffer a proportionate punishment;
2. that it is intrinsically morally good—good without reference to any other goods that might arise—if some legitimate punisher gives them the punishment they deserve; and
3. that it is morally impermissible intentionally to punish the innocent or to inflict disproportionately large punishments on wrongdoers.

The idea of retributive justice has played a dominant role in theorizing about punishment over the past few decades, but many features of it—especially the notions of desert and proportionality, the normative status of suffering, and the ultimate justification for retribution—remain contested and problematic.”

Notice that retributivism is committed to the claim that punishing wrongdoers is *intrinsically* good – that is, it is good *in itself*, regardless of its practical implications. This is why Immanuel Kant, a German philosopher and one of the most important defenders of retributivism, famously insisted that a group of people preparing to dissolve their society would have a moral obligation to execute the last murderer in prison before doing so. Even though such an act would bring about no social benefits under those circumstances (since the society in question would dissolve anyhow), according to Kant it ought still to be performed because it would be intrinsically good.

Many arguments have been offered in defense of retributivism. One of the best-known is presented by Michael S. Moore, professor of law and philosophy at the University of Illinois. According to Moore, our moral intuitions – in more familiar terms, our conscience – tells us that those who do wrong deserve to suffer proportional punishments. That is, when we see examples of injustice, *we just can't help believing* that the guilty ought to be punished. Moore contends that our moral intuitions allow us to “see” into the moral realm in the same way that our eyes allows us to “see” into the physical realm, so we should trust our intuitions when they tell us that retributivism is true.¹

¹ Michael S. Moore (Michael S. Moore is Professor of Law and Philosophy at the University of Illinois and is widely regarded as the most prominent living defender of retributivism.), *Placing Blame: A Theory of the Criminal Law* (Oxford: Oxford University Press, 2010): Ch. 2-3.

PHILOSOPHY

Another influential argument for retributivism comes from Hebert Morris, professor emeritus of philosophy at UCLA. According to Morris, when human beings participate in society, we voluntarily accept certain “burdens” in return for the benefits that come with living in community. The most important of these burdens is obedience to the law. When someone breaks the law, they gain an “unfair advantage” over others by shunning that burden, and punishment is necessary to *restore balance*, or *level the playing field*, so to speak. In a just society, no one should benefit unless they are willing to follow the rules. Proportional punishment is, in a way, a policy for preventing free-riding.²

One last argument, which traces back to the work of Joel Feinberg, professor of philosophy at the University of Arizona, is that retribution is necessary to communicate condemnation or disapproval. It is plausible that we have a moral duty to condemn what is evil; it is also plausible that verbal condemnation is sometimes not enough. If, in response to an act of homicide, a government official expressed his disapproval merely by *saying* that he disapproved of the act, he would appear disingenuous. Taking crime seriously – in other words, *really* condemning it – requires something more than verbal censure. It requires punishment. And the more serious the crime, the stronger the condemnation it demands, and the harsher the punishment it justifies.³

Despite boasting so many defenders, retributivism has been subject to many criticisms. The first is that it is merely a kind of vengeance. As we saw in the introduction to the resolution, this objection is misguided: there are plenty of differences between retribution and vengeance. However, there is a second, more serious objection related to the first, and that is that retribution is *sublimated* vengeance; in other words, although it is distinct from vengeance, it is motivated by vengeful desires. According to this view, retribution is simply a restrained – and therefore more socially acceptable – way of exacting vengeance, in the same way (one might argue) that watching or playing excessively violent films or video games is simply a restrained way of satisfying the desire to act violently.

Another serious objection is that retribution can never be put into practice, since it is impossible to determine what counts as a “proportional” punishment in any sort of objective way. The closest thing to an objective standard of proportionality is *lex talionis*, or the law of retaliation, which is often summarized as “an eye for an eye and a tooth for a tooth.” This rule,

² Herbert Morris (Herbert Morris is professor emeritus of philosophy and law at the University of California, Los Angeles.), “Persons and Punishment,” *Monist* 52, no. 4 (1968), <https://www.jstor.org/stable/27902099>.

