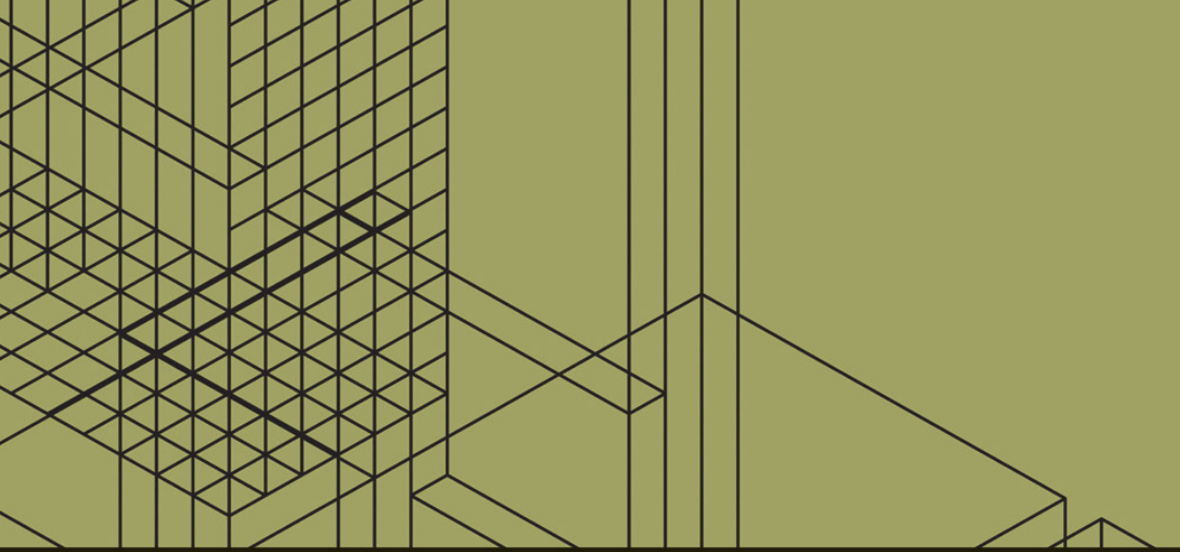
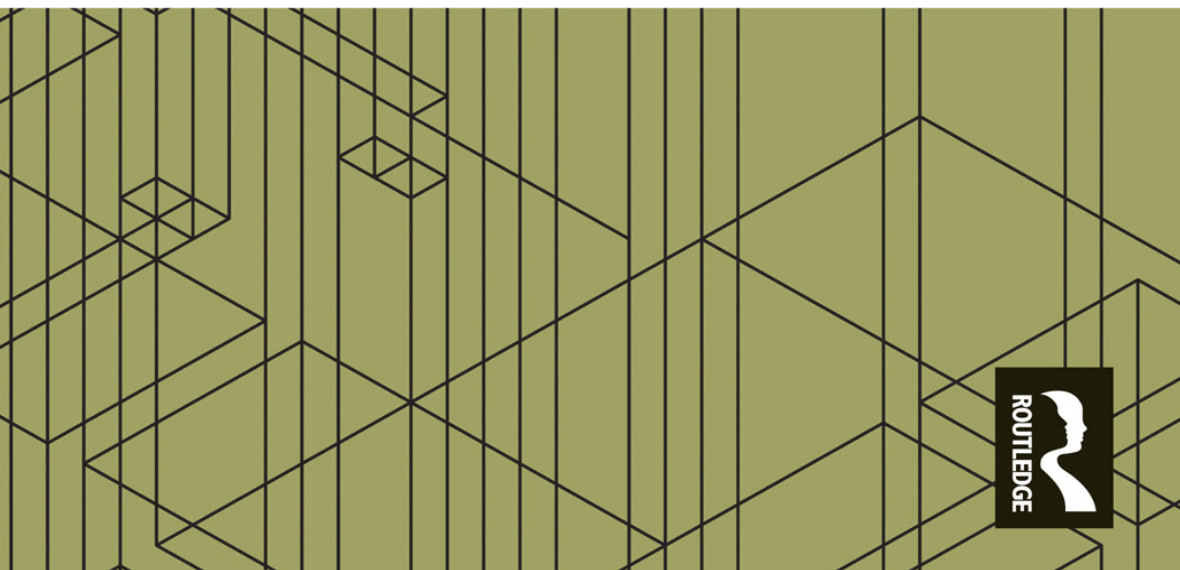


