PART ONE

Introduction

Retributivism and the Criminal Law

CHAPTER

I

Criminal Law, Punishment, and Desert

Ultimately, what underlies the criminal law is a concern with harms that people suffer and other people cause – harms such as loss of life, bodily injury, loss of autonomy, and harm to or loss of property. The criminal law's goal is not to compensate, to rehabilitate, or to inculcate virtue. Rather, the criminal law aims at preventing harm.

This admission may seem puzzling, given that the authors of this book have argued in previous writings, and will continue to argue here, that whether a criminal defendant actually causes harm is immaterial to whether he should be deemed to have violated the criminal law and is likewise immaterial to the amount of punishment he should receive. But these claims do not entail that the criminal law is not ultimately concerned with harm causing. Quite the contrary.

I. The Criminal Law and Preventing Harm

To explain how we can admit that the criminal law's primary concern is the prevention of harm yet still maintain that the actual occurrence of that harm is immaterial, we will begin by exploring ways that harm might be prevented. One way to prevent the harms with which the criminal law is ultimately concerned is to make the causing of harms to others more difficult. There are three strategies for doing this. One strategy is to *increase the difficulty* of causing harm by increasing the effort or natural risk required to cause harm. We put money into safes that are difficult to crack. We put our castle behind a deep moat, perhaps filled with alligators, and build high walls. We put our high-security establishment behind an electrified fence. In all sorts of ways, we try to make harming us difficult by making it impossible, costly, or risky.[1]

The second strategy for making harming more difficult is to impose penalties on those who attempt or succeed in harming us. Penalties are meant to raise the expected cost of the harming act (amount of penalty times likelihood of detection, conviction, and so forth). In this respect, penalties are quite similar to the first strategy. If I trespass by jumping into your moat, the alligators might scarf me, or I might drown. Trespassing, therefore, looks less appealing. The fine – a penalty – that I might have to pay similarly makes trespassing look costlier and thus less appealing. Here, the strategy is one of *deterrence through prospective penalization*.

Notice that these two strategies bear no relation to the would-be harmer's desert. Take prevention. If the trespasser drowns or is killed by the alligators, we do not consider his death as what he "deserved" for trespassing. We may place limits on prevention strategies, particularly because they do not distinguish between the culpable and the innocent (alligators might find both equally tasty). Indeed, prevention seems to require both a wrongful act and notice of the consequences risked – especially if they exceed the wrongdoer's desert. On the other hand, these limitations do *not* include the requirement that the prevention

[1] *See* Larry Alexander, "The Doomsday Machine: Proportionality, Prevention, and Punishment," 63 *Monist* 199, 210 (1980).

strategy be proportional to the wrongdoer's desert.[2] Although we would say that in some sense, by risking death through an act he had no right to undertake, an actor brought his death on himself, we would not say that a tulip thief *deserved* to be eaten by alligators.

Just as the enterprise of prevention may be disproportionate in terms of the harms risked by the wrongdoer relative to his desert, so too may penalties premised on deterrence. For example, if possession of marijuana is a crime that many people are tempted to commit, under a deterrence theory the state may be justified in imposing a significant jail term to prevent the possession of even the smallest amount of marijuana. As is frequently pointed out, when we impose harsh treatment solely to deter, there is no necessary connection between the penalty we impose and the offender's desert. Indeed, because any penalty we impose will have failed to deter at least the offender we are imposing it on, deterrence would have warranted a higher penalty. Indeed, from a pure deterrence standpoint, the ideal penalty is one so draconian that it achieves 100 percent deterrence and therefore never has to be imposed.

There is actually a third strategy for preventing harm-causing conduct, and that is the strategy of *incapacitating those who we predict are likely to cause harm* if they are not incapacitated. Again, preventively detaining those predicted to be harmful bears no relation to the desert of those detained. One can be dangerous without being deserving of bad treatment. Assume, for example, that we can predict with some reasonable degree of certainty that if a four-year-old boy enjoys torturing puppies, he will later harm his fellow human beings.[3] If we lock him away now, we are locking him away not for what he has done (to human beings) but for what he might do. He is dangerous for what he *is*. He can *deserve* harsh treatment, however, only for his chosen acts (or, in some cases, his chosen omissions). Although preventive detention may likewise be subject to limitations, desert is not among them.

The alternative to these three strategies for preventing harmful acts, all of which attempt to make harmful acts physically difficult,

[2] *Id*. at 213.

[3] *Cf.* Jim Stevenson and Robert Goodman, "Association between Behavior at age 3 Years and Adult Criminality," 179 *British J. of Psychiatry* 197, 200 (2001) (finding that "[e]xternalising behaviours such as temper tantrums and management difficulties [e.g., non-compliance] were associated with adult convictions, in particular with violent offences").

impossible, or risky, is to inculcate norms that are meant to guide people's choices. The norms inform people of the reasons that should govern their choices, *and the inculcation of such norms involves as its corollary the inculcation of reactive attitudes toward those who comply with and those who violate the norms*.[4] The negative reactive attitudes, to be directed at those who choose to violate the norms, include both blame and the sense that punishment is fitting. When we say that, by choosing as the norms forbid, the chooser deserves punishment, we are invoking the reactive attitude that punishment of a certain amount is a fitting response to the choice. Thus, the criminal law both creates and reflects value by announcing which conduct is sufficiently wrong to deserve blame and punishment.

Such a view presupposes that people act for reasons and that the law can influence those reasons. Moreover, it considers an actor *deserving* of punishment when he violates these norms that forbid the unjustified harming of, or risking harm to, others – that is, failing to give others' interests their proper weight. This approach to preventing harm, although setting forth the types of harms and risks that are forbidden, focuses on the actor's reasons and thus derives its ability to prevent such harms from the capacity and opportunity that agents have to act or abstain from acting for reasons.

It is this last alternative that we believe the criminal law should, and to some (imperfect) extent does, adopt. What we intend to do in this book is to explore what the doctrines of the criminal law would look like if they were structured (primarily) by the concern that criminal defendants receive the punishment they deserve, and particularly that they receive no more punishment than they deserve. We argue that the elements of crimes and defenses thereto should pick out those factors bearing on the defendant's negative desert, either to establish it or to defeat it. In our view, it is the defendant's decision to violate society's norms regarding the proper concern due to the interests of others that

[4] The essentially constitutive relation between "Don't do that because it's wrong" addressed to a responsible moral agent and the reactive attitudes implied thereby is frequently noted. For a recent example, *see* John Tasioulas, "Punishment and Repentance," 81 *Phil.* 279, 294–301 (2006); James Lenman, "Compatibilism and Contractualism: The Possibility of Moral Responsibility," 117 *Ethics* 7, 11–12 (2006).

establishes the negative desert that in turn can both justify and limit the imposition of punishment.

II. Questions about Retributivism

Hence, what we elaborate can be called a retributive theory of the criminal law because the structure of the criminal law that we propose is dictated by a retributive theory of criminal punishment. However, our argument in the remaining chapters deals solely with the culpable choices that give rise to retributive desert and does not focus on the retributive theory in which they are embedded. We do not more fully defend retributivism against competing theories than we just have, although our arguments about what makes an individual culpable and worthy of punishment no doubt implicitly reflect our position. We also recognize that there are a number of outstanding issues regarding retributivism and hence a retributive theory of criminal law. We touch on these issues, although our theory does not depend upon their full resolution.

A. WEAK, MODERATE, OR STRONG RETRIBUTIVISM?

First, even for those, like us, who believe that desert is a necessary condition for punishment, there remain questions about the exact relationship between desert and punishment. There are three possible positions. The first is that negative desert is merely necessary but not sufficient for punishment (weak retributivism – or perhaps, more accurately, desert-free consequentialism side-constrained by negative desert). The second is that negative desert is necessary and sufficient for punishment but that desert does not mandate punishment (moderate retributivism). The final position is that desert is necessary and sufficient for punishment and mandates punishment (strong, Kantian retributivism).

In our view, the middle, moderate position seems most preferable. We believe that weak retributivism is too weak to guide the criminal law substantially; so long as no criminal receives *more* punishment than he deserves, the criminal law could be structured completely by

consequentialist considerations. In such cases, unless some additional good were served, individuals who deserve punishment would be beyond the reach of the criminal law. To us, however, deserving punishment seems to be a weighty enough reason to punish someone.[5]

On the other hand, strong retributivism is too strong. We could spend all available resources and risk all sorts of terrible harms – for example, accidentally convicting the innocent, taking resources from health and safety, and so forth – trying to ensure that all of the negatively deserving receive their due. Surely government should not be monomaniacally concerned with punishing the guilty at the expense of all other interests.[6]

Thus, only moderate retributivism looks eligible for our purposes. In contrast to strong retributivism, moderate retributivism entails that some of the guilty will not receive their negative due. In contrast to weak retributivism, however, it entails that sometimes punishment will serve no purpose other than to see that the guilty get what they deserve.

Notice that moderate retributivism has the following notable features. First, the moderate retributivist position has both a deontological and consequentialist aspect. The moderate retributivist position is deontological in placing a side constraint on punishment, namely, that no one should be (knowingly) punished more than that person deserves. (What *risk* of undeserved punishment we may subject people to is taken up later in this chapter and then again in Chapter 8, in which we also raise further questions about the implications of the deontological side constraint.) The position is consequentialist in that it rejects a deontological duty to see that all the guilty receive the punishment they deserve. Instead, it counts just punishment as one good among

[5] *See generally* Leo Zaibert, *Punishment and Retribution* 214 (2006) ("[D]eserved punishment is an intrinsic good"); Mitchell N. Berman, "Punishment and Justification," esp. note 59 at 32 (working paper, December 15, 2006), available at SSRN: http://ssrn.com/abstract=956610. Berman distinguishes the justification for the criminal's suffering punishment from the justification for inflicting it, pointing out that one might concede that criminals deserve to suffer while at the same time arguing that inflicting such suffering is a violation of their rights. He goes on, however, to deny that any such right against infliction of deserved suffering exists. *See id.* at 42–48. *See also* John Martin Fischer, "Punishment and Desert: A Reply to Dolinko," 117 *Ethics* 109 (2006).

[6] *See* Zaibert, *supra* note 5, at 153–155; Mark C. Murphy, *Natural Law in Jurisprudence and Politics* 144–146 (2006).

many, and one that can be outweighed by other goods that punishing the deserving puts at risk.[7] Not only does the side constraint of not punishing more than is deserved prevent maximizing the number of punishments of the deserving – otherwise, it would be permissible to punish an innocent person or to punish a culpable person more than he deserves in order to maximize the number of just punishments – but it countenances less than monomaniacal pursuit of deserved punishment within the bounds of that deontological constraint. Deserved punishment is a positive value, but it is not the only positive value. Seeing that people receive their negative desert may be an aim of criminal punishment for the moderate retributivist as it is for the strong retributivist; but for the moderate but not the strong retributivist, other values define the circumstances in which pursuit of that goal is properly undertaken.[8] On the other hand, for the moderate retributivist as for all retributivists, undeserved punishment, if administered with knowledge that it is undeserved, is always a trumping *dis*value.

Beyond this asymmetry between the positive and negative aspects of deserved punishment, a further feature of moderate retributivism is that it covers a wide range of positions on just how weighty a positive value deserved punishment is. One might deem deserved punishment to be a very weighty value, justifying huge social costs in its pursuit. On the other hand, one might deem it to be of much less weight, justifying very little expenditure of resources or risk to other values. Moderate retributivism occupies a large territory between weak and strong retributivism, with weak and strong retributivism serving as the limiting cases of the weight of the positive value of deserved punishment. (Weak retributivism represents zero weight relative to the strength of the side constraint forbidding giving anyone more punishment than is deserved, *even in order to achieve a greater number of just punishments*; strong retributivism represents infinite weight relative to other values.)

[7] For a discussion of how the law currently trades off retributive desert against other values, *see* Paul H. Robinson and Michael T. Cahill, *Law without Justice: Why Criminal Law Doesn't Give People What They Deserve* (2006).

[8] Murphy regards these other values as the sine qua non conditions of retributive punishment, not its aim. *See* Murphy, *supra* note 6, at 146. We believe that punishment can be called retributive if it is both constrained by negative desert and regards inflicting deserved punishment as a positive value rather than its sole aim.

The moderate retributivist must at the end of the day come up with a theory for how the value of retributive justice trades off against the values of societal welfare, distributive justice, and corrective justice. We endorse moderate retributivism, but we take no position on the weight of deserved punishment relative to other values. For our purposes here, it is unnecessary that we do so. It is enough that the weight of deserved punishment be sufficient to make desert a central focus of criminal law.

B. MEASURING DESERT

A corollary to our rejection of strong retributivism is the potential for comparative and noncomparative injustices, and this effect introduces a second conundrum for retributivism – the question of how retributive desert is measured. We must first ask whether desert is comparative or noncomparative. In other words, is the punishment an offender deserves a function solely of how much similar offenders are punished, or is there a specific amount of punishment that each offender deserves irrespective of how much others are punished? A system of punishment that required only comparative justice would be satisfied if all bank robbers received one-day imprisonment. In contrast, in a system that viewed desert from a noncomparative ("cosmic") perspective, if a bank robber A received ten years (the "cosmic" amount) but bank robber B received one day, then from the standpoint of retributive justice, this system could be criticized only insofar as B did not get his "just" deserts but not because of the seemingly unfair discrepancy between A and B. In our view, desert is itself noncomparative, but there are additional constraints on the imposition of punishment that speak to fairness, including that similarly situated defendants be treated similarly. If defendants A and B are cosmically – noncomparatively – deserving of equal punishment, but only A, who is black, receives his just deserts, whereas B, who is white, receives a lesser punishment, A's complaint is not that *his* punishment is inappropriate but is, instead, that an illegitimate norm – "whites should be treated better than blacks" – is at work and has led to B's being treated better than he deserves. The remedy may (or may not) be to reduce A's punishment if B's punishment for some reason cannot be increased. But A can have no complaint against *his* punishment.

This leads to a second measurement question: how does negative retributive desert, with which the criminal law is concerned, mesh with positive retributive desert (reward) and with distributive desert? Desert appears to be a single positive or negative unit of measurement. The currencies we employ in rewarding positive and negative desert – for example, pleasure and pain, liberty or its loss, or money or its loss – are fungible across any positive-negative desert divide. If a person serving ten years in prison performs a heroic act, he might be rewarded by getting special privileges in prison, or by having his term of imprisonment shortened, or both. Is it meaningful to ask whether his negative desert and hence his punishment were decreased, or whether instead his positive desert and reward were increased?

If one believed that everything – benefits as well as harms – should be distributed according to desert, positive as well as negative, then retributive punishment would just be an aspect of a more general scheme of distribution according to desert. This leads to yet another set of questions. If A and B commit the same crime, but A is happy and wealthy and B is unhappy and poor, do they receive differential amounts of punishment so that they are similarly situated once the punishment is imposed? Moreover, does it matter who does the distributing? What if C leaves a bank robbery and her criminal conduct warrants an "alpha" level of punishment, but as she flees the scene, she is hit by lightning and suffers an "alpha" amount of pain? Should the state still inflict the same degree of harm? In practice, a court may inflict a "shaming" punishment,[9] wherein a defendant is subject to public disapproval for his conduct; but if the defendant's loss of public respect is not the product of a judicially imposed sanction but just the product of the defendant's conviction, courts may ignore this "fall from grace" as irrelevant to what further sanction should be imposed.[10]

[9] This example is for illustrative purposes only. We are not taking a position on whether shaming is an appropriate form of punishment.

[10] *See, e.g.*, Gertrude Ezorsky, "The Ethics of Punishment," *in Philosophical Perspectives on Punishment* xi (Gertrude Ezorsky, ed., 1972); Jeffrey Moriarity, "Ross on Desert and Punishment," 87 *Pac. Phil. Q.* 231, 232–236 (2006); United States v. Bergman, 416 F. Supp. 496 (S.D.N.Y. 1976). *But see* Douglas N. Husak, "Already Punished Enough," 19 *Phil. Topics* 79 (1990) (arguing that public disapproval can reduce the amount of deserved punishment).