³ Joel Feinberg (Joel Feinberg was Regents Professor of Philosophy and Law at the University of Arizona. He is known for his work in the fields of ethics, action theory, philosophy of law, and political philosophy as well as individual rights and the authority of the state.), *Doing and Deserving* (Princeton, NJ: Princeton University Press, 1970): 98-101.

PHILOSOPHY

which is present in many ancient legal codes, including the Code of Hammurabi and the Law of Moses, requires that criminals suffer harm that is equal in *kind* to the harm that they caused. But this rule, applied absolutely, would require that mass murderers be executed several times over, that torturers be tortured, and that rapists be raped. Even if it were possible to dispense each of these kinds of inhumane treatment, virtually no one would contend that a just system of punishment should include them. (In defense of the Code of Hammurabi and the Law of Moses, they did not apply *lex talionis* in all situations. But that shows that it was not supposed to be a general rule of proportionality in the first place.) The retributivist might agree but insist that justice requires punishments equal in *magnitude* to the crimes for which they are given, even if it does not require punishments equal in *kind*. But that brings us back to the problem of determining which punishments are proportional in magnitude to which crimes, and different civilizations have had wildly different answers to that question. Just four centuries ago, many westerners believed that it was just to draw and quarter heretics or burn them alive (or both simultaneously). And no objective standard for determining whether they were wrong seems to be forthcoming.

Part III: Consequentialism

It seems doubtful to many that punishment is intrinsically good, given that it involves the intentional infliction of suffering on a person and suffering is almost always an intrinsically *bad* thing. In light of this problem, it may seem more fitting to treat punishment as an *extrinsic* good, or something that is good only because it is a means to an intrinsically good end. Theories that try to justify punishment by its consequences are called, quite appropriately, “consequentialist” theories of punishment. The most influential such theory is deterrence theory, famously espoused by the utilitarian philosopher Jeremy Bentham. Bentham argued that punishment is good because it reduces crime by threatening potential wrongdoers with harm. Since this upholds peace and security, and those are intrinsic goods, punishment is extrinsically good and thus justified.⁴

Deterrence theory is not the only consequentialist theory of punishment on offer. Mitchell N. Bermann, professor of philosophy at the University of Pennsylvania, offers a helpful survey of the options:

Michael N. Bermann, “The Justification of Punishment,” in *The Routledge Companion to Philosophy of Law*, ed. A. Marmor (London: Routledge, 2012), 145.

⁴ Jeremy Bentham (Jeremy Bentham was an English philosopher, jurist, and social reformer regarded as the founder of modern utilitarianism.), *An Introduction to the Principles of Morals and Legislation* (Ontario: Batoche Books, 1999): Ch. XIV-XV.

PHILOSOPHY

Deterrence theory, in both general and specific forms, involves both a view about the good state of affairs that punishment can produce—namely, the state in which there are fewer acts of antisocial violence or rights violations than would otherwise obtain—and a view about the mechanism by which punishment produces that state—namely, by bolstering the credibility of conditional threats. But deterrence is not the only means by which punishment might cause a reduction in antisocial aggression and the like. For one thing, punishment can incapacitate offenders from reoffending, either permanently (say, by execution) or for a time (by incarceration)—though incarceration in practice might do more to shift the incidence of violence and aggression, from the general community to the prison population, than to substantially reduce the total number of harmful or wrongful acts. Also, punishment might in theory reduce violence and aggression by reducing offenders' inclinations to offend, by reforming their character, strengthening their attachment to moral or legal norms and even by providing them with job-related skills. These mechanisms, especially the last, are what people have in mind when they invoke "rehabilitation" as a justification for punishment. (It is worth noting, though, that a wrongdoer does not likely develop socially productive skills as a consequence of punishment itself, but as a consequence of whatever training prison officials can provide auxiliary to an offender's punishment.) Relatedly, if punishment can deter what we might call initiatory aggression, it can also deter retaliatory aggression—aggression by victims and their families against initial offenders—by assuring victims that the wrongs they have suffered will be vindicated.