Routledge Research in Legal Philosophy

A THEORY OF LEGAL PUNISHMENT

DETERRENCE, RETRIBUTION, AND THE AIMS OF THE STATE

Matthew C. Altman



A Theory of Legal Punishment

Deterrence, Retribution, and the Aims of the State

Matthew C. Altman

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2021
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
52 Vanderbilt Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Altman, Matthew C., author.
Title: A theory of legal punishment: deterrence, retribution, and the aims of the state / Matthew C. Altman.
Description: Milton Park, Abingdon, Oxon; New York, NY: Routledge, 2021. |
Series: Routledge research in legal philosophy |
Includes bibliographical references and index.
Identifiers: LCCN 2020049904 (print) | LCCN 2020049905 (ebook) |
ISBN 9780367698102 (hardback) | ISBN 9781003143352 (ebook)
Subjects: LCSH: Punishment–Philosophy. | Criminal law–Philosophy. |
Punishment in crime deterrence. | Lex talionis.
Classification: LCC K5103 .A477 2021 (print) | LCC K5103 (ebook) |
DDC 364.601–dc23
LC record available at https://lccn.loc.gov/2020049904
LC ebook record available at https://lccn.loc.gov/2020049905

ISBN: 978-0-367-69810-2 (hbk)
ISBN: 978-0-367-69816-4 (pbk)
ISBN: 978-1-003-14335-2 (ebk)

Typeset in Galliard
by Deanta Global Publishing Services, Chennai, India

Preface

Punishment, like slavery, is a peculiar institution. We as a society deprive people of their freedom, have them perform forced labor, and sometimes even kill them,[1] yet we seldom give it a second thought. One big difference is that we punish people who are responsible for crimes rather than enslaving people because of their race. Following the American Civil War, with the country thinking about when we are justified in taking away people's freedom, the Thirteenth Amendment was enshrined in the U.S. Constitution: "Neither slavery nor involuntary servitude, except as a punishment for crime whereof the party shall have been duly convicted, shall exist within the United States, or any place subject to their jurisdiction." Slavery violates people's rights unless it is used as a form of legal punishment. There are different possible reasons for exempting criminals specifically from this prohibition, including the belief that they deserve hard treatment for their past acts or that the threat of hard treatment will deter potential criminals in the future. Historically, these two different lines of argument – retributivism and consequentialism – have dominated the philosophical conversation.

Recent developments, especially in the United States, are prompting us to reexamine the purpose and value of punishment. A punitive approach to criminality and exaggerated worries about public safety have led to a huge increase in the U.S. prison population since 1970, far outpacing population growth, changes in the crime rate, and incarceration rates in other countries. Prisoners are subjected to brutal conditions, including sexual violence, torture, and death, due to prison overcrowding, lack of funding, and lack of oversight. And the effectiveness and cost of capital punishment has come under scrutiny, with declining executions, recent abolitions and moratoriums, and near-record low levels of public support.

Historians can explain how these things have happened and social scientists can measure the effects on offenders and the community, but it is up to philosophers to determine whether a penal system is justified. Philosophers of law, in particular, investigate the normative foundations of law and reasons to punish. At its most useful, this sort of inquiry can give us moral standards to evaluate existing policies and guide reform efforts, if necessary. For example, because mass

1 … especially people of color in the American South. See Garland (2010).

incarceration is due in part to punitive practices such as increased sentences for drug possession and three-strikes laws, we ought to evaluate the theoretical justification of legal punishment in order to decide whether such practices further the cause of justice and the consequences are worth it.

This book adds to a growing body of philosophical research on punishment. Although a lot of excellent work has been done recently, many theorists fail to appreciate the implications of distinguishing punishment as an institution from the distribution of individual punishments, do not adequately take social scientific research into account, make overly general claims about people's supposed lack of responsiveness to incentives, and do not explicitly relate theories of punishment to the political institutions that would apply them. By identifying separate roles for the legislature and the criminal judiciary, I incorporate both consequentialism and retributivism into a mixed theory. Although there are key differences, the resulting two-tiered model of punishment echoes theories that were developed about seventy years ago. Many philosophers abandoned consequentialism in the 1970s, primarily because of concerns about the social sciences' lack of predictive power. However, with recent advances in criminology and a recognition that our current approach to punishment is failing, it is time to revive this tradition.

Introduction

Punishment, it seems, has always been with us. Gods have visited divine retribution on humanity and parents have corrected their children ever since there have been gods and parents. Legal punishment exists throughout the world. If a nation were to do away with punishment entirely, we would think it is at best a dangerous social experiment and at worst a devolution into anarchy. In *Brave New World*, Aldous Huxley imagines a world without punishment, where people are controlled by genetic engineering, drugs, propaganda, and social reinforcement. The cure is worse than the disease. History, literature, and common sense, as misguided as it sometimes is, all seem to be telling us that punishment is necessary if we want to have both individual freedom and social order.

The ubiquity of punishment makes it seem natural and hardly in need of justification. However, the things that we do to criminals would be wrong in just about any other context: forcibly taking their property (fines), putting them in cages (prison), and killing them when they pose no immediate threat (execution). Punishment also involves a number of normative claims that are not settled: about the wrongness of criminals' behavior, the value of society, the necessity of hard treatment, and the justice of enforcement. Practical philosophy can help to address these questions, not only because it includes the study of ethics and the foundation of law, but also because of philosophy's interdisciplinarity and the fact that several intersecting fields of research converge on the study of punishment. This book is an attempt to bring philosophical thinking to bear on punishment: to propose and defend a mixed theory, and to explain its real-world implications.

Two major theories

Although legal punishment had been addressed by prior philosophers, it only became a distinct object of philosophical scrutiny in the eighteenth century. Since then, two theories have dominated the conversation: consequentialism and retributivism. According to consequentialism, the state reduces crime by threatening potential criminals with punishment if they break the law (general deterrence). The actual imposition of punishment not only makes the threat credible, in general and in the minds of those who are punished (specific deterrence), but it also keeps criminals from committing other crimes (incapacitation) and improves

their characters so they are less likely to reoffend upon release (rehabilitation).[1] Historically, the main proponents of consequentialism are Cesare Beccaria (1738–1794) and Jeremy Bentham (1747–1832), both of whom appeal to the principle of utility to justify the institution of punishment as a mechanism for maximizing overall happiness. As Beccaria puts it, punishment is justified only insofar as it is necessary for "defending the repository of the public well-being from the usurpations of individuals" (1995, 10). Although they defend the law as an incentive system to reduce crime, they also criticize contemporary penological practices, with Beccaria focused on the overuse of the death penalty (1995, 66–72) and Bentham focused on prison reform (1970, esp. ch. XII–XV [pp. 143–86]). If the same social aims can be achieved more effectively and with less suffering – with imprisonment instead of the death penalty and a more humane prison system that encourages moral improvement – then we ought to change the system.

Consequentialists evaluate a punishment or policy by comparing its consequences to the consequences of other options. We ought to do what produces the best outcomes. In this way, consequentialism is forward-looking. Different versions of consequentialism are distinguished by which ends they say we ought to achieve. According to utilitarianism, for example, any rule or action ought to maximize happiness and minimize suffering (Bentham 1970, ch. I, §§2–5 [pp. 11–12]; Mill 2001, 7). Punishment can be justified if the burden imposed on criminals efficiently reduces the greater harm caused by criminal activity.

By contrast, retributivism is backward-looking in the sense that we ought to punish offenders because they deserve it for their past conduct. Even if nothing else is accomplished as a result, on this view punishment is intrinsically valuable. Criminals may deserve to be punished because they incurred a debt to society (repayment theory) or took unfair advantage of the system (fair-play theory), such that punishment would annul the wrong (annulment theory) or publicly denounce it (denunciation theory) – there are different ways of understanding the basis of that desert claim (Cottingham 1979; Walker 1999).[2] On any of these variations, retributivism refers to people's guilt in determining both who ought

1 Bentham writes:

> The immediate principal end of punishment is to control action. This action is either that of the offender, or of others: that of the offender it controls by its influence, either on his will, in which case it is said to operate in the way of *reformation*; or on his physical power, in which case it is said to operate by *disablement*: that of others it can influence otherwise than by its influence over their wills, in which ease it is said to operate in the way of *example*. (1970, ch. XIII, §2n [p. 158n])

2 Michael Davis says that desert theory is only one kind of conceptual theory among others. However, he also says that "desert theories seem to be the direct descendants of traditional retributivism" (2009, 90–91). Like Davis, Cottingham (1979) and Walker (1999) claim that desert theory is only one retributivist position, and that other versions of retributivism such as fair-play theory are alternatives to desert theory rather than kinds of desert theory. When I refer to retributivism, however, I mean it in the classical sense that I have defined here, as a synonym for desert theory. Thus, I agree with Honderich, who writes:

to be punished and how much. On the latter issue, retributivists disagree about whether a punishment ought to be equivalent or only proportionate to the crime (equality retributivism versus proportional retributivism). Retributivists also disagree about what desert claims entail regarding punishment, whether

1) guilt is necessary and sufficient to justify punishment (positive or bold retributivism), and punishment is morally obligatory (extreme or maximalist retributivism);
2) guilt is necessary and sufficient to justify punishment (positive or bold retributivism), and punishment is morally permissible (moderate or minimalist retributivism); or
3) guilt is only necessary for punishment, and there must be some other positive (usually consequentialist) reason to punish, such as promoting the common good (negative, weak, or modest retributivism; or side-constrained consequentialism).[3]

Having desert as a positive reason to punish does not necessarily, by itself, establish that one has a duty to punish. It could mean either (1) that one has a duty to punish (that the reason *is* overriding) or (2) that one would be justified in punishing even if there is no other reason to do so (that the reason *may be* overriding).

It is or becomes pretty clear that all of these claims about desert, desert claims, are ways of saying one main thing, whatever it is. That some have slightly different implications does not much affect the fact. They are all ways of saying that a certain penalty or punishment stands in a certain relation to a past offence or something about a past offence, which relation makes the punishment right – morally obligatory or at least permissible. The relation is what is crucial and fundamental to retribution. (2006, 22)

Duff also says that the different varieties of retributivism set out by Cottingham are "different attempts to articulate and explain" the claim that "punishment is justified if and only if it is deserved in virtue of a past crime" (1996, 7). At some point, this becomes a verbal quibble, since both sides recognize that there are different varieties of retributivism.

3 Dolinko identifies two views: "bold" (positive) retributivism and "modest" (negative) retributivism (1991, 541–44). Golding calls negative retributivism a "*minimalist* position" (1975, 85). L. Alexander (1980) distinguishes among "weak retributivism" (desert as a necessary condition), "moderate retributivism" (desert as necessary and sufficient), and "extreme retributivism" (we *must* punish the deserving). *Black's Law Dictionary* distinguishes "maximalist retributivism," "minimalist retributivism," and "negative retributivism" (Garner 2019, 1575). The maximalist position requires punishment for the guilty, the minimalist position allows it, and the negative position is explicitly consequentialist and uses retribution only as a side constraint ("side-constrained consequentialism"). Mackie defines negative retributivism as the view that someone who is innocent may not be punished, and separates this position from permissive retributivism, which says that the guilty may be punished (1982, 4; 1991, 678–79). I cannot think of a form of negative retributivism that would not hold *both* that the guilty may be punished *and* that the innocent may not be punished, so I conflate negative retributivism and permissive retributivism. Zaibert rejects these labels and claims instead that retributivism provides a *prima facie* reason to punish that may be outweighed by "other factors," including "the sorts of considerations typically associated with utilitarianism" (2006, 213–24).

Those two positions differ regarding the weight of the desert claim. So, talking only about positive (1 and 2) and negative (3) retributivism, as many philosophers do, elides an important distinction between the extreme positive position (1) and the moderate positive position (2).

Historically, the main proponents of retributivism are Immanuel Kant (1724–1804) and G. W. F. Hegel (1770–1831). Kant claims that we have a positive moral duty to punish the guilty and that failing to give people what they deserve implicates us in their crimes (1996b, 6:331–35).[4] Hegel claims that criminals upset the relation of right in a law-governed state. By punishing them, the state negates, annuls, or cancels (*aufzuheben*) the wrong (or negates the negation of right), thus reestablishing a just social arrangement (1991, §§99–100 [pp. 124–27]).[5] For both Kant and Hegel, giving criminals what they deserve acknowledges that they are responsible for their actions and, by holding them accountable, treats them with a kind of respect. Kant says that basing punishment decisions on what will be produced as a result demeans offenders by treating them "merely as means," or as instruments of social engineering (1996b, 6:331–32). And Hegel says that people have a right to be punished because they have a right to be treated as responsible agents (1991, §§99–100 [pp. 124–27]).

One way of understanding the differences between consequentialist and retributivist theories of punishment is to adapt the classic distinction between consequentialist and deontological (duty-based) theories of ethics. Actions can be right or wrong, but experiences, states of affairs, and people can be good or bad. Ethical theories are distinguished by how they define and relate the right and the good. Value-based theories, including consequentialism, define the right in terms of the good. According to utilitarianism, for example, happiness is good, and the right thing to do is maximize happiness. Similarly, the utilitarian theory of punishment says that happiness is good, and the right thing to do is maximize happiness by preventing

4 As is customary in Kant scholarship, references to Kant's writings throughout the book give the volume and page number(s) of the Royal Prussian Academy edition (*Kants gesammelte Schriften*), which are included in the margins of the translations.

5 Neither Kant's nor Hegel's theory of punishment is as clear as we are often led to believe. Kant was a proponent, in various parts of his writing, of retributivism (1996b, 6:331–35) and consequentialism (1997, 27:286–88). And Hegel writes:

> Punishment … has a variety of determinations: that it is retribution; and also a deterrent example, a deterring threat made by the law; and also a contribution to the self-awareness and betterment of the culprit. Each of these different determinations has been regarded as the ground of punishment, on the ground that it is the essential determination, and by default the others, since they are different from it, have been regarded as only accidental. But the one determination which is assumed as ground does not amount to the whole punishment; the latter, as something concrete, also contains all of the rest, and in it these are only linked to the first without having their ground in it. (2010, 405–6)

Although Kant and Hegel are usually considered retributivists – proponents of desert theory and annulment theory, respectively – it is a matter of scholarly debate how to interpret their positions.

crimes through general and specific deterrence, restraint, and rehabilitation. Duty-based moral theories, by contrast, define the rightness of actions independently of the good. According to Kant's ethics, we ought to do our duty regardless of the consequences; it is valuable in itself. Happiness is defined separately from duty and is only conditionally good, depending on whether someone also has a good will (Kant 1996a, esp. 4:393–401). Similarly, the retributivist theory of punishment says that giving guilty people what they deserve is intrinsically good, regardless of the consequences. Punishment is justified because of what offenders did, without considering whether it deters potential criminals in the future.[6]

There are other theories of punishment, including expressivism, moral education theory, societal safety-valve theory, restitutive theory, and self-defense theory, to name a few. Often, however, these theories are more about specific methods to use rather than different kinds of justification, distinct from retributivism and consequentialism. According to expressivism, for example, punishment serves the symbolic function of expressing the community's attitudes of resentment and indignation at being wronged, and to signify a judgment of disapproval. On one interpretation, this is primarily forward-looking: it signals to potential and actual criminals that the behavior is wrong and will not be tolerated, so it has a deterrent effect on future crimes. On another interpretation, the hard treatment that criminals get in prison is itself symbolic because it is a kind of secular penance that they owe to the state. This is right in itself, regardless of whether criminals are reformed as a result.

6 The distinction between retributivism and consequentialism has been challenged by Michael Davis (2009), who claims that even traditional retributivist positions such as Kant's and Hegel's are consequentialist in some senses of the term. For one thing, they rely on arguments in which a conclusion follows from the premises, making them consequentialist in a logical sense. Retributivists also claim that punishment should be inflicted in order to achieve some aim, such as justice, as a consequence of the punishment. Furthermore, Davis says, it is misleading to distinguish the two theories based on whether they are backward-looking or forward-looking. After all, consequentialists must believe there is some relationship between the crime and the punishment (lest they risk unjust punishment), and retributivists believe that the institution of punishment is there to ensure justice in the future, such as restoring a relation of equality between offender and victim, annulling the crime, or negating the offender's advantage. As an alternative to the consequentialist/retributivist distinction, Davis suggests that we distinguish theories based on whether the rightness of a punishment depends on facts in the world (e.g., whether punishment has a deterrent effect) or on some conceptual relationship between crime and punishment (e.g., that a wrongdoer deserves hard treatment). The former are what he calls empirical or externalist theories, and the latter are conceptual or internalist theories.

Davis is right that theories of punishment are more varied and complicated than the traditional distinction between consequentialism and retributivism seems to allow. However, I wonder whether his clarification is more illuminating. In the logical sense of "consequentialist," nearly every philosophical position is consequentialist. And defining the "aim" of retributivist punishment as justice in the future obscures what defines justice for the retributivist: giving people what they deserve based on past actions. In short, although Davis is correct, his distinction does not seem more useful than the more common distinction between consequentialism and retributivism, especially since the latter distinction also tracks a common distinction in ethical theory.

Like John Stuart Mill's utilitarianism and Kant's deontological theory in ethics, we keep returning to consequentialism and retributivism in our moral, legal, and political discourse on punishment. For example, debates about the death penalty usually focus on a cost-benefit analysis about deterrence and whether the money could be better spent on other crime-prevention efforts; whether death is the only appropriate, deserved punishment for murderers; and whether it is worth it given the inevitable execution of some innocents, who do not deserve it. According to the U.S. Supreme Court, "the death penalty is said to serve two principal social purposes: retribution and deterrence of capital crimes by prospective offenders" (*Gregg v. Georgia*, 428 U.S. 153 [1976], 183; see also *Enmund v. Florida*, 458 U.S. 782 [1982], 798; *Atkins v. Virginia*, 536 U.S. 304 [2002], 318–19; *Roper v. Simmons*, 543 U.S. 551 [2005], 571; and *Kennedy v. Louisiana*, 554 U.S. 407 [2008], 441).[7] It is unclear, however, whether these two contrary positions can be maintained within the same system, given that they often have different practical implications. If the death penalty does not deter but murderers deserve to die, where should we put our limited resources? Without any rule-governed way to apply these (sometimes) competing principles, they provide little guidance for policymakers. Indeed, adopting both principles encourages arbitrary decision-making because public officials can pick and choose which theory to use depending on which position they want to take, resulting in a kind of moral subjectivism. It seems that we need to prioritize one of the two theories so that we know what is justified in cases where consequentialism and retributivism come into conflict.

The contest between these two theories defined the field until Paul Johann Anselm Feuerbach (1775–1833) introduced an important distinction. Feuerbach (1801, 2014) distinguishes the function of the legislature from the function of the criminal judiciary. Legislators write penal laws to discourage potential criminals and preserve the public order (on his view, to prevent rights violations), and judges apply the law as written in condemning and punishing the guilty.

7 Dolinko notes that retributivism has been in the ascendency in U.S. appeals courts since the mid-twentieth century, while utilitarianism is on the wane (1991, 537). With regard to the death penalty at least, U.S. courts want to hold both theories. As for rehabilitation, the U.S. Supreme Court has both flatly rejected it as a reason to extend an offender's prison term (*Tapia v. United States*, 564 U.S. 319 [2011]) and endorsed is as a goal of juvenile imprisonment (*Graham v. Florida*, 560 U.S. 48 [2010], 74, 79). The Court also has permitted civil confinement solely for the purpose of incapacitation, when, as is the case with some sex offenders, they are likely to reoffend and not be deterred (*Kansas v. Hendricks*, 521 U.S. 346 [1997], 372–74). Like the U.S., most countries do not explicitly commit themselves to one theory of punishment or another and, when punishment rationales are given, they are often inconsistent. According to Bagaric, Australia and the United Kingdom have both utilitarian and retributive objectives (1999, 543–44). Fox and Freiberg claim that governments have appealed to a number of different principles to justify criminal sentencing, often at the same time: deterrence, rehabilitation, denunciation, incapacitation, education, and community protection (1985, 444). For his part, Duff believes that trying to classify existing systems according to philosophical theories is bound to fail: "It would clearly be absurd to try to explain any existing penal system, whose historical development reflects an unsystematic diversity of competing influences, in terms of some unitary set of coherent values and purposes" (1986, 5).

Thus, punishment serves a social purpose insofar as it is a threat, and the threat is carried out by judges. This is a mixed theory that justifies legislative action on consequentialist grounds and judicial action on retributivist grounds, specifically legalistic retributivism (i.e., criminals deserve to be punished because they break the law).

This distinction was refined by H. L. A. Hart (1907–1992) in the twentieth century. When we try to justify punishment, we could be talking about different things: we could be talking about why we should adopt an institution of punishment, or we could be talking about why and how much we should punish particular criminals. Hart posed these as three separate questions: "What justifies the general principle of punishment? To whom may punishment be applied? How severely may we punish?" The first question concerns punishment's "*General Justifying Aim*," while the second and third questions concern its "*Distribution*," both its "*Title*" and its "*Amount*" (2008, 3–4). So long as these different topics are distinguished, Hart says, one does not have to choose either consequentialism or retributivism. They could function together but at different levels, with regard to either punishment policies or judgments of specific offenders.

Early proponents of mixed theories included such philosophers as Anthony Quinton, Stanley Benn, and John Rawls. In the 1950s, all of them proposed utilitarian justifications of the practice of punishment while also claiming that we should punish individual criminals because they deserve it. However, these mixed approaches have different conceptions of how the general justifying aim of punishment is related to its distribution. Quinton (1954) and Benn (1958) both argue that the answer to the question of whom specifically to punish follows logically from the meaning of the concept "punishment." Retributivism is thus a logical, conceptual, or definitional claim rather than a normative philosophy: only the guilty are able to be punished, lest we contradict the meaning of the term. According to Rawls (1955), on one conception of rules (as practices), we can think of legislators as establishing the laws and assessing their consequences, since their purpose is to protect the community. Judges then follow the rules by applying the laws in particular cases and giving people what they deserve. If they do not give people what they deserve – by, say, purposely imprisoning innocents – then they are not acting as judges anymore and they are not punishing people, since they would be violating the practice as defined by the rule.

Because these mixed theories incorporate both utilitarianism and retributivism, inevitably there is the possibility that they will have conflicting implications. If utilitarianism is overriding – we should punish individual offenders if and only if it maximizes utility – then it would seem to constrain retributivism to the point where the theory collapses into utilitarianism, which is not a mixed theory. For example, if a murderer escaped to an isolated desert island, and they posed no additional threat and punishing them would not deter other murderers, then they should not be punished, even though they deserve to be. And if utilitarianism is not overriding – we should punish the guilty regardless of the consequences – then non-utilitarian considerations could also shape the institution of punishment, in which case it becomes a disguised form of retributivism. For example,

the death penalty would remain in place, as a deserved punishment, even if it had no deterrent effect and siphoned funding from more effective methods of crime control. To avoid both horns of this dilemma, some mixed theorists try to maintain a strict distinction between the general justifying aim of punishment and its distribution. But that can be difficult. For example, what should we do if utilitarianism dictates that we punish a class of criminals in a way that is disproportionate (either too much or too little) to the guilt of individual criminals who fall under that class? If neither theory is overriding, we are at an impasse.

As a theory of punishment, consequentialism became less popular in the 1970s mostly as a result of sustained critiques and perceived failures of the social sciences.[8] Prior to that time, legislators, criminologists, and the public generally believed that imprisonment and the threat of imprisonment were supposed to deter potential criminals and rehabilitate convicted criminals. However, if we cannot tell what kinds or degrees of punishments affect the crime rate, and if we are unable to measure the impact of rehabilitative programs on recidivism, then consequentialism gives us no guidance about which policies and practices to adopt. There are so many psychological, social, economic, political, and legal factors that affect the crime rate that we cannot isolate punishment, let alone its kind or degree, as one cause among many. And without being able to measure the effects of punishment, we cannot do a utilitarian calculus to determine the best of the available options. If we can be morally obligated to do only what we are able to

8 Histories of the rise and fall of rehabilitation as an ideal, and of consequentialism more generally, can be found in Tonry and Morris (1978), Bottoms and Preston (1980, ch. 1–3), Allen (1981), Galligan (1981), Radzinowicz and Hood (1981), von Hirsch (1985, 3–18), Duff and Garland (1994, 8–16), Ryberg (2004, 3–5), M. Davis (2009, 84–85), and V. Bailey (2019). In addition to concerns about the predictive power of the social sciences, political opposition to rehabilitation came from both the left and the right. People on the left objected to the unfairness of indeterminate sentencing, since judges and parole boards tended disproportionately to punish Black offenders. People on the right objected that rehabilitation programs were too lenient and that victims deserve justice. See Bushway and Paternoster (2009, 121–23). The general population supported more punitive policies such as mandatory minimums and imprisonment for drug offenders because of their fear of crime and anger at criminals. However, philosophers may have different reasons for their shift toward retributivism, as Cederblom points out:

> The resurgence of the idea of just deserts among the general public might be interpreted simply as a reaction to the increase in crime. To a public which feels threatened by and angry toward lawbreakers, the idea of "giving them what they deserve" is undoubtedly satisfying. But among scholars, the renewed interest in retributivism is surely not the result of anger over crime. One suspects rather that it results from a confluence of factors such as the growing documentation of the failure and abuse of rehabilitation programs, the general decline in respectability of utilitarianism (the chief rival of retributivism among philosophic theories of punishment), and the rise of contractarian and natural rights theories which might serve as an underpinning for a retributivist approach to punishment. (1977, 2)

Since this is a philosophy book and not a history book – about justifications and reasons rather than causes of events – I will relegate these historical musings to a footnote.

do ("ought implies can"), then our inability to calculate the consequences means that we should not be consequentialists when it comes to punishment.

The apparent failure of utilitarian theories, and indeed of all forward-looking deterrence theories, paved the way for the ascent of retributivism, which, in different forms, remains the dominant theory of punishment. Although most philosophers are willing to grant that punishment has some impact on reducing crime and thus that consequentialism is true in this general sense, we should decide which actions to punish and how much based on what criminals deserve. On this view, the severity of the punishment should match the seriousness of the offense.

The two-tiered model of punishment

In this book, I defend a mixed theory that I call the two-tiered model of punishment. My view is not entirely original. It has its roots in the views of Feuerbach, Hart, and Rawls. But there are significant differences, in part because of developments in philosophy and criminology in recent decades, that make the two-tiered model stronger as a theoretical alternative to either utilitarianism or retributivism considered on its own.

Although the position I defend is a mixed theory of punishment, I usually refer to it as the two-tiered *model* rather than the two-tiered *theory* because two theories – consequentialism and retributivism – are plugged into it as a kind of organizing structure. On my view, the institution of punishment as well as designated offenses and statutory penalties, as set by the legislature, are justified based on their costs and benefits, in terms of deterrence and rehabilitation. The law exists to preserve the public order. Criminal courts, by contrast, use retributivist claims to determine who is punished and how much, based on what offenders deserve. David Wood calls this a form of "split-level dualism" (2002, 304–6). Distinguishing the two roles makes sure that criminals are not punished beyond what advances the general welfare (because individual judgments are constrained by legislation), and it precludes punishing the innocent because judges and juries consider only the guilt of the accused and not the consequences of a given judgment (so that they do not consider, for example, whether framing an innocent defendant would maximize the good).

The two-tiered model of punishment avoids the two biggest criticisms of mixed theories. First, it is not a disguised form of either consequentialism or retributivism. Unlike the mixed theories of the 1950s, the two-tiered model does not merely make a conceptual distinction between practices and specific decisions within a practice. Instead, both consequentialism and retributivism are independently justified, meaning that one cannot simply be absorbed into the other without losing something of moral significance. Although the consequentialist concerns of the legislature constrain the retributivist aims of the criminal courts, that constraint is not absolute. When a prescribed range of punishments is unreasonably excessive or inadequate, there is a deliberative process to determine what is appropriate. A second reason why a mixed theory is now more defensible

than it was seventy years ago is because there has been significant progress in the social sciences, such that they can more clearly and justifiably guide public policy. We can now make reliable (or reliable enough) generalizations regarding punishment to comparatively assess more and less effective punishments or alternatives to punishment, and more and less effective rehabilitation methods. Therefore, we can choose among the available options on consequentialist grounds.

Outline of parts and chapters

The book is divided into five parts. In Part I, I define what punishment is (Chapter 1). Although I largely follow the standard definition set out by Antony Flew, Benn, and Hart, I do not focus exclusively on legal punishment. Seeing how legal punishment is different from other kinds of punishment will help to illuminate its distinctive function in advancing the aims of the state. Because there seems to be no way to define punishment without begging the question in favor of one theory or another, however, I propose a few different stipulative definitions that will help to guide the evaluative discussions in the rest of the book.

In Part II, I defend consequentialism and retributivism. Punishment helps to maintain the social order, and the rightness of a given set of punitive policies depends on how well it accomplishes its aim and whether the social order itself is justified (Chapters 2–3). Yet there is also value in the state's expressing the community's resentment at being wronged (Chapter 4). Drawing on Hart's distinction between the general justifying aim of punishment and its distribution, I propose the two-tiered model as a way that we can simultaneously be committed to consequentialism and retributivism in, respectively, the legislature and the criminal judiciary (Chapter 5).

I give three arguments in Part III for why the two-tiered model is the right way of organizing these theories. Legislators and criminal court judges are in different positions, epistemically, with regard to the social effects of punishment in general and the guilt or innocence of particular defendants. Their different frames of reference entail that the legislature is more bound by consequentialist concerns and the courts are more bound by retributivist concerns (Chapter 6). Following that, I argue that the two-tiered model can accommodate a compatibilist theory of freedom. Under compatibilism, people can be held responsible, even if determinism is true, so long as they are determined in the right way. The legislature and the judiciary take two different perspectives on a person's action, as determined or free (Chapter 7). Finally, I show how the implications of the two-tiered model accord with our moral intuitions regarding, for example, the punishment of innocents and the prohibition on cruel punishments (Chapter 8). One shortcoming of previous mixed theories is that they often sound like mere proposals or suggestions, with little argument about why the system *should* be organized in this way. I correct that oversight in Part III.

Having established the reasons behind consequentialism and retributivism, and the reasons why they should be combined in a two-tiered model, in Part IV I explain how to make punishment decisions at both the policy level and the

level of individual judgments. I begin by explaining how consequentialist reasoning regarding deterrence can be validated by the empirical social sciences (Chapter 9). We need not adopt a retributivist system, which faces problems of its own, in order to devise a punishment schedule (Chapter 10). A recurring challenge to mixed theories is that combining two opposing positions leaves us with the choice of either prioritizing one over the other or following one or the other arbitrarily when they conflict. I explain the two-tiered model's decision procedure in Chapter 11: conflicts between the legislature and the courts are solved not by subsuming one to the other, but through a public process of deliberation by which the community achieves reflective equilibrium.

In Part V of the book, I demonstrate the value of the two-tiered model by applying it to two controversial topics regarding punishment. First, I make the case that the death penalty should be abolished because of its lack of deterrent effect. Assessing its effectiveness is crucial for formulating public policy (Chapter 12). Second, I show how restorative justice practices can be used instead of or in addition to traditional punitive measures in ways that reinforce the law's retributive function (Chapter 13). These two examples illustrate how the two-tiered model can be action-guiding as we consider penological reform.

Conclusion

With legal punishment, the state intentionally and systematically inflicts harm on its own people. It is incumbent on us to examine the practice, to evaluate its supposed justifications, and, if necessary, to reform our policies. Because philosophers are trained in analyzing concepts, investigating values, and assessing arguments, we are uniquely situated to perform this task. My (admittedly lofty) hope is that this book will not only add to the body of existing research on punishment, but that it will raise the quality of discussion about issues of pressing public importance.