

5-31-2014

Can Retributivism be Saved?

Chad Flanders

Follow this and additional works at: <http://digitalcommons.law.byu.edu/lawreview>

 Part of the [Law Enforcement and Corrections Commons](#), and the [Legal Theory Commons](#)

Recommended Citation

Chad Flanders, *Can Retributivism be Saved?*, 2014 BYU L. Rev. 309 (2014).
Available at: <http://digitalcommons.law.byu.edu/lawreview/vol2014/iss2/4>

This Article is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

Can Retributivism Be Saved?

*Chad Flanders**

Retributive theory has long held pride of place among theories of criminal punishment in both philosophy and in law. It has seemed, at various times, either much more intuitive, or rationally persuasive, or simply more normatively right than other theories. But retributive theory is limited, both in theory and practice, and in many of its versions is best conceived not as a theory of punishment in its own right, but instead as shorthand for a set of constraints on the exercise of punishment. Whether some version of retributive theory is a live possibility in the contemporary world remains very much an open question.

In my essay, I consider three interrelated lines of attack against retributive theories of punishment: first, that it relies on philosophical assumptions that are either unrealistic or false; second, that the notion that offenders deserve to be punished, whatever its intuitive appeal, is possibly an empty idea and in any case one unsuited to a politically liberal state; and third, that the abstractness of most versions of retribution render it unable to offer much in the way of useful or concrete policy advice. If retributivism is to be taken seriously as a robust theory of punishment, it needs to provide replies to each of these lines of criticism.

* Assistant Professor of Law, St. Louis University School of Law. Thanks to the faculty at DePaul Law School for comments on a draft I presented to them, especially Susan Bandes, David Franklin, Andrew Gold, Steven Resnicoff, Daniel Morales, and Allison Tirres. I have certainly not answered all of their concerns. I am also grateful to William Baude, Christopher Bradley, and Zach Hoskins for comments and conversations on earlier drafts. David Svolba helped me think through early versions of the ideas presented here and Danny Priel gave me numerous, detailed comments on a near-final draft. Vincent Chiao has influenced my thinking away from retribution and toward a broader understanding of the criminal law. Finally, a Chicago junior faculty workshop challenged me to make my argument clearer, and also much more modest in its conclusions. Alex Muntges and Yiqing Wang provided excellent research assistance. All errors are my responsibility.

I. INTRODUCTION	310
II. RETRIBUTIVISM: A CRITIQUE OF THE MAJOR THEORIES .	316
A. <i>Kant and Hegel</i>	316
1. <i>Respect</i>	317
2. <i>Proportionality</i>	322
B. <i>The Modern Retributivists: Morris, Hampton</i>	327
1. <i>Herbert Morris</i>	328
2. <i>Jean Hampton</i>	330
III. DOING WITHOUT DESERT	333
A. <i>Desert as Misunderstood Side Constraint</i>	335
B. <i>Does Desert Exist?</i>	338
C. <i>Desert in Comprehensive Doctrines</i>	343
IV. RETRIBUTIVE THEORY IN PRACTICE	347
A. <i>Retribution as Limited in Practice</i>	348
1. <i>Cost containment</i>	348
2. <i>Mass incarceration</i>	350
3. <i>Mandatory minimums</i>	351
4. <i>The death penalty</i>	352
B. <i>Why Retribution Fails, Practically Speaking</i>	354
1. <i>Closure</i>	354
2. <i>Compartmentalization</i>	356
C. <i>Is Retribution a Theory of Punishment at all?</i>	358
V. CONCLUSION: REVENGE, FAIRNESS, AND DESERT	360

I. INTRODUCTION

Retribution is often taught as one of the major philosophies of punishment, along with (prominently) deterrence and (less prominently) rehabilitation. No major casebook omits at least a cursory discussion of Immanuel Kant on retribution, or James Fitzjames Stephen on revenge.¹ Retribution, no mere historical

1. One could cite almost any introductory criminal law casebook here, but Joshua Dressler's seems exemplary in the emphasis he gives to the philosophy of punishment, including retribution. See JOSHUA DRESSLER & STEPHEN P. GARVEY, CASES AND MATERIALS ON CRIMINAL LAW (6th ed. 2012).

artifact, has also enjoyed a tremendous theoretical resurgence in the twentieth and twenty-first centuries as more and more philosophers and lawyers spend time articulating, defending, and elaborating versions of retributivism. Several books published recently testify to retributivism's enduring prominence in the halls of the academy.² The so-called retributive revival in theory is here to stay, or so it seems.³

There is also good reason to believe that retributivism, in perhaps a cruder and much less sophisticated form, has also been influential at the practical level. The demand for uniformity in sentencing reflects the idea that sentences should be roughly the same for all offenders, which reflects the retributive ideas of proportionality and just deserts.⁴ More generally, the sentiments that wrongdoers should be blamed or that wrongdoers deserve punishment are surely retributive in nature.⁵ We believe that when a crime has been committed and prosecuted, punishment should be imposed irrespective of whether this fully maximizes deterrence or rehabilitates the offender.

But despite this seemingly overwhelming support, retributivism faces serious challenges both in theory and practice. In theory, retributivists have ignored or skated over what seems to be a major (and almost decisive) flaw in their theory: they cannot give any

2. In just the past two years, two major collections of essays on retribution have appeared. See *RETRIBUTIVISM: ESSAYS ON THEORY AND POLICY* (Mark D. White ed., 2011); *RETRIBUTIVISM HAS A PAST, DOES IT HAVE A FUTURE?* (Michael Tonry ed., 2011) [hereinafter Tonry, *RETRIBUTIVISM HAS A PAST*]. White's book is mostly essays in defense of retribution; Tonry's is more balanced. In addition, Michael Tonry has recently edited an anthology of writings in philosophy of punishment, in which significant space is given to retributive theories. See *WHY PUNISH? HOW MUCH?* 111–95 (Michael Tonry ed., 2011) [hereinafter Tonry, *WHY PUNISH?*] (section on retributive theories of punishment).

3. For citations on the “retributive revival,” see Chad Flanders, *Retribution and Reform*, 70 MD. L. REV. 87, 87 & n.1 (2010).

4. For a good overview of mandatory minimums in sentences, see Erik Luna & Paul Cassell, *Mandatory Minimalism*, 32 CARDOZO L. REV. 1 (2010); see also *id.* at 11 (connecting mandatory minimums to retributive justifications for punishment). The classic defense of uniformity in sentencing is MARVIN E. FRANKEL, *CRIMINAL SENTENCES: LAW WITHOUT ORDER* (1973). See generally ANDREW VON HIRSH, *DOING JUSTICE* (1976) (defending the strict retributive theory of punishment). For an excellent overview of “equality” in sentencing, see James Q. Whitman, *Equality in Criminal Law: The Two Divergent Western Roads*, 1 J. LEGAL ANALYSIS 119 (2009).

5. See James Q. Whitman, *A Plea Against Retributivism*, 7 BUFF. CRIM. L. REV. 85, 85 (2003). Jeffrie Murphy makes a more general defense of the so-called “retributive emotions” (blame, hatred, resentment) in JEFFRIE G. MURPHY & JEAN HAMPTON, *FORGIVENESS AND MERCY* (1990).

meaningful account of proportionality in punishment.⁶ This has been recognized before by scholars, but no retributivist has so far given an entirely satisfactory answer to the basic question: how do we tell exactly how much punishment offenders deserve? The fact that retributivism places so much emphasis on proportionality (or fairness) in punishment, but cannot offer much illumination about what proportionality is—or how to achieve it—is a key insight that something has gone wrong with retributive theory, even on its own terms. Indeed, it suggests deeper problems about the very idea of “deserved” punishment. Retribution has too easily been allowed to “punt” on spelling out what proportionality in punishment means.⁷

Retributivism has also been increasingly seen as being limited practically. One major example is that by describing deserved punishment in terms of justice, scholars and practitioners have made deserved punishment seemingly resistant to fiscal considerations.⁸ But it turns out, unsurprisingly, that punishment is a serious financial drain on state and federal budgets, so prudent legislators and citizens are now saying that punishment needs to be looked at pragmatically and not as a matter of punishing offenders come what may. This is a serious, indeed monumental,⁹ shift in the politics of punishment on the ground, and unfortunately retributivists have little to say about it.¹⁰

6. A particularly good statement of this problem is in Hugo Adam Bedau, *A Retributive Theory of the Pardoning Power?*, 27 U. RICH. L. REV. 185, 198 (1993), who calls defining proportionality the “very heart of darkness” for the retributivist. For an extensive discussion of the problem, see MICHAEL J. ZIMMERMAN, *THE IMMORALITY OF PUNISHMENT* (2011). In his introduction to *RETRIBUTIVISM HAS A PAST, DOES IT HAVE A FUTURE?*, Michael Tonry notes the problem, but does not (to my mind) appreciate its depth. See Tonry, *RETRIBUTIVISM HAS A PAST*, *supra* note 2, at 1 (expressing frustration at recent retributivist thought to say anything meaningful about proportionality).

7. I have been guilty of punting in the past. See Dan Markel & Chad Flanders, *Bentham on Stilts: The Limited Relevance of Subjectivity to Retributive Justice*, 98 CALIF. L. REV. 907, 951–52 (2010) (rejecting the insistence that retribution answer questions of proportionality, and saying that such questions are institutional, not conceptual).

8. For a sensitive examination of this problem, but with (in my opinion) a less than satisfactory answer, see Michael T. Cahill, *Retributive Justice in the Real World*, 85 WASH. U. L. REV. 815 (2007).

9. Especially given that it has meant conservatives are now mobilizing against long (and expensive) sentences. For an excellent survey of this development, see David Dagan & Steven M. Teles, *The Conservative War on Prisons*, WASH. MONTHLY, Nov.–Dec. 2012, at 25, 25–31. For more on the role of cost in criminal justice more generally, see Chad Flanders, *Cost as a Sentencing Factor: Missouri’s Experiment*, 77 MO. L. REV. 391 (2012).

10. See my discussion *infra* Part III.

In my essay, I try to diagnose in what ways traditional retributivism has gotten things wrong and how to fix them. In truth, the solution may be to largely abandon retribution as a *theory* of punishment.¹¹ What retributivism gets right is that there should be some constraints on how and why punishment is implemented. We should not punish the innocent. We should not, as a rough matter, punish less severe crimes more harshly than how we punish more severe crimes. These two features of retribution—call them the *innocence constraint* and the *proportionality constraint*—are two things that retribution has gotten right. But retribution goes wrong when it translates these features into insights on the aims of punishment generally instead of limiting them to how punishment should be distributed. Retribution goes wrong when it endeavors in the first instance to explain *why people deserve* punishment. In fact, deterrence theory is largely right that people should be punished because it improves the safety of society, and for no other reason. If retributive theory is to find its proper place, it will be as part of an overall theory that takes deterrence as the ultimate justifying aim of punishment, not one that begins and ends with retributive considerations.

The consequences of the retributive mistake have been unfortunate both in theory and practice. By seeking out a theory of desert, retributive theory has spent time explaining how punishment might be a good thing, for both those being punished and for the rest of society, beyond the fact that punishment might increase society's safety. This idea, *that punishment is an intrinsic good*, is not only fruitless, it is harmful because it encourages the notion that those who are punished benefit somehow from their punishment, so that if we failed to punish them, we would be hurting them by violating their right to be punished.¹² This rather strange notion has its roots in Kant and Hegel, and makes little sense considered

11. But, as I go on to note, this does not mean abandoning certain norms usually associated with retribution. The problem comes in thinking that *only* if we adopt retribution as a theory can we endorse those norms.