Of course, the most important alternative to deterrence Bermann mentions for our purposes is rehabilitation. Rehabilitation may be a better justification for punishment than deterrence on the consequentialist view if it is possible to show that it brings about more social good than deterrence does. (Indeed, for a through-and-through consequentialist, that is the *only* question that matters.) The proponent of rehabilitation could try to show this either by showing that rehabilitation reduces crime more effectively than deterrence or by showing that rehabilitation leads to other social goods that deterrence cannot secure, like lower costs, better mental health, fewer fatherless homes, higher rates of employment, and so on.

The most serious objection to consequentialist theories of punishment is that they appear to sanction unjust punishments in certain circumstances. The clearest statement of this problem comes in the form of a thought experiment offered by the Australian philosopher H. J. McCloskey, in which a judge must decide whether to condemn an innocent black man to death in order to prevent a deadly race riot. The trouble for the consequentialist is this: condemning the man would lead to the best overall consequences (no riot) and yet would be unjust.⁵ Given the constraints of his theory, in order to overcome the objection, the consequentialist must explain why allowing the riot would lead to a greater good in the long run. As it turns out, there are plausible ways to do this, for instance by arguing that the justice system would collapse if judges

⁵ H. J. McCloskey (Henry John McCloskey was an Australian moral philosopher and writer and was Professor of Philosophy at La Trobe University in Melbourne.), "A Non-Utilitarian Approach to Punishment," *Mind* 8, no. 1–4 (1965): 249–63.

habitually condemned innocent men when they judged it to be expedient. Consequentialism has its weaknesses, but it is far from untenable.

Part IV: Mixed Theories

As you have likely gathered by now, retributivism and consequentialism have their own unique advantages: consequentialism nicely explains why punishment is sometimes a good thing, and retributivism explains why disproportional or undeserved punishment is a bad thing. Some philosophers have tried to combine the advantages of both theories by defending a mixed theory. The most widely accepted mixed theory is called negative retributivism. According to Richard Lippke, a philosopher of criminal law at Indiana University,

Richard Lippke (Richard Lippke is chair of the department of philosophy at Indiana University. He works at the intersection of philosophy, law, and criminology, where he writes about legal punishment, sentencing, criminal law and criminal procedure.), "Elaborating Negative Retributivism," *Philosophy and Public Issues* 5, no. 1 (2015): 58-9, http://fqp.luiss.it/files/2016/05/4_Lippke_Elaborating-Negative-Retributivism_PPI_vol5_n1_2015.pdf.

As it is standardly formulated, negative retributivism posits that the general justifying aim of legal punishment is the reduction of crime, whether such reduction is affected by deterrence, incapacitation, or rehabilitation. However, negative retributivism insists that the pursuit of that aim is to be limited by two retributive constraints, one forbidding the intentional or knowing punishment of the innocent, the other forbidding disproportionate punishment of the guilty. According to negative retributivism, as contrasted with positive retributivism, we need not punish the guilty (if doing so would not, for some reason, reduce crime) or proportionately punish the guilty (again, if doing so would not, for some reason, reduce crime). It is undeserved punishment that negative retributivism rules out; deserved or (fully) proportional punishment is not required by it. Brooks' characterization of negative retributivism includes the first constraint, but misses the second, as he claims that the theory permits disproportionate punishment of the guilty if crime reduction considerations so dictate. This is mistaken, or so I contend. Negative retributivism is premised on an awareness of the pitfalls of a simple crime reduction approach to legal punishment's justification. It has been argued that such an approach permits both punishment of the innocent and over punishment of the guilty. The two retributive constraints have been thought necessary to address these defects. Recently, I argued for a third constraint, one forbidding degrading punishment of the guilty. I come back to this in the next section.

It is worth noting that a mixed theory can accommodate retribution and rehabilitation at the same time: rehabilitation provides the justifying aim of punishment (achieving some social good or other by reforming criminals), while retribution provides the limits within which punishments are just.

The challenge for negative retributivism is to explain why wrongdoing *permits* us to punish someone without *requiring* us to punish them. Remember, each of the arguments we examined for standard retributivism (sometimes called "positive" retributivism in contrast to negative retributivism) were supposed to show that proportional punishment is good *in itself*, and

that is not what the mixed theory says. This challenge is not impossible to overcome, but it requires creative arguments.

Part V: Abolitionism

A lesser-known but intriguing family of theories state that punishment *cannot be* justified – in other words, there is no good answer to the problem of punishment. According to abolitionism, governments should not be in the business of intentionally inflicting *suffering* on criminals at all. Rather, governments should seek to rectify criminal acts in some other way.

One such theory is defended by Derk Pereboom, professor of moral philosophy at Cornell University. Pereboom says that free will, and thus moral responsibility, is an illusion, so criminals do not deserve to suffer harm. Thus, we ought to deal with them in ways that protect the public but do not impose harsh penalties on them: for example, by rehabilitating or confining them with a view to helping them and others flourish.⁶ Pereboom's theory is based on rather obscure (though interesting) philosophical arguments and is not at all likely to persuade non-specialists. But a more moderate version of the theory is easier to defend. The abolitionist may argue that our institutions of punishment *as they currently stand* ought to be abolished, because they impose harsh treatment primarily on those who are so destitute that they have no choice to turn to crime, perhaps because they are disabled, illiterate, addicted, or mentally ill. If the majority of criminals fit this description (and there is some empirical evidence out there that they do), then it arguably makes little sense to orient our criminal justice system around punishment. Instead, we ought to focus on assisting them in overcoming those barriers through rehabilitation.

A second, more widely accepted abolitionist theory is the theory of pure restitution, most vigorously defended by David Boonin. According to Boonin, the criminal justice system ought to focus on upholding the rights of victims, not inflicting suffering on wrongdoers. When victims suffer harm, justice demands that they be compensated, and punishing criminals does nothing to compensate them. Instead of imposing harsh treatment, governments should require criminals to do whatever is necessary to restore their victims to the condition they enjoyed prior to being victimized. While this may be (and usually is) *unpleasant* for the offender in question, Boonin argues that it does not constitute punishment, since the *purpose* of requiring restitution is to uphold the victim's right to compensation, not to make the criminal suffer. Any suffering the

⁶ Derk Pereboom, "Free Will Skepticism and Punishment," in *The Future of Punishment*, ed. T. A. Nadelhoffer (Oxford: Oxford University Press, 2013), 49-78.

PHILOSOPHY

criminal experiences is merely a side-effect of the administration of justice.⁷ (Note that it is possible to treat restitution as a form of punishment and thus as a non-abolitionist answer to the problem of punishment – in particular, it can be integrated into a version of retributivism that addresses what both victims *and* criminals deserve.)

The greatest challenge for the theory of pure restitution is explaining what should be done when restitution cannot be made to the victim – for example, in cases of homicide. Boonin answers this objection systematically and in great detail; I do not have space to rehearse his arguments here, but you can read Chapter 5 of his book for more details (see footnote below).

Conclusion

It should be clear by now that the philosophical problem of punishment is tremendously relevant to this year's resolution. Indeed, the resolution *may just be* a simplified version of the problem. Every one of the arguments presented above could (and likely will) make an appearance in debate rounds this season. Beyond that, though, the reasoning surveyed above leads to a rather unexpected consequence for NEG, one that will make certain AFF cases very difficult (though by no means impossible) to refute. I will explain that consequence and ways to address it in the next installment.

⁷ David Boonin (David Boonin is Professor of Philosophy at the University of Colorado Boulder and Director of the Department's Center for Values and Social Policy.), *The Problem of Punishment* (Cambridge: Cambridge University Press, 2012), Ch. 5.