12. Alice Ristroph has emphasized this point in her recent work. Alice Ristroph, *Respect and Resistance in Punishment Theory*, 97 CALIF. L. REV. 601 (2009). However, Ristroph seems to me to take her point too far, into a generalized skepticism about state authority, a skepticism which I reject. See Flanders, *supra* note 3, at 133–37 (contrasting Hobbes and Smith on the legitimacy of state punishment). On the “right to be punished,” see John Deigh, *On the Right to Be Punished: Some Doubts*, 94 ETHICS 191, 195–97 (1984).

outside of their metaphysics, yet it persists in modern philosophical writing about punishment.

Moreover, by focusing on the intrinsic benefit of punishment, retributivists have seemed to shift their focus away from the practical problems of crime and punishment: how to reduce crime by preventing it from happening, how to improve the lives of those in prison, how to ease the reentry of prisoners back into society. These are irrelevant to the question of what justifies punishment in the way retributivism often conceives of it. This focus on punishment may in some sense be excused as methodological: retributivists are interested just in justifying the act of punishment and nothing else. But punishment is just one part of a complex network of criminal justice, and the focus on one part of that network at the expense of others (which is what retributivists do) can result in a distorted picture of crime and punishment. There is some exciting empirical work being done on stopping crime and reducing punishment, which retributivism takes very little account of, if at all.¹³

Retributivism, I conclude, risks ringing false when it stops merely being a theory of some rational constraints on punishment and instead turns into a theory of why we have punishment at all: it makes what should be at the outer bounds of a theory of punishment into its core. For the fact is, we have punishment in order to make society safer by controlling crime, and not primarily to give people what they deserve.¹⁴ To be sure, not every kind of punishment should be tolerated in pursuit of this goal, and this is where retributivist constraints on punishment can be both necessary and helpful, by providing the outer framing of the pursuit of our criminal justice goals. But retributivism goes wrong when it presents itself as an independent good, something we should seek out for its own sake. Everything that is valuable about retribution as a full-

13. See, *inter alia*, DAVID M. KENNEDY, DON'T SHOOT! ONE MAN, A STREET FELLOWSHIP, AND THE END OF VIOLENCE IN INNER-CITY AMERICA (2011) (describing the author's experience with various crime-reducing programs); MARK A.R. KLEIMAN, WHEN BRUTE FORCE FAILS: HOW TO HAVE LESS CRIME AND LESS PUNISHMENT (2009) (providing empirical studies and theoretical analysis of what reduces crime).

14. Here I follow H.L.A. Hart in his analysis of the "justifying ground" of punishment. See H.L.A. Hart, *A Prolegomena to the Principles of Punishment*, in Tonry, WHY PUNISH?, *supra* note 3, at 195–205.

blown theory of punishment can be reformulated into retribution as a theory of side constraints on punishment.¹⁵

My essay divides into three major parts, each developing a part of the critique of retributivism. In the first part, I look at the substance of some of the major historical and contemporary theories of retribution: Kant and Hegel on the one hand (historical), and Herbert Morris and Jean Hampton on the other (contemporary). Kant and Hegel's theories either rely on barely comprehensible, and probably false, metaphysics, while Morris and Hampton's theories—while dispensing with the dubious metaphysics—only obliquely touch facts about real people. The problem with both types of retribution (the historical and the contemporary) turns out to be similar: retributivists take an important point about the *limits* of punishment—that it should not be imposed on the innocent and should not be grossly disproportionate to the crime—as an insight into the *purpose* of punishment. But Rawls and Hart, in remarkably similar essays published at roughly the same time,¹⁶ had it right: the *overall* purpose of punishment should be utilitarian, not retributive.

The second part of my essay attacks the intuition that grounds much sophisticated retributive thought: the idea that people deserve punishment for their crimes. What is this notion of desert? Using insights from John Rawls's work on distributive justice, I question whether we have any firm grasp on the idea of preinstitutional retributive desert. Indeed, the inability of retributivists to say anything meaningful about proportionality shows the truth of this insight. We have no grasp on what a “proper” punishment would be because the idea of determinate desert prior to institutions that can give desert a shape and meaning is incoherent. If we insist, contrary to Rawls, that we do have a concrete idea of desert prior to institutions, either this idea of desert cannot give us meaningful guidance when it comes to sentencing, or it relies on a set of metaphysical beliefs that should not ground a political conception of justice.

The final part focuses on the practical failings of retributive theory. When we look to retribution for solutions or even suggestions on what to do about criminal justice in our present

15. For the idea of “side constraints” on action, see, e.g., ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA (1974).

16. See *infra* note 79.

condition, we either find that retributivism is virtually silent, or it points us in the wrong direction. With cost containment, retribution says cost is irrelevant and we should ignore it. With mandatory minimums, retribution says we should work to make *all* sentences mandatory. On the death penalty, retribution gives a strong justification *for* it, despite recent (heroic, but misguided) efforts by retributivists to assert the contrary. Finally, with the phenomena of mass incarceration, retribution says that it is justified if the punishments were deserved, which is simply to deny that mass incarceration itself is a problem at all.

My conclusion touches on the old question of whether retribution is just a stand-in for revenge.¹⁷ I conclude that it is not, or at least not entirely, but rather that retribution and revenge share a very important characteristic, which makes them both inappropriate to be taken seriously as theories of punishment: they isolate themselves from consideration of consequences, and any purposes other than the “intrinsic” good of punishment itself—a good which, if it exists, is no good at all.

II. RETRIBUTIVISM: A CRITIQUE OF THE MAJOR THEORIES

A. Kant and Hegel

The hero and godfather of retributivists is surely Immanuel Kant.¹⁸ Although he is widely cited, or at least parts of him, Kant’s philosophy is little understood and perhaps for good reason. Kant’s major discussion of retribution in his *Metaphysics of Morals* includes some passages that are hard to understand outside of Kant’s sometimes extravagant and dubious metaphysics. Indeed, few accept all of what Kant says in these passages, though many think that there are still ideas worth salvaging in his work, if only we are careful. But it remains to be seen whether these “good” ideas can be separated from Kant’s larger metaphysical beliefs, and *even if they can be*, whether they are plausible ideas taken by themselves.

The two major ideas that modern retributivists take over from

17. For the classic rebuttal to this charge, see ROBERT NOZICK, PHILOSOPHICAL EXPLANATIONS 366–68 (1981).

18. In referring to Kant, I will be mostly using the excerpt found in Immanuel Kant, *The Penal Law and the Law of Pardon*, in Tonry, WHY PUNISH?, *supra* note 2, at 31–36 [hereinafter Kant, *Penal Law*].

Kant are these: the idea that punishment necessarily expresses *respect* for those who are punished, and second, that punishment should be *proportional* to the crime. The first idea, I think, is largely mistaken. The second idea is correct, but it turns out to be a mostly empty idea. Modern interpreters of Kant are hard pressed to give content to Kant's idea of proportionality once it is divorced from Kant's metaphysics.

1. *Respect*

Start with the idea that punishment expresses respect for those who commit crimes. According to Kant's conception of the social contract, anyone who wills a crime, also wills that he be punished—he has done the crime *to himself*.¹⁹ To fail to punish the person who commits the crime, then, is to fail to treat him as a willing subject in his own right, to fail to respect his choices. It is to treat him instead, perhaps, as someone who is to be managed or controlled and not as a person.²⁰ That, for Kant, is almost the paradigm case of disrespect, and any self-respecting person will accordingly *want* to be punished.²¹ Hegel, in this seemed to be substantially in accord with Kant. Indeed, he seemed to pick up Kant's idea and run with it. To not punish a person who had done wrong would be to dishonor him, Hegel said; punishing is a way of “honoring” the choosing subject.²²

This idea has some intuitive appeal. If we do not hold someone responsible for their actions, and instead say that is the result of social factors beyond their control, we may diminish them as persons. C.S. Lewis gave great rhetorical support to this idea in his

19. *Id.* at 32 (“Accordingly, any undeserved evil that you inflict on someone else among the people is one that you do to yourself.”).

20. *Id.* at 31–32 (stating that a person can only be punished “on the ground that he has committed a crime; for a human being can never be manipulated merely as a means to the purposes of someone else and can never be confused with the objects of the Law of things.”).

21. *See, e.g., id.* at 33 (expressing that a man of honor would choose to be punished according to his desert, even if this meant death). Hegel would later pick up and run with this idea. *See* G. W. F. Hegel, *Wrong [Das Unrecht]*, in Tonry, *supra* note 2, at 45–46 (“In so far as the punishment which [the act of the criminal] entails is seen as embodying *the criminal's own right*, the criminal is *honoured* as a rational being.—He is denied this honour if the concept and criterion of his punishment are not derived from his own act; and he is also denied it if he is regarded simply as a harmful animal which must be rendered harmless, or punishment with a view to deterring or reforming him.”).

22. *See* G.W.F. Hegel, *Wrong [Das Unrecht]*, in Tonry, *supra* note 2, at 45–46.

classic essay on the humanitarian theory of punishment: treating someone as not responsible for his actions because he couldn't really control himself can be insulting; it is as if we see him more as a child than as an adult human being.²³

But we need to question each part of Kant's analysis. Initially, is it true that a person who chooses to commit a crime chooses to be punished? As an empirical matter, this is certainly not true: those who commit crimes by and large do not want to be punished, they want to commit a crime and they want to get away with it. They do not want to suffer the consequences of their criminal act. They want to enjoy the benefit of the laws, without having to pay the price when they break those laws. Kant thought that no one would object to being sentenced to death for murder, and that people would laugh at anyone condemned to death said he was being punished too harshly;²⁴ in that, we can see, he was clearly wrong.²⁵

Of course, Kant was not simply making a descriptive claim. Rather, his point was that as a transcendent matter when people choose to commit a crime they are in fact choosing to be punished. And so, even if they do not actually want to be punished, *they are committed transcendentally to willing their punishment.*²⁶ So we respect people when we punish them, even if their empirical wills are suggesting (pleading?) that they would rather not want this form of "respect." We respect criminals by taking their actions seriously, as acts of responsible agents and not of children or animals or robots who could not have known better or who could not have done otherwise. Kant's version of this argument requires a lot of suppositions about the original social contract, but we can put those to one side.²⁷ Most contemporary defenders of Kant do not appeal,

23. See C.S. Lewis, *The Humanitarian Theory of Punishment*, in Tonry, *supra* note 2, at 91.

24. Kant, *Penal Law*, *supra* note 18, at 34.

25. Corey Brettschneider, *The Rights of the Guilty: Punishment and Political Legitimacy*, 35 POL. THEORY 175-99 (2007) (arguing that, on a contractualist theory, people would not consent to the death penalty).

26. *Id.* at 35 ("When, therefore, I enact a penal law against myself as a criminal it is the pure juridical legislative reason (*homo noumenon*) in me that submits itself to the penal law as a person capable of committing a crime, that is, as another person (*homo phaenomenon*) along with all the others in the civil union who submit themselves to this law.")

27. *Id.* Kant imagined that we "transcendentally" agreed to a contract where, if we were to commit a crime, we consent to being punished. See Immanuel Kant, *Theory and Practice Concerning the Common Saying: This May Be True in Theory but Does Not Apply in Practice*, in

explicitly anyway, to them: they reject the idea that we are literally honoring people's choices by punishing them.²⁸ Instead, they focus on the idea that punishment is a way of showing respect for people: by taking their choices seriously, as choices of responsible agents.

I am suspicious of this form of "respect." For one thing, it does not seem true that we respect people *only* by punishing them. For example, we could try to rehabilitate or even forgive them.²⁹ Both of these methods of responding to a crime do not seem to diminish our respect for that person *qua* person, or treat them merely as products of social forces. Indeed, both rely on treating that person *as a person* either by encouraging him to see the wrongness of his ways (forgiveness) or helping him change into a better person (rehabilitation). It seems cruel and wrong to think that the only way (or the best way) of relating to a person who has done wrong is punishing him. Even the idea that therapy shows lack of respect for a person seems hopelessly outdated: the remnants of a culture that was fearful of psychological explanations of human behavior, fearful that it would simply show us as a product of dark, unconscious drives.³⁰ Therapy now seems positively humanistic in contrast to the now-dominant biological (and reductive) understandings of human nature.

In short, there are many ways to treat a person as a responsible agent, and only a narrow set of them involves actually punishing the person, by subjecting them to either physical or psychological suffering. Think of the different ways we hold others responsible for doing something wrong in our daily lives. We hold others responsible by criticizing them, shaming them, expressing disappointment in them, rebuking them, shunning them, badmouthing them, yelling at them. All of these acts presuppose that a person has done something wrong, and that this wrongdoing requires censure. But they all stop short—some very far short—of punishment. Even the plausible, lesser claim by Kant, which is not that we *will* our punishment but that we ought to be taken seriously

THE PHILOSOPHY OF KANT 421–22 (Carl J. Friedrich trans., 1949).

28. Most contemporary Kantians reject the idea of a transcendental contract, but hold on to the idea of a "hypothetical contract." The question still remains, however, whether such a contract is binding. On this latter question, see Dan Brudney, *Hypothetical Consent and Moral Force*, 10 LAW & PHIL. 235 (1991).

29. See Deigh, *supra* note 12.

30. See, e.g., C.S. Lewis, *The Humanitarian Theory of Punishment*, in Tonry, *supra* note 2, at 91 (expressing distress about treating crime as a "disease" and needing to be cured).

as responsible agents when we do wrong, does not give us an argument for punishing people.

But there is in fact another flaw in the idea that punishment is the only way to show respect, one relating to the moment the person chooses to disobey the law. Kantian respect-as-punishment seems to rely on the idea that those who commit crimes are fully autonomous when they choose their crime, and they express their full, true selves in choosing. This also seems to be empirically incorrect, and not many would hold to the idea that those who grow up in rotten social circumstances have full autonomy, and fully choose their crimes. Kant of course felt otherwise: his view of the will was that *no matter our empirical circumstances* we can rise above them, and choose the right path.³¹ This is our transcendental or noumenal freedom, which we always retain, even in the most trying of circumstances. Social psychology has shown *this* type of freedom—freedom from any and all circumstances—to be simply nonexistent. A lot depends on group dynamics, or how we frame the situation, not on our transcendental freedom.³²

Now, consider: does it diminish our respect for a person if we show the victim has had a history of abuse in his family, or that he has brain damage, or that he has been addicted to drugs? It would diminish *respect* for him only if we saw a fundamental incompatibility between sympathy and respect such that acknowledging difficult social circumstances or mental or psychological difficulties meant no longer relating to a person *as* a person.³³ But there seems no necessary conflict between these two points of view. At the very least,

31. Kant held that it was always *possible* that we act in accord with the dictates of morality. See, e.g., KANT, CRITIQUE OF PRACTICAL REASON 5:30 (Mary McGregor trans., 1996) (“Ask [someone] whether, if his prince demanded, on pain of . . . immediate execution, that he give false testimony against an honourable man who the prince would like to destroy under a plausible pretext, he would consider it possible to overcome his love of life. . . . He would perhaps not venture to assert whether he would do it or not, *but he must admit without hesitation that it would be possible for him.* He judges, therefore, that he can do something because he is aware that he ought to do it and recognizes freedom within him, which, without the moral law, would have remained unknown to him.”) (emphasis added).

32. This has been the upshot of much work in behavioral psychology. See generally JOHN M. DORIS, LACK OF CHARACTER: PERSONALITY AND MORAL BEHAVIOR (2005); LEE ROSS & RICHARD E. NISBETT, THE PERSON AND THE SITUATION: PERSPECTIVES OF SOCIAL PSYCHOLOGY (2d ed, 2011)

33. For a sensitive reflection on the relationship between knowing the details of a person’s past and still respecting him or her as a person, see Martha Nussbaum, *Equity and Mercy*, 22 PHIL. & PUB. AFF. 83 (1993).

we might acknowledge that not all crimes are the result of fully aware, fully informed, fully controlled choice, and that this should make a difference as to whether punishment shows respect for a person's choices.³⁴

A large amount of the rhetorical strength of the Kantian view involves contrasting it with a deterrence based view: if we punish only to scare people, or to threaten them with terrible consequences if they commit a crime, then this treats people as means rather than as ends. We treat them as dogs, as Hegel famously said.³⁵ But in fact, this strikes me as eminently debatable. Certain threats require treating people as autonomous, choosing agents, who understand the consequences of their actions and can be expected to alter their behavior accordingly. Nor does it seem incompatible with respect to punish someone if this is necessary for the safety of the community, or to set an example for others. It would be wrong to punish someone only for these goals and not because he had committed an offense. But this is a side constraint on deterrence, not a replacement of deterrence with some other (retributive) goal.³⁶

A final point on respect: whatever the merits of the idea of transcendental respect by punishing someone who has committed a crime, it may still be the case that the *actual conditions* of American prisons and jails mean that those who are incarcerated there are not being treated with respect.³⁷ Respect suggests some constraints on how we treat people, and not merely that we treat them as responsible agents. Unless it is *any* punishment that will show respect (and I am doubtful that modern Kantians believe this; certainly Kant thought there were limits to how harsh and disgusting punishments can be³⁸), there may be some punishments that fail to

34. It seems to me a fairly common feature of modern life that we ascribe some things in a person's behavior to chemical or psychological factors without at the same time absolving them from all responsibility *or* thinking that they are no longer "persons." Note that the question of whether people with bad social backgrounds or psychological disorders should be *punished* is a separate question than whether punishing them is always necessary in order to *respect* a person, when he or she has committed a crime.

35. See ROBERT R. WILLIAMS, *HEGEL'S ETHICS OF RECOGNITION* 168 (1997) (using deterrence to justify punishment is like threatening a dog with caning).

36. For an excellent development of this point, see Zach Hoskins, *Deterrent Punishment and Respect for Persons*, 8 OHIO. ST. J. CRIM. L. 396 (2011).

37. See, e.g., *Brown v. Plata*, 131 S. Ct. 1910 (2011) (illustrating conditions in California prisons).

38. See the discussion of Kant on rape, *infra* at note 18.

show respect to prisoners, because they are cruel, or demeaning, or inhumane. Such may be the reality of many prisons and jails in America.³⁹

Of course, the Kantian may respond that this is an argument to reform prisons and jails; indeed, I have suggested this argument in the past, in a retributive spirit.⁴⁰ But this supposes that such an argument can be made from within Kantian philosophy; that Kant can give us a sound argument—one which modern Kantians can adopt and adapt—that some conditions are disproportionate to the crime that has been committed, so that we can say some punishments are in fact disproportionate. This is an argument at bottom about proportionality, which I turn to next.

2. *Proportionality*

Kant's argument that punishment is a gesture of respect is, I think, ultimately wrongheaded and should be rejected. It is not a bar to treating people with respect to have an institution of punishment that works primarily to reform them, or to deter others from committing similar crimes (or even to have no institution of punishment at all). But Kant's claim about respect was linked to another one, which some have also found persuasive. Kant says that those who commit a crime will their punishment upon themselves, but this begs the question: how much should they be punished? And Kant's answer is *exactly to the extent that they broke the law*. For Kant, this was the old doctrine of *jus talionis*, or an eye for an eye and a tooth for a tooth.⁴¹

Kant again reinterpreted the doctrine, rather brilliantly, in terms of the criminal as willing his punishment. A person who commits a crime not only wills that he be punished, but also that his punishment be the very crime he committed. He has agreed to this, again, transcendently speaking. Kant's best example is the death

39. This is yet another way in which Kant's theory may be better considered a theory for the ideal world, and not for our real world.

40. Flanders, *Retribution and Reform*, *supra* note 3, at 104.

41. Kant, *Penal Law*, *supra* note 18, at 32 ("What kind and what degree of punishment does public legal justice adopt as its principle and standard? None other than the principle of equality . . . that is, the principle of not treating one side more favorably than the other. Accordingly, any undeserved evil that you inflict on someone else among the people is one you do to yourself.").

penalty. A person who kills another deserves nothing less than to die, because he has chosen it.⁴² There is no possible substitute for death as his punishment. For other crimes, however, it is harder to see how this principle of reciprocity should work. Kant said that the thief makes another's property insecure, and so *he* should have his property made insecure.⁴³ But how do we do this? Do we do it by stealing from the thief? Or do we simply demand restitution? Kant, somewhat convolutedly, suggested that the person who stole may be used by the state for "convict labor", "so he becomes a slave, either for a certain . . . time or indefinitely, as the case may be."⁴⁴ But it is not obvious that slave labor is anything like a commensurate punishment for theft.

Rape presented an even harder case for Kant, who said that to rape the rapist would be undignified.⁴⁵ He suggested castration instead. There are two problems with this response, however. The first is that the idea that some punishments might be undignified is an external constraint on *lex talionis*, not something that is contained within it: the true retributivist, one might surmise, would grit his teeth and rape the rapist, if that is what the just punishment was in that case, given the requirements of proportional punishment. Moreover, if Kant were consistent, he might use the *same* argument against the death penalty, for killing somebody might be thought to treat someone in an undignified way. But Kant does not make this move, because for him, an execution is not the same as an unjustified killing. One wonders, then, why Kant could not say the same thing about a (judicially authorized, morally required) rape against a rapist.⁴⁶

The second point against the rape example is that castration does not seem *as proportionate* as the act of rape. We could run the argument in either direction. We could say that, because castration is permanent, and rape is not, castration is a worse harm to the rapist than the rape was to the victim (and certainly castration was not

42. *Id.* at 33 ("If . . . he has committed a murder, he must die. In this case, there is no substitute that will satisfy the requirements of legal justice.").

43. *Id.* For an illuminating (but also inconclusive) discussion of this example, see David Gray, *Punishment as Suffering*, 63 VAND. L. REV. 1619, 1663–1665 (2010).

44. Kant, *Penal Law*, *supra* note 18, at 33.

45. IMMANUEL KANT, *THE METAPHYSICS OF MORALS* 130 (Mary Gregor ed. & trans., Cambridge Univ. Press 1996) (1797).

46. But see *id.* (explaining that to rape the rapist would be contrary to "punitively justice").

painless in Kant's day). Or we could run the argument the other way. Rape is psychologically and emotionally devastating in a way that castration is not, so that the rapist actually gets *less* punishment than he deserves. Which one of these views is right? More importantly, how do we even go about answering this question (supposing we even wanted to ask it)?

This indeterminacy gets to the core of the point about proportionality. Proportionality in the abstract sounds desirable, a necessary constraint on under- and (especially) over-punishment. But unless we have a reasonable way of determining proportionality, we are left in the lurch. Kant, as the rape example shows, was not ready to submit the offender to the exact same punishment as the offense he committed. In some cases, this would be simply impossible logically, as Kant himself well knew. A person who sodomizes an innocent, vulnerable child cannot himself be turned into a child and then be sodomized.⁴⁷ So, there are difficulties that force us to move away from strict proportionality and into rough commensurability: most crimes can be punished in an equal way, consistent with *lex talionis*, by putting the criminal offender in prison or fining him. Hegel conceded the problem but then tried to finesse it: fines and imprisonment do not have the "specific external shape" of, say, theft or robbery, but they are "comparable," and it is a matter "for the understanding to seek an approximate equivalence in this common value."⁴⁸

Hegel seemed to assume that fines and prison were the place to look for an "approximate value" in punishing crimes; Kant suggested something similar.⁴⁹ But Hegel's move here is questionable. Why is time behind bars viewed as *the* common denominator in punishment?⁵⁰ Because it is the most often used in advanced, Western democracies? But this simply assumes, uncritically, that Western democracies have discovered the correct kind of punishment. We should ask for a deeper, conceptual explanation.

47. *Id.*

48. G.W.F. Hegel, *Wrong*, in Tonry, WHY PUNISH, HOW MUCH?, *supra* note 2, at 47–48. See also KANT, METAPHYSICS OF MORALS, *supra* note 43, at 130 (punishments should be in accord with the spirit, and not necessarily the letter, of the crime).

49. Kant, *Penal Law*, *supra* note 18, at 33 (person guilty of theft should be put to work by the state "for a certain period of time").

50. Why not whips of a lash?

We can imagine some explanatory possibilities. Is imprisonment usually the most appropriate form of punishment because it is the closest to getting at what is wrong with most crimes, namely a deprivation of liberty? This has been a regular answer, ever since it was defended by Hebert Morris in his classic essay on “Persons and Punishment.”⁵¹ But we might wonder whether deprivation of liberty is the chief wrong done by criminals, the one thing about crime that deserves a response from the state. One might think that the essence of the wrong done is not merely a loss of liberty, but the horror of being a victim of a crime (the wrongness of rape is not the wrongness of being constrained, but something much deeper and more wounding to the individual’s dignity; the wrongness of rape is different from the wrongness of kidnapping, or even the wrongness of an assault). There is more—much more—normative work to be done here.

Suppose, for the sake of argument, that imprisonment is our best approximation to a “common denominator” in punishment. Still, there are problems with the very idea of measuring the severity of punishments along any denominator once we abandon a strict *lex talionis* approach, as do almost all modern commentators (including Kant).⁵² There are two steps we need to take in matching crimes with punishments.

The first step is to scale crimes in terms of their badness. This is difficult, if not impossible. We might think that an unjustified *killing* is the worst crime, and the *more unjustified the killing*, the worse the crime. But this ignores how some killings that are motivated by racial or ethnic hatred might be worse than other killings, so that two killings motivated by racism might be worse than two killings done for money (or for honor, or for some other, more seemingly respectable motivation). Or it might be that killing a police officer (or the president, etc.) is worse than killing an ordinary civilian. And certainly, after we move beyond causing a death without justification as the worst crime, we are in very amorphous territory: Is an armed home robbery worse than a wounding? Is kidnapping worse than rape? Is picking a pocket worse than shoplifting? On and on it can go, down the line. Is there any principled way we can say that some

51. Herbert Morris, *Persons and Punishment*, in ON GUILT AND INNOCENCE: ESSAYS IN LEGAL PHILOSOPHY & MORAL PSYCHOLOGY, 31, 36 (1976).

52. KANT, *supra* note 45, at 130.

crimes really are worse than others? Perhaps there is, but I have yet to find it. The fact that we might make some, localized comparisons (stealing \$50 is worse than stealing \$25) should not blind us to the fact that we have no *global* scale on which to rank crimes.

Even assuming there is a principled way to scale crimes in terms of their badness, there are problems. Assume for the sake of argument that we have a rough ranking of the badness of crimes, with murder at the top and then rape, then assault, then robbery, and so on down the line. How do we determine how many years each of these crimes should receive? This may have a deceptively obvious answer (one which many have given), that the worst crime should get the most years (or death) and the less serious crimes should get successively fewer years, until we get to probation or to a fine for the *least bad crime* that we continue to call a “crime” (rather than an “infraction”). But then we have to determine, first, the *top end* of the scale, and second, the *intervals* between punishments for crimes.⁵³ We could fix the top end at death, but then how far down do we go to the next punishment? Do we fix intervals at one, five, or ten years (or days)? How do we decide how finely grained the scale of punishments should be? No matter that we don’t know even how to determine what is in the universe of available punishments in the first place.⁵⁴

Consider an example. In Norway, a mass murderer was sentenced to twenty-one years in prison.⁵⁵ Many were outraged, and said so at the time, but how do we know that this punishment is retributively unjust? If we believe in the death penalty, then we may have our answer: he should be put to death. But take the death penalty off the table (or if you want to leave it on the table, imagine another crime, such as a violent rape, for which one would not deserve the death penalty under *lex talionis*). If the killer deserves more than twenty-one years, how do we know this? What is it about his crime that makes it a matter of retributive justice that he serve

53. See also Russ Shafer-Landau, *The Failure of Retributivism*, 82 PHIL. STUD.: INT’L J. FOR PHIL. ANALYTIC TRADITION 289, 308 (1996).

54. For example, is death a permissible punishment? Is torture?

55. Laura Smith-Spark, *Norway Killer Anders Breivik Ruled Sane, Given 21-Year Prison Term*, CNN.COM (Aug. 24, 2012), <http://www.cnn.com/2012/08/24/world/europe/norway-breivik-trial>. Breivik killed seventy-seven people. His sentence is subject to extension at the end of the 21 year term.

any time in prison at all, let alone twenty-one years? Why not simply public shaming?

If we say, in response, that public outrage, or public sentiment more generally, should determine the appropriate level of punishment,⁵⁶ then retribution emerges as something nearly empty philosophically speaking, a theory that can give us no critical perspective on how much punishment a given crime should receive.⁵⁷ The same also applies if we say that the right punishment is up to the discretion of the judge because we do not know how to determine what limits the judge should work within. In short, retribution promises a proportionality constraint, but it gives us no resources to see how that constraint might yield not just a determinate result but *any* result, even a *range* of results. This has to be considered a serious, indeed almost devastating, failure of retributivism. It says we should punish, but it cannot say how much we should punish: even worse, it cannot even provide a principled way we would go about discovering the answer to the “how much” question.

B. The Modern Retributivists: Morris, Hampton

Two of the most interesting and colorful, and indeed most compelling, philosophical theories of punishment in the twentieth century have both been retributivist. They take their inspiration from Kant and his retributive theory. They, like Kant, express a concern for respect and equality. But where Kant speaks in lofty and transcendental terms, they seek to interpret retribution in slightly more real-world terms. Herbert Morris puts retribution in terms of a binding contract that we enter into with others of our society: when we violate those terms we owe it to society to “pay back our debt.” Jean Hampton develops her retributive theory as one that involves giving an expression of respect to the victim of a crime to “balance out” the contempt that the criminal has shown her or him.

With Kant, I developed my criticism of him on a largely theoretical basis. I questioned whether punishment was necessary for

56. In recent years, Paul Robinson has been the most vocal and consistent advocate of this approach. See his classic essay with John Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997) (arguing that sentences should be assigned based on a community’s perception of “just deserts”). At the very least, the critical potential of such an account depends on the (possibly shifting) perceptions of the community as to what is “just.”

57. If we rely on public opinion to determine sentences, then we are subject to changes in those opinions: they may be quite harsh at one time, and lenient at the other.

showing respect to an offender and his choices. I also wondered whether he could follow through on his claim that retribution could specify “proportionality in punishment” in any meaningful sense. Here I proceed a little differently. I think that Morris and Hampton’s theories are much more attractive, metaphysically speaking, than Kant’s, although I suspect they do not escape the theoretical difficulties that bedeviled Kant. But this is not my main concern here. Rather, I proceed against Morris and Hampton by noting the serious *practical* difficulties their theories encounter.

1. Herbert Morris

Herbert Morris was certainly one of the most influential, if not the most imaginative, retributivists of the twentieth century. For Morris, the controlling metaphor for punishment is quite literally *paying back a debt that we owe*. Society for Morris is a contract we make with others: we obey the laws, and in exchange for that, we get the benefit of living in a society where the rule of law is enforced. When we break the law, this bargain is shattered, but not irrevocably. If we pay an extra burden back to society, then we can regain the equilibrium that our crime has disrupted. In a word, it is tit for tat.⁵⁸

Accordingly, when we break the law, we have to pay society back, and not because of societal stability or some other good. We can imagine that society could survive a certain number of crimes that go unpunished (indeed, it almost certainly does). But that does not remove the unfair advantage that those who commit the crimes have over the rest of society. They get to commit their crime, and in so doing they do not follow the constraints society has laid down on everyone. Punishment’s purpose is to rather forcefully put those constraints back on those who have recklessly cast them off.⁵⁹

In an admiring but critical essay, Jeffrie Murphy criticized Morris’s theory in a way that has, to my mind, never been sufficiently rebutted.⁶⁰ The premise of Morris’s theory is that we all start out roughly equal. Indeed, Morris explicitly said so in his essay.⁶¹ But this is obviously not true.⁶² We do not all get the same

58. Morris, *supra* note 51, at 33.

59. *Id.* at 34–36.

60. Jeffrie G. Murphy, *Marxism and Retribution*, in Tonry, *WHY PUNISH, HOW MUCH?*, *supra* note 2, at 127.

61. Morris supposes that men are “roughly equal in strength and abilities.” Morris,

benefits out of society. Murphy cast his objection in strictly Marxist terms, by thinking of the inequality that mattered as inequality of resources.⁶³ For some, society provides them with lots of resources and the law protects those resources. Others have little or nothing and their material and social deprivations prevent them from getting more. Those on the losing end of the societal bargain do not start out on equal terms with the rest of society, so when they commit a crime and are punished, that punishment is not returning society to an *ex ante* state of equality. It is simply reasserting the earlier, unequal state. In short, if the opening conditions of society are not fair, punishment isn't fair, because it isn't restoring an initially fair starting point. The person who was never given anything by society to start with doesn't "owe" it anything.

But we can make a claim against Morris that is broader than simply the Marxist one about economic inequality; it is in this form that the argument is at its most persuasive because it avoids the Marxism of Murphy. For Morris might respond that even if there is inequality of resources and social conditions, there is still a rough legal equality—equality of the protections of the law against violations of property, person, and life. We are all equally "formal" citizens before the law: we are protected by the rule of law, especially when it comes to the criminal law, even if the rule of law sometimes perpetuates and deepens real, actual, material inequality. This objection seems especially compelling against Shawn Bayern's reformulation of Murphy's argument: Bayern emphasizes the private burdens that people may individually suffer in society and argues that Morris's argument fails to take these private burdens into account.⁶⁴ But if the criminal law protects everybody, Morris can still say that at least our protections under the law are equal, whatever our other private burdens and disadvantages might be.

But it seems patently untrue that the law is enforced equally. In America, the law that protects people against acts of violence or theft is not enforced and implemented in a fair or impartial way such that

supra note 51, at 33.

62. Murphy, *supra* note 60, at 130.

63. *Id.* at 144. This simplifies Murphy's brilliant and sophisticated article to a great degree; still, I think I get at the most important aspect of it.

64. Shawn Bayern, *The Significance of Private Burden and Lost Benefits for a Fair-Play Analysis of Punishment*, 12 *NEW CRIM. L. REV.* 1 (2009).

all equally benefit from the law.⁶⁵ The poor and African Americans simply do not receive as much in the way of legal justice as rich and white people do.⁶⁶ In fact, many times, they are *victims* of the unfair enforcement of the law. To say that the poor and African Americans have gotten a “fair deal” by the law as it is actually enforced is by and large empirically false.⁶⁷ To rephrase it in Morris’s terms: for many in society, the “social contract” is a bad deal because it is unequally enforced against them, and indeed ends up being a tool of maintaining social inequality rather than correcting it.

All of this is to say that Morris’s retributivism does not describe modern society, and its unequal enforcement of the laws. Morris might admit that if society does not fit his description, then society should not endorse retributivism as a theory of punishment, at least not yet. His theory, in short, becomes an ideal theory of punishment. We should work towards social and legal equality (although perhaps for reasons other than that such equality would make retributive justice possible), but until we get there, Morris’s theory does not pull much weight. Murphy reinterpreted Morris as offering just such an ideal theory, and it explains why Murphy was able to hold on to Morris’s theory for as long as he did.⁶⁸ Some do not interpret Morris this way, however, and assert it is a suitable justification for punishment in the status quo. I think this is wrong, unrealistic, and harmful. I think that the most we can say about Morris’s theory is that *it* is a theory for a society very far from our own, one that we may never reach.⁶⁹ Something similar holds for Jean Hampton’s theory of retribution.

2. *Jean Hampton*

If Morris gives Kant’s ideas content by relying on a “bargain” between those in society, Hampton instead refers to the esteem or respect that is promised—and guaranteed—by a liberal society to its

65. See, e.g., JEFFREY REIMAN, *THE RICH GET RICHER, THE POOR GET PRISON* (8th ed. 2006) (documenting the many ways the criminal justice system is biased against the poor).

66. See the especially powerful indictment of the American criminal justice system in MICHELLE ALEXANDER, *THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS* (2010).

67. See *id.*

68. Murphy, *supra* note 60, at 130 (calling for a restricting of society in order to “render punishment permissible”).

69. There is a problem with construing the theory even this way, as I explain *infra*.

members. When a person commits a crime, that person announces to the victim by his or her actions that he or she is “better” than the victim: that he or she gets to do something *to* the victim, without the victim being able to respond.⁷⁰ Note that the offender doesn’t have to lord it over the victim; rather, the crime itself expresses the message of superiority. What punishment does, then, is restore the balance of esteem in a liberal society: the criminal needs to be made low and brought down from his previous “high” of superiority over his victim.⁷¹ The criminal deserves his punishment as a way of balancing out the distribution of benefits and burdens, but a way of canceling out the bad message promoted by the criminal.

By putting things in terms of esteem or respect, Hampton avoids at least one aspect of the main objection to Morris, which is that he ignores the existing inequality of benefits and burdens in society. If punishment is designed to “balance out” resources, and the existing balance is already inequitable, then punishment cannot restore a balance; it will only perpetuate an unbalanced state of affairs. Esteem, or respect, does not work the same way. Respect is not a finite resource. A just state can and should show respect to everybody, by punishing people for the wrongs that they have done.

But the same implausibility that doomed Morris’s argument also looms large for Hampton. If we say that punishment is necessary to show respect to the victim of the crime, don’t we also need to take into account the preexisting distribution of respect and esteem in society (something that is probably not exogenous to the distribution of wealth and other resources)? And here we find a situation that is perhaps even worse than the material and legal case. Is the person who is poor and underprivileged, who commits a theft, or even a violent crime, in a position of overall superiority to the victim of that crime? It is hard to say, at least in the abstract. But the fact that it is an open question (at the very least) shows the difficulty of saying in the abstract that all crimes are crimes in which a criminal asserts his or her superiority, rather than need or desperation or frustration, in committing the crime.

70. See generally MURPHY & HAMPTON, *supra* note 5, at 115.

71. *Id.* at 134 (“The higher wrongdoers believe themselves to be (and thus the more grievously they wrong others), the harder and farther they must fall if the moral reality of the parties’ relative value is to be properly represented.”).

Again, we have to descend from the abstractions of an imagined just legal system and look at the reality of crime and justice in America and elsewhere. In a system that works in some areas, but not at all in others, where the rules are sometimes enforced fairly, but often are not, it is hard to say that the enforcement of the laws will always reassert the balance of respect. Moreover, in a system where it is hard to say that some crimes express anything at all, we cannot say simply as a general matter that crimes are ways of some people lording it over other people. Is the juvenile who robs a store out of pressure from members of a gang, and who does it not to gain money, but simply to gain acceptance and a measure of economic security, expressing his *superiority* over the storekeeper? He may feel a shiver of domination as he holds the gun in front of him, but this cannot be an overall judgment of what he is doing and his situation. For his situation is, overall, a matter of oppression more than domination. He may be dominated by his peers, or by society at large; he may have had a poor to nonexistent education; he may have few opportunities for meaningful employment; he may get little or no support from his family and friends.⁷² This is not to say that the storeowner who is robbed is in any greater position of superiority, or that a failure of the law to respond at all would not diminish *his* moral standing in the eyes of the community.⁷³ But this may show even further the folly of trying to conceive of the criminal law as a mechanism of upholding the respect people are owed; respect is a complex and fluid phenomenon, and it is hard to say who has more or less, or how in any given situation we can distribute it more evenly.

In short, when the juvenile who sticks up a store is punished, is the law canceling out his domination, or is it simply dominating him even more? The juvenile could well see it that way. That the odds were stacked against him, and now they are even more stacked against him: he may face time in prison, and then a wrecked path out of prison. We might imagine ideally, that the law would punish those who commit crimes of domination, and reassert the power of society to cancel out that domination. But things in the real world are rarely

72. See, e.g., CORNELL WEST, RACE MATTERS 12–13 (2001) (noting the “profound sense of psychological depression, personal worthlessness, and social despair so widespread in black America”).

73. That is to say, there must be *some* response of society to the wrong, but how great this response should be, and what should motivate it, are further questions.

that simple. The social meaning of any given criminal act cannot be distilled into a clear message of “I’m better than you,” rather than “I have no other choices in my life but to do this sort of thing,” or “I don’t want to do this but I feel trapped,” or “This is all I can do to feel like a person.” Nor can the message of incarceration be translated into a sign of reducing the offender’s expressive advantage over society, as opposed to being interpreted as *another* (unfair, and disproportionate) way in which society shows its advantage over those who are poor or disaffected, or both. Is the person punished in this way *restored* to his original position, or has he made even lower? The answer, it seems to me, is obvious, especially given the well-known difficulties prisoners have in reentering society.

Hampton’s retributive theory, like Morris’s, attempts to explain why it is that offenders deserve punishment. It is a positive theory because it doesn’t just say that we shouldn’t punish those who haven’t committed a crime, but that we have a duty to punish those who commit crimes. Morris says we should punish because we need to balance out an unjust state of affairs. Hampton says we have to reset the economy of esteem, so that the high are brought low, and the low lifted up. While they are both well-intentioned, they are too abstract, and they imagine an equality of resources, or of esteem, that simply doesn’t exist in any present-day society.⁷⁴

III. DOING WITHOUT DESERT⁷⁵

Part II canvassed some major theories of retributive justice: the classic theories of Kant and Hegel and the more modern theories of Morris and Hampton.⁷⁶ I think that when we consider the theories, and how they spell out their notion of retributive justice, there is much left wanting: for Kant and Hegel, the idea of proportionality cannot be fully cashed out (despite the vague assurance that it *can* be

74. *But see* Jean Hampton, *Correcting Harms versus Righting Wrongs: The Goal of Retribution*, 59 UCLA L. REV. 1659, 1699 (1992) (“So of course there is a problem in punishing the inner city teenager given this theory of retribution-but that is a strength of the theory, because punishing such a person *is* problematic.”).

75. I borrow this title from an essay by Erin Kelly, with which I am in considerable agreement. Erin Kelly, *Doing Without Desert*, 83 PAC. PHIL. Q. 180 (2002).

76. There are other retributive theories, of course; but I have confined myself to the ones I consider the most foundational and the most recently influential. There is also some question in my mind whether all theories that bear the label of “retributive” are *really* retributive, or borrow more from other theories. I hope to address this point in future work.

cash out); the ideas of Morris and Hampton have more (or perhaps further) troubles with application. They (Morris and Hampton) seem to dream of a world of equal resources, or of equal esteem, that is far from actually existing. But there persists a stubborn intuition that criminals *deserve* their punishment—in other words, that bad things should happen to bad people.⁷⁷ Even if we cannot say how this punishment should be measured, whether according to *lex talionis* or in terms of years or something else, or if we are not yet at a place of equality where that punishment can be given fairly or justly, the intuition remains that those who do wrong deserve state punishment of some kind, no matter what the social conditions. So long as we hang onto this intuition in some form, the retributive idea will still remain with us; we will be unwilling to sign on to the program that says punishment is first and foremost about crime control and not about giving people what they deserve, which is at best a secondary and derivative concern.

So we need to get rid of the supposed primacy of this intuition, at least for the purposes of punishment. When we do so, we will find that punishment theory becomes much more open to the real world, and it ceases to be cut off from other aspects of the criminal justice system, such as defining what crimes there are, or what procedures should exist to safeguard the rights of the accused, or even measures to alleviate poverty. In short, when we give up an obsessive focus on desert, we are able to take a much more holistic attitude towards the problem of punishment and will cease to see it as a “discrete” problem at all. This, as it were, is the prize for abandoning the notion that in political society, prisoners deserve their punishment, and the point of the institution of punishment is to give people what they deserve.

Because the idea of desert in criminal punishment goes so deep, and is so pervasive, it requires an attack on three levels. First, I will attempt a direct attack that the idea of desert has any role to play in criminal justice apart from setting up certain side constraints on punishment. The problem comes when we turn the eminently reasonable and acceptable idea that punishment must have constraints into an end in itself and we start entertaining ideas of fundamental desert. The first step in getting over this misconception

77. With apologies to HAROLD S. KUSHNER, WHEN BAD THINGS HAPPEN TO GOOD PEOPLE (2004).

is to diagnose the root of the idea, to explain how it arose in thinking about punishment.

My second attack is more basic and more wide-ranging, and that is to question the idea of pre-institutional desert itself, or the idea that we can be said to deserve anything in political society before we have set up the institutions of political society in the first place. The philosopher John Rawls is famous for making this attack against ideas of pre-institutional distributive justice, but the point has been less common in discussions of pre-institutional retributive justice. There are intuitive reasons why this argument should also apply in the retributive justice debate as well. The idea that desert only emerges after we have decided on the basic institutions of society further underscores the point that desert is at best a derivative consideration in designing and implementing institutions of criminal justice.

The third attack considers what remains after the first two attacks have been completed. It says, insofar as there remains any other idea of desert, any other idea that we may merit reward or punishment, praise or blame for our actions, that idea is properly considered only outside of politics, and not in the design of our political institutions. It would be foolish to deny that there may be ideas of religious desert, or some such thing, so that the wicked deserve to be in hell and the saintly deserve to be in heaven. But such a conception would surely be ill-suited for politics. The same holds for any other idea of desert that treats “desert” as a fundamental concept. One can feel guilty and the need to atone, but these are in the end *private feelings* that are at best secondary to the state’s purposes in punishing. These feelings are matters, to put it again in Rawlsian terms, of one’s own comprehensive doctrine;⁷⁸ your bad feelings about yourself or about others cannot be the basis of a properly political conception.

This is the preview of my attack on desert; now, on to the attack.

A. Desert as Misunderstood Side Constraint

In one of the more striking instances of great minds converging, both John Rawls and H.L.A. Hart had similar insights on the nature of punishment and on the roles that utilitarian and retributive ideas

78. The idea of a comprehensive doctrine was developed in JOHN RAWLS, *POLITICAL LIBERALISM* (1992).

play in the justification of punishment.⁷⁹ The state of the argument prior to Rawls and Hart was a stalemate: neither utilitarians nor retributivists were able to persuade the other that *their* justification was the one that best captured the true justification for punishment. Retributivists insisted that any theory that could countenance the punishment of the innocent—as they claimed utilitarianism could and would do—was wrongheaded and evil. Utilitarianism responded that retributivism would allow punishment that served no purpose—executing last murderer on the island—and so was really, at bottom, a form of revenge: why have people suffer for *no good reason*? The debate between these two *isms* is a staple of introductory philosophy classes, and of first year criminal law courses.⁸⁰

What Hart and Rawls both saw was that the debate, if considered as a debate where one must choose one theory over the other and only one theory could not be solved, and shouldn't be—or at least not on those terms. Both theories contain an element of the truth and both have some fundamental failures. Utilitarianism, without a retributive constraint, is blind to the possibility that it could justify punishing the innocent, or punishing someone without end, if either of these things are necessary for deterrence or the safety of the community more generally. Many utilitarians admit this failure, which is why many utilitarians adopt a side constraint against punishing the innocent. But such a constraint seems rather *ad hoc*. Retributivism as well has its failing, chief among them is that retributivism, without *any* utilitarian aim, is empty: empty of a satisfying reason *or any reason at all why* punishment was necessary or fitting for a wrong done. Many retributive theories, if only implicitly, acknowledge this failure of retribution by “stuffing” their retributive theories with intrinsic goals that are quasi-utilitarian at bottom: such as ensuring democratic self-defense, or promoting equality before the law, or reforming the offender by “communicating” the state’s disapproval of his conduct.⁸¹

79. See Hart, *supra* note 15, at 195; John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3 (1955), reprinted in JOHN RAWLS, COLLECTED PAPERS 20, 23 (Samuel Freeman ed., 1999).

80. See, e.g., DRESSLER, *supra* note 1.

81. These are all goals that Markel has said, at one point or another, his retributive theory embodies. See Markel & Flanders, *supra* note 7, at 938. It seems that many of these ideas can be fully accommodated by a consequentialist theory, especially the need for self-defense, and even the idea of communication, which could be part of a suitably fleshed-out idea of specific deterrence (those offenders who internalize the right values are less likely to

Hart and Rawls saw that neither theory could do well on its own, because both are right, but at different levels of understanding criminal punishment.⁸² The utilitarians are right that the reason we have punishment at all, why it is necessary for society, is that we need to ensure the safety of communities, to control crime. If we had no need for this, then the institution of punishment, established and enforced by states, would have no sensible *raison d'être*. There would be no reason for a society on its own to set up punishment just so it could punish the guilty if that had no relationship to some other, deeper, utilitarian purpose like maintaining order or preventing private acts of vengeance. Nor is it clear that an institution of punishment could exist for very long if it fulfilled none of these utilitarian purposes or worked in a way contrary to them. Whatever the origins of punishment (if it began, say, to institutionalize private acts of revenge), the function of punishment in the modern, bureaucratic state is crime control.

But this is not the end of the story. We may agree that the reason we have the institution of punishment is to control crime, but this fact doesn't dictate how the institution should be run, once we have it up and running. And it is at the level of running the institution that retributivism comes into play, not in explaining *why* we punish, but in dictating *how* we should punish. This insight was the great innovation of Hart and Rawls. We punish people in general because this keeps society safe, but we have rules about what merits punishment. For instance, one can be punished only when he has committed a crime, and he can be punished only proportionally to his wrong. These are retributive rules about punishment. They do not derive from utilitarian reasons, or at least, they do not derive obviously from utilitarianism. They are (more or less) independent rules about when and how we can punish people. We might wonder *why* we have these rules, whether they reflect concerns about human dignity, or about fair notice, or something else.⁸³ But Hart and

offend again). I deal with the idea of equality before the law in the conclusion to this essay.

82. For a longer look at Rawls and Hart in the context of debates about cost and sentencing, see Flanders, *supra* note 9.

83. Hart grounded the prohibition against punishing the innocent on legality principles, not directly on desert-based ideas. See Hart, *A Prolegomena to Punishment* in TONRY, WHY PUNISH?, *supra* note 2, at 200 (retribution is distinct from the "simple insistence that only those who have broken the law—and voluntarily broken it—may be punished"). We might also have a prohibition against "excessive" punishment without this being founded on anything

Rawls say that we best understand these rules as coming in at the level of operation, and not at the level of the justification of the institution as a whole. We don't need to be retributivists all the way down to see that retributivism had some good ideas about when we *shouldn't* punish.

But for Hart and Rawls, we make a basic mistake when we take the retributive constraints on punishing (don't punish the innocent, don't punish excessively) as the reason why we punish in the first place. To do this is to turn those constraints inside-out, to make them do work that it is not appropriate for them to do: to place on retributive constraints the burden of justifying the institution of punishment as a whole. It would be like saying the reasons we have government is so that we can have Fourth Amendment constraints on the power of the government. But this gets things exactly backwards. We have the Fourth Amendment, among others, to constrain the power of government, and government exists for reasons and purposes other than what the Fourth Amendment aims to protect. The retributivists make the mistake of turning something that restricts punishment into something that explains why we have punishment.

Of course, retributivists might ask: why do we have this constraint? Don't we need a notion of desert to show why we don't punish the innocent? But this supposes that only the notion of desert can do this work, and not anything else. And this supposes, in turn, that we can make sense of desert as a stand-alone idea. But what does it mean to say not just that the innocent shouldn't be punished, but that the guilty actually deserve their punishment? Does desert, in short, exist?

*B. Does Desert Exist?*⁸⁴

My argument in the previous section was, in part, both genealogical and descriptive. I wanted to explain how we might get our intuitions about desert in the first place. We may be tempted to

desert-based: just an intuition that some punishment is just more than we can or should stand, or that is incompatible with human dignity (one might object to the death penalty or torture on this ground). Of course, a punishment can also be excessive on deterrence grounds, when the punishment is greater than that which is necessary to optimally deter. For another attempt to justify proportionality outside of punishment theory, see Alice Ristroph, *Proportionality as a Principle of Limited Government*, 55 DUKE L.J. 263 (2005).

84. Here I summarize an argument I make at length in Chad Flanders, Punishment and Political Philosophy: The Case of John Rawls (unpublished manuscript) (on file with author).

think that the norm of not punishing the innocent exists as part of a larger principle that says we ought to punish the guilty. But as I suggested above, we don't need to think of desert in this way. Instead, we can look at desert not as an independent principle governing punishment, but rather as a short-hand way of thinking about certain side constraints on punishment, i.e., don't punish the guilty, don't punish disproportionately, punish like cases alike, etc. The goal of punishment is still utilitarian, but it has side constraints. We make a basic mistake when we turn what are really constraints into stand-alone principles.

But we can also levy an even stronger claim against the retributivist. For he has not just made a mistake in thinking that what are constraints on punishment provide the fundamental reason for punishment. In fact, he has settled on a concept—the concept of desert—which is, at the end of the day, empty. We cannot give any meaningful content to desert, because desert in the sense the retributivist needs it to exist, does not exist.

We have already seen one aspect of this puzzle in our discussion of Kant. There, we asked how we decide what proportionality in punishment requires. How many years is “fitting punishment” for a rapist? For a kidnapper? Once we get away from a strict *lex talionis* approach to proportionality, all we have are our intuitions, and our intuitions have precious little to go on. In fact, our intuitions tend to “anchor” in our culture, finding what is the conventional punishment to be the “appropriate” punishment in that case.⁸⁵ But we have little reason to trust that our culture has magically hit on the exact right number of years for punishment, any more than present-day Norway has, or for that matter, China in the 1500s.

There is an easy explanation for this puzzlement when it comes to retributive proportionality, and it is this: desert does not exist as an independent concept. If we ask, in a vacuum, what people deserve, we have a hard time coming up with a determinate answer. The questions get harder to answer once we get more specific: what does a battered wife who kills deserve? What does an armed robber who is a juvenile deserve? Does anyone deserve the death penalty? The questions are hard here, but their difficulty derives from an even more fundamental problem, which is the absence of any

85. For a discussion of “anchoring” in the context of sentencing, see Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2464, 2515–16 (2004).

guidelines about the concept desert itself.

The best argument against desert comes from Rawls again, but this time from his later, more mature work, *A Theory of Justice*.⁸⁶ Rawls made the argument about desert in the context of distributive justice, but we can easily transpose it into the key of retributive justice as, indeed, many have.⁸⁷ Rawls can be read in a radical way, as to say that there is no such thing as an “independent” concept of desert. Or it can be read in a less radical way, as to say that desert is a concept not appropriate for a modern, political society such as ours. Both arguments are, I think, devastating to the concept of desert. I consider the first one in this section; I consider the second one in the next. Even if we don’t find the first version of the argument persuasive, the second one is really all that we need to do away with desert as an independent concept.

Begin with Rawls on distributive justice. Rawls says that one principle we might think of when it comes to desert is this: give resources to those who deserve them, who score high on some metric of merit, or worth, or deservingness. We look at how hard people work, or what they produce, or how good of people they are, and then reward them accordingly. Rawls says that to do this is impossible. Why?

One way Rawls is commonly read is that no one can deserve anything because we don’t deserve our talents.⁸⁸ Even our ability to work hard is not our doing; it is something we were born with, or were lucky enough to acquire. Because we don’t deserve our talents, we can’t be said to deserve anything that flows from those talents. This is certainly one way to read Rawls, and it is supported in some parts by the text; I do not think it is the only way (although it seems to me the impulse behind this reading is one worth investigating). In this way of reading Rawls, he is a skeptic about our ability to deserve *anything*, because we—as individual humans—don’t have the proper “desert basis” to do so. To properly deserve anything, we would have to be responsible *all the way down* for being who we were, for creating ourselves almost.⁸⁹ But no one creates themselves;

86. JOHN RAWLS, *A THEORY OF JUSTICE* (1971).

87. See especially the discussion in Samuel Scheffler, *Justice and Desert in Liberal Theory*, 88 CALIF. L. REV. 965 (2000).

88. A reading given forcefully in MICHAEL SANDEL, *LIBERALISM AND THE LIMITS OF JUSTICE* 34 (1992).

89. See ROBERT NOZICK, *ANARCHY, STATE, AND UTOPIA* (1974) for this argument.

we need society, we need our genes, and we can't be said to have created those things about us. On this line of reasoning, we can't be justly punished, because we aren't responsible for being who we are.

Again, some have read Rawls this way, but I don't think it is the only way to read him. Consider a more "institutionalist" reading.⁹⁰ Read in this way, Rawls says that how we deserve something does not exist apart from the institutions of society. We can only say that we deserve something if we know the scheme or structure we are working with: desert is a measurement that does not exist prior to society; it exists *after* society and is used for the purposes of that society. To take an example from Rawls's earlier essay, "Two Concepts of Rules," we only know who deserves to win a baseball game once we know what the rules of the game are.⁹¹ We can only say that someone deserves something after we know what the rules of the game are. Similarly, it is only after we have set up the basic institutions of society that desert exists as a meaningful concept at all. The question of who deserves to win the game of baseball doesn't get any traction until we actually *have* an institution called "baseball," and certain skills, talents, and results become salient.

The same applies to other institutions of society. How much money should doctors earn? What about firemen? Should we pay public school teachers more?⁹² How do we determine *these* things? Well, we need to know what system we are working with; we need to know the rules of the game. Once we have established this, we can go on to say what doctors should earn, but not before then. Because it turns out there is no fact of the matter about what doctors ought to earn prior to us knowing what, in fact, society is for and what the rules of society are in order to achieve the goals of society. What people "deserve" will be different depending on what the organization of that society is, what principles it is run by, and what those principles say about desert.

For Rawls, society should be organized around the difference principle: differences in the distribution of resources should be made in order to help the least well off.⁹³ How much people ought to earn

90. See Scheffler, *supra* note 87.

91. RAWLS, *supra* note 79.

92. For a perceptive discussion of these issues (focusing on desert in *distributive* justice), see MICHAEL SANDEL, JUSTICE: WHAT'S THE RIGHT THING TO DO? (2009).

93. RAWLS, *supra* note 79.

depends on what the principles of justice say about how society should be organized. Until we know what those principles are, we simply don't know what people deserve. To borrow from Samuel Scheffler, there doesn't exist anything like "prejudicial desert."⁹⁴ Desert is a concept that we derive from justice; it has no independent moral force, because whatever moral force it *does* have comes from the principles of justice by which we define what desert is in the first place.

The same argument applies, if we make some adjustments, to punishment.⁹⁵ How much time do rapists and murderers deserve? What crimes should be punished by prison, and what by probation? What crimes should even be considered *crimes*? We can't say what the answers to *these* questions are, either, until we have set up the institution of punishment, and know what the goals of punishment are. But this means that desert can't be understood as a concept prior to the institution of punishment: we can't know what we ought to punish, and for how long, until we have certain institutions set up. This is not the same as the more skeptical reading of Rawls which says that no one deep down has the capacity to deserve anything, because no one is responsible for all he or she is. Rather on this less skeptical reading, people do deserve things (and, in a sense, they "really" deserve them to the extent that anyone can deserve anything). But they deserve things only after we have decided what the rules of the game are.

Now, there will be constraints on awarding desert or meting out punishment. There are constraints in Rawlsian distributive justice, after all. We do not use the difference principle in *all* areas of life: we don't let it rule over individual decisions, for instance, or require families to distribute resources according to it.⁹⁶ The difference principle is the rule that organized the basic structure of society, and there can be other rules that limit how the difference principle operates. The same goes for punishment, too. We have the basic institution, set up for a particular purpose or purposes, and then we put constraints on how we can achieve those purposes.

94. See Scheffler, *supra* note 87.

95. See Douglas Husak, *Holistic Retributivism*, 88 CALIF. L. REV. 991 (2000), for similar arguments.

96. The limitations of the difference principle are examined—and critiqued—by G. A. Cohen in G. A. COHEN, *RESCUING JUSTICE AND EQUALITY* (2008).

Rawls insists in *A Theory of Justice*, somewhat confusingly, that retributive justice should not be thought of in the same terms as distributive justice, suggesting perhaps that pre-institutional desert belongs in the former, but not the later, area.⁹⁷ Rawls seems to think that our pre-institutional instincts about punishment are stronger than our pre-institutional instincts about desert. But his remarks on this point are less than clear. It is true that retributive and distributive justice have different goals. This of course is obvious: the goal of a scheme of distributive justice is to encourage people to work and to make sure that all inequalities of resources go to the least well off. It would be bizarre if this were the goal of a system of punishment! A system of punishment, rather, tries to discourage crime and protect society, within fair limits (fair notice must be given, the innocent should not be punished). But setting apart the obvious difference in goals, the structure of the two institutions is the same, and so is the place of desert within them. Desert arises as a meaningful term only once we know what the rules of the game are. In other words, people “deserve” their punishment because the institutions of society say they do, not because there is some pre-institutional notion of desert that the institutions of society somehow have to live up to.

C. Desert in Comprehensive Doctrines

But this may seem to move too quickly and require qualification because surely we have *some* idea of pre-institutional desert. We have this in both the distributive and the retributive context. We have some idea that good acts should be rewarded and bad acts should be punished, independent from any institutional context. Institutions, some might object, really do exist to give shape and form to these intuitions, not the other way around; it is not as if the institutions create these intuitions of desert, which we then follow. Rather, we have the intuitions first, even if they are inchoate, and hard to define. When children see bad things happening, they sense that wrongs should not go unanswered, that bad men should not be allowed to get away with doing bad.⁹⁸ Similarly, when we see someone do a

97. RAWLS, *supra* note 79, at 314–15; *see also* Flanders, *supra* note 84.

98. Here I am indebted to unpublished work by Thomas Simon on the nature of our “brute” emotions. *See* Thomas Simon, Injustice Emotions (unpublished manuscript) (on file with author).

good act, we want to benefit that person: we instinctively want to praise her. These are not feelings created by institutions; they are, as it were, part of human nature. Our intuitions are responding to these feelings. Moreover, we use these intuitions to critique existing institutions; when people get more or less than what we think they deserve, we balk.

I think this is probably true, even undeniable, but it does not, or does not necessarily, affect the earlier point. For example, consider the wide distance between our simple (“brute”) intuitions of praise and blame, and the actual institutions of distributive and retributive justice. We again, of course, face the problem of indeterminacy in reward and punishment. How much should the good person gain from his good act? How much should the bad person suffer in proportion to his wicked act? The idea that there is some pre-institutional, and yet *firm*, idea of praise and blame that can give us determinate answers to these questions seems to me to be far-fetched. So even if we agree that we have these feelings, we still need institutions to give a shape to them, to give them form and precision.

Further, there is still a long way between the fact that we have these feelings of praise and blame and the point that it is *the purpose of political institutions* to give expression to these feelings. There is a gap here, which we have to take account of, and notice. We might get mad or upset at all sorts of things, but we do not think they should be punished, or, that they should be punished by the state. At the very least, we need an argument as to why some things should be punished and for how long, and we can’t make this argument just in terms of our brute feelings.

We can make this point in terms of what Rawls would later call “political liberalism.”⁹⁹ Rawls says, rightly, that modern society is characterized by deep conflicts over ways of looking at the world, what he calls “comprehensive doctrines.”¹⁰⁰ Some look at the world religiously; others don’t. Some subscribe to an ethics of absolute duty; some are pragmatists. If we tried to convince one another about these doctrines, we would be bound to fail. The arguments, at the very least, would stretch out interminably as we would be

99. RAWLS, POLITICAL LIBERALISM, *supra* note 7878. Jeffrie Murphy explores a similar line of thought in JEFFRIE MURPHY, PUNISHMENT & THE MORAL EMOTIONS 64–93 (2012).

100. RAWLS, *supra* note 79, at 12.

starting from different places, with different experiences, and different values. We'd never get to an agreement.

But still, we have to live with one another, so we find a way to agree on political matters, matters that do not rely on our "comprehensive doctrines." To take one example: some may think that there are religious wrongs, that some things offend God, and that they should be punished. But others may disagree. They may think that God is not so easily offended, or they might think that there is no God. There is probably no way of successfully settling this debate. So we try to put it off the table. We try to say, instead, let us talk about things we can all agree should be punished, that don't depend on one's belief in God, or in any particular "comprehensive doctrine."

I think the idea that there is a pre-institutional notion of desert is an example of an appeal to a comprehensive doctrine. People will, it seems fair to say, disagree fundamentally about desert. Some will deny there is such a thing: Rawls, read one way, does precisely this. He says (as interpreted by some) that we are not responsible for anything because we are not responsible for our genes or our parents or where we were born. Some will affirm that there is such a thing as desert, but still say that people need material preconditions before they are properly held responsible for anything.¹⁰¹ Some will say that people always have the freedom to rise above whatever material circumstances they face, and to do the right thing and refrain from criminal activity. Still others will deny the existence of free will. These are debates that are about pre-institutional, or pre-political, desert, about what we might be said to deserve, before there is any society, or apart from any societal constraints. They involve what are quintessentially metaphysical matters, about the nature of persons, and their relationship to society. They are debates that are likely to be interminable. They qualify (if anything qualifies) as debates we would want to keep "off the table" if we are trying to figure out how to govern ourselves, how to live together.¹⁰² We would want to keep that debate metaphysically simple. We would want to have a debate about the shape of the intuition of punishment, about what

101. As Murphy suggested in his critique of Morris. See Murphy, *supra* note 60.

102. See also the similar argument made by Jeffrie Murphy in Jeffrie Murphy, *Retribution and the State's Interest in Punishment*, in NOMOS XXVII: CRIMINAL JUSTICE 156-64 (J. Roland Pennock & John W. Chapman eds., 1985).

punishment is *for*, that does not depend on deciding deep questions of responsibility and desert.¹⁰³

So we might agree that there is something like pre-institutional desert, some first or original principles of merit and demerit before we enter into society, but whatever that idea is, whatever shape it might take (determinate or indeterminate, religious or secular, Kantian or Marxist), it cannot be used as the basis of political society. When it comes to politics, we have to look for rules that we can all agree on about the aims and purposes of punishment.

This debate, I think, will look a lot like the one Rawls and Hart described in their early essays. We will look at punishment as a matter of crime control first and foremost, which is precisely why we would need an institution like punishment.¹⁰⁴ It would also include a discussion of side constraints, about not punishing the innocent. But the debate about side constraints would need to be (at best) incompletely theorized.¹⁰⁵ We would not want to rely on a deep theory about why we should not punish the innocent and punish only the guilty. We should give notice, we should punish fairly, and we should not punish excessively. All of these requirements can and should be separated from the idea that we should punish people because they deserve it in some pre-institutional sense.

Those who wanted to make the argument that their brute feelings demand justice would need to make a political argument about why the institution of punishment should correspond to their feelings. Their feelings alone are not sufficient, publicly acceptable grounds for punishment. Thus, when Michael Moore writes that he would feel so guilty he would want to die if he had committed a murder, this is *not* an argument that he should be punished.¹⁰⁶ It is simply a report of his feelings. Even group feelings are not a political argument. This is not to say that they cannot be made into one, but as they stand, they do not rise to that level.

103. This does not mean that there are no ways in which *parts* of retributivism may be part of a properly political conception. Mostly, I am doubtful that any notion of desert can properly be called “political.”

104. We should not, however, “beg the institution” of punishment. What we need is crime control: whether criminal punishment is the best way to achieve this end is an open question. For an analysis of this question, see DAVID BOONIN, *THE PROBLEM OF PUNISHMENT* (2008) (advocating the abolition of prison).

105. See CASS SUNSTEIN, *LEGAL REASONING & POLITICAL CONFLICT* (1996).

106. MICHAEL MOORE, *PLACING BLAME* 145 (1997) (stating that if he had killed a 23-year-old woman, he “would feel guilty unto death”).

So does pre-institutional desert exist? It may, but if it does, it is not a suitable ground for designing institutions of punishment. For one, it is too indeterminate: we cannot get specifics about what to punish, and how, from intuitions about bad acts. For another, pre-institutional desert does not give us the right kind of reasons for building and maintaining just institutions.

IV. RETRIBUTIVE THEORY IN PRACTICE

Traditional versions of retribution fail as a matter of a theory. In its “popular” theoretical forms, the ones familiar from first-year criminal law classes and given sophisticated shape say, in Kant or Hegel’s philosophy, or in Morris and Hampton’s modern-day updates, retribution doesn’t work. Retribution doesn’t work because when we try to spell it out, we either get incoherence and indeterminacy, or worse we get a type of unworldliness and unreality: a society in which the laws are enforced equally, or that people start out on a level playing field, which crimes upsets and punishment restores. But the problems go even deeper. The very idea that anchors retribution, the idea that gives it its moral force and power, the idea that we deserve punishment when we have done wrong, is either empty, waiting for content from utilitarian norms, or at best an idea beset by controversy, and inappropriate for political debate. For in political debate we need consensus, we need agreement, and we can’t get this with as fuzzy a notion as desert.

In the retributive picture of the world, people freely commit crimes, and they subsequently deserve punishment. Punishment shows respect for those who are punished, it shows them that they are taken seriously, that their actions have not gone unnoticed, that (in a slightly perverse way) the community cares about them. If society did not care for the criminal, they would treat the criminal, or try to threaten him or her, or worse, ignore him or her. Punishment demonstrates that the criminal and his acts matter.

Society on the retributive picture is by and large fair (in the metaphor of Morris, a fair bargain), either as a matter of empirical fact, or as a kind of transcendental truth. Crime damages that fairness: someone is getting away with something at the expense of all of us. We, as a community, have to react to this harm against us, not merely the physical harm that, say, murder or assault or robbery or rape represents, but the moral and political harm that the criminal act represents. The punishment symbolically negates that harm, by

either forcibly resetting the balance of benefits and burdens (Morris) or correcting the false message the criminal presents by his crime (Hampton).

This is a vision of the world that is stirring, almost ennobling, certainly inspiring, and possibly even aspirational (when charitably read). It helps that some of the writers in the retributive tradition are brilliant prose stylists. Their writing *moves*. In fact, it is almost tempting to read some modern retributivists, such as Morris and Hampton, as pointing to ideals of punishment. We may not be able to punish, justifiably, on their theories *now*, but someday (when things are more equal) we might be able to. But even this seems a little strange. We should not try to inhabit an ideal world because that would enable us to punish more justly; an ideal world is a worthy goal in its own right (if anything is).¹⁰⁷ Nor does it answer the pressing question, what do we do in the meantime about crime and punishment?¹⁰⁸

On this last and most vital of questions, I maintain that the retributive theory has little to say: it is too caught up in ideals and abstractions. And when it does have something to say, traditional retributivism leads us in the wrong direction. Retribution faces practical as well as theoretical difficulties.

A. Retribution as Limited in Practice

Consider four areas of criminal justice that are among the most hotly debated today, and that will be exemplary for my purposes. In these three areas, retribution either has nothing to say, what it says is useless, or what it says is positively harmful.

1. Cost containment

Amazingly, we have reached a point in American criminal justice where there is actually positive legislative pressure against building more prisons and increasing sentences for criminal offenders.¹⁰⁹ One would have thought that this might have never come to pass. There has traditionally been considerable political upside to being tough on

107. Cf. Murphy, *supra* note 60, at 144 (changing “basic social relations” the key to just punishment).

108. I discuss the gap between the ideal and the real at greater length in my *Pardons and the Theory of the “Second Best,”* 65 FLA. L. REV. 1559 (2013).

109. See Dagan & Teles, *supra* note 9.

crime. And conversely, one gains little by being seen as sympathetic to prisoner's rights or to the rights of criminals. Moreover, prisoners have no organized lobby, save some civil rights groups and the members of the prisoner's families. The prison industry (public and private), however, *does* have considerable lobbying force, and has an interest in filling beds and building prisons.¹¹⁰ Moreover, criminal punishments may tend to ratchet "up" over time, as long sentences become the new baseline, and the new outrageous crime needs to be more than the current baseline.¹¹¹

But the recent financial crisis has proven to be a strong enough counter-trend to this "natural" state of affairs. Prisons and prisoners cost money to house and feed. Guards and new buildings are expensive. So something amazing has happened: conservatives, previously "tough on crime" have now become, in the current phrase, "smart" on crime.¹¹² They have realized that many states cannot afford to spend more on criminal justice, or at least on criminal punishment. They are looking at alternatives to prison: they are looking to probation, to treatment, and to decriminalization.

Can retributivism say anything about this issue? Not really, at least historically. Or if it does say something it says that this move to controlling costs is mistaken. Punishment is a matter of justice, Kant says in several places, as do many contemporary retributivists. We should not change or alter the proper punishment for someone because we need to save money, or because an alternative is cheaper. The murderer must be executed, even if there is no social purpose provided by the execution, or even if leaving him in exile would be cheaper. This harshness is part of the legacy of Kant on punishment.

Some retributivists have tried to make retribution more "real world" by saying that giving people what they deserve is only *one factor* among many that judges and legislators have to weigh, that retributive punishment is not an absolute duty, but a good among

110. Charles Blow, Op-Ed., *Plantations, Prisons, and Profits*, N.Y. TIMES, May 5, 2012, at A21 ("[O]n almost a daily basis, local prison officials are on the phones bartering for prisoners with overcrowded jails in the big cities.").

111. See Chad Flanders, *Shame and the Meanings of Punishment*, 54 CLEV. ST. L. REV. 609, 631 (2006).

112. See, e.g., *Statement of Principles*, RIGHTONCRIME.COM, available at <http://www.rightoncrime.com/the-conservative-case-for-reform/statement-of-principles/> (last visited Mar. 15, 2014) (describing the "conservative case" for reforming prisons and jails). See also *supra* note 9.

other goods.¹¹³ But how should we weigh this good compared to other goods, how do we manage the trade-off? How do you weigh the justice of a punishment, the necessity of punishing evil, with other factors, such as saving money? Traditional retributivists would say you can't: justice can't be "balanced off" against other factors, without missing the point of what *justice* is. Contemporary retributivists give us no guidance as to how we should balance off retributive justice and why we would ever want to.¹¹⁴ Retribution here mimics the problems with absolute rights (rights as "trumps")¹¹⁵ in other areas. It gives, at best, no guidance and at worst guidance of the wrong sort by saying that we should simply ignore cost, as if this were possible in the real world.

2. *Mass incarceration*

The term "mass incarceration" has now become familiar in penological discourse.¹¹⁶ The statistics are also drearily and depressingly familiar. Several numbers stand out: the percentage of adult Americans (and especially African-Americans) in prison, and the sheer number of Americans in prison compared to the rest of the world.¹¹⁷ What does this phenomenon mean about America and what should be done about it? Like the problem of prison costs spiraling out of control, many remedies have been talked about, and floated.¹¹⁸ But most importantly, mass incarceration is now treated as a problem in itself, over and above other problems with the administration and functioning of criminal justice.¹¹⁹

113. See Cahill, *supra* note 8, at 866–67 (author's approach "does not view desert as a fundamental *ex ante* obligation (or restraint), but rather as one good to be balanced against other possible goods in determining whether punishments appropriate in a specific case").

114. A point I make in more detail in Flanders, *In Defense of Punishment Theory, and Contra Stephen*, 10 OHIO ST. J. CRIM. L. 243 (2012).

115. See RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* (1978) (defending notion of "rights as trumps"); see also Shafer-Landau, *supra* note 53, at 306 (strong retributive theory "forbids any cost-benefit calculations in the creation or application of sentencing guidelines").

116. See ALEXANDER, *supra* note 66.

117. Adam Gopnik, *The Caging of America*, NEW YORKER, Jan. 30, 2012, at 72 (describing the rise of mass incarceration in America).

118. See, e.g., Carol Steiker, *Mass Imprisonment: Causes, Consequences, and Exit Strategies*, 9 OHIO ST. J. OF CRIM. L. 1 (2011).

119. Dorothy Roberts, *The Social and Moral Cost of Mass Incarceration in African American Communities*, 56 STAN. L. REV. 1271 (2004).

Can retributivism even recognize mass incarceration as a problem? I am not sure it can. To do so, it would have to be more specific about particular sentences and why they are unjust; it would have to say more about what kinds of things should be punished. If we leave *these* questions to elected legislatures, then we risk getting no critical purchase on the problem of mass incarceration because the problem is by and large a democratically created one. Or if we leave the problem of long sentences open because we cannot say what proportionality requires, then we cannot criticize some sentences as excessive, at least not without any precision. Nor does retribution as a theory of punishment have anything to say about overcriminalization, a phenomenon that is, for obvious reasons, closely related to mass incarceration.

But if we are worried about the social and economic costs of punishing people too long, we might see mass incarceration as a waste of both money and lives. Or as a matter of political philosophy, we might see mass incarceration as presenting a problem, even a crisis, of state legitimacy.¹²⁰ Retribution is at best silent about the problem of mass incarceration, or says: if prisoners are justly in prison, then mass incarceration is required as a matter of justice. There can be no “further” problem of the large number of people behind bars on the retributive way of looking at things.

3. Mandatory minimums

Not unrelated to the previous problems is the existence of mandatory minimums—that is, fixed minimum sentences for certain crimes. Mandatory minimums for some nonviolent drug crimes have been the focus for some critics, but the problems with them can be described in general terms: mandatory minimums give prosecutors too much power, force judges’ hands at sentencing, and frequently result in sentences that are too long. The solution to this is to get rid of the minimums, or lower them substantially, and give judges more discretion in sentencing offenders.

Does retribution have anything to say about this? We might think it does. Retribution might say that some of the minimums are too long, and so are disproportionate, but it lacks any firm basis to

120. See in this regard, the important pieces by James Forman and Glen Loury. James Foreman, Jr., *Why Care About Mass Incarceration?*, 108 MICH. L. REV. 993 (2010); LOURY, RACE, INCARCERATION, & AMERICAN VALUES (2008).

say so for reasons we have already canvassed. Again, retribution as a general theory is too indeterminate on what counts as proportional or not: it cannot say *what is too long* for any given crime. If the retributivism says that, instead, we fix the minimum by asking people what is an appropriate minimum sentence, then retribution is openly giving up on its critical purchase. For the minimums that we have are precisely the consequences of the “people” acting through the legislature to fix the appropriate sentences.¹²¹

In fact, if anything, retribution would seem to favor something like mandatory sentencing not just for selected areas of the criminal law, but for every area. Retribution favors a principle of equality in sentencing, at least as a formal principle: like crimes should be treated alike, and no differences in punishment based on the sensibilities of the offender or on whether a longer or shorter punishment is necessary for deterrence. By treating like offenders alike, the retributivist treats people equally; as such, the retributivist will tend to be against giving judges too much discretion in punishing offenders because this leads to unequal treatment of similarly situated offenders.

4. *The death penalty*

The death penalty has long been an object of contention and criticism,¹²² with some of the chief arguments defending it coming from retributivists (Kant, again, is the main target). Where the deterrence argument for the death penalty has seemed indeterminate or vague or worse, retribution has seemed to be a reliable “go-to” justification for putting people to death for their crimes. But recently, some have maintained that retribution shows why the death penalty is immoral. Dan Markel and Thom Brooks, for very different reasons, have maintained that retribution positively excludes the death penalty. Their arguments are not convincing.¹²³

Retributivism, at its core, says that wrongdoing deserves punishment, and the greater the wrongdoing, the greater the punishment. Nothing in this rules out the death penalty; indeed,

121. Paul Cassell & Erik Luna, *Mandatory Minimalism*, 32 CARDOZO L. REV. 1 (2010).

122. See DAVID GARLAND, *PECULIAR INSTITUTION: AMERICA'S DEATH PENALTY IN AN AGE OF ABOLITION* (2010).

123. See generally Chad Flanders, *The Case Against the Case Against the Death Penalty*, 16 NEW CRIM. L. REV. 595 (2013).

Kant thought that it was one of the clearest cases of retributivist philosophy that murderers deserve the death penalty. This was proportionality at its purest. And yet, Markel's argument relies (in part) on the idea that the death penalty offends dignity, and dignity is a value intrinsic to retribution.¹²⁴ But Markel's burden here is twofold: first, to show that dignity is indeed intrinsic to retribution rather than extrinsic, and second, to show that the type of dignity that *is* intrinsic to retribution is one that rules out the death penalty. I am not sure this can be done. If a person gets what he deserves, the retribution is satisfied; there is nothing inherent in retribution (as opposed to principles taken from liberal political philosophy, or elsewhere) that says that punishment has to be done in a dignified rather than an undignified way. Any connection between retributivism and human dignity is, I think, a contingent one, not a necessary one. Nor is it obvious under a retributivist paradigm that putting someone to death for a crime is not to treat him with dignity; Kant certainly felt otherwise.¹²⁵ Dignity, here as elsewhere, is a stirring word that means something, but what that something is can be notoriously hard to define.

Brooks bases his argument on the possibility of error in meting out the death penalty.¹²⁶ To him, retribution means not punishing the innocent, and putting the innocent to death is a mistake that we cannot undo. But this is, at best, merely a practical argument about how retribution rules out the death penalty; it is also a perfectly general one, that would apply to (and thus possibly rule out) all punishment.¹²⁷ Moreover, it assumes that, on balance, it is worse to kill the innocent than to not kill those who deserve to be executed. That the balancing should go like this (rather than the other way, where not giving people what they deserve is worse than killing an innocent person) is nowhere entailed by retributivism.

124. See Dan Markel, *State, Be Not Proud: A Retributivist Defense of the Commutation of Death Row and the Abolition of the Death Penalty*, 40 HARV. C.R.-C.L. L. REV. 407, 466–67 (2005).

125. Kant, *Penal Law*, *supra* note 18, at 32 (death only appropriate punishment for murder; man of “honor” would choose death rather than penal servitude).

126. Thom Brooks, *Retribution & Capital Punishment*, in RETIBUTIVISM: ESSAYS ON THEORY AND POLICY, *supra* note 2, at 236–38.

127. A common rebuttal to this is that death is irrevocable, but a prison sentence is not. For an effective response to this see MATTHEW KRAMER, *THE ETHICS OF CAPITAL PUNISHMENT* (2011).

If we want a cogent argument for abolishing the death penalty, we should not look to retribution for it. Indeed, given how empty retributivism is in its prescriptions, it is unclear that retribution can rule out *any* mode of punishment. Why couldn't torture be an acceptable punishment, if it gives the guilty what they deserve? If the retributivist cannot say what the guilty deserve, only that they deserve punishment, the retributivist cannot give us a positive reason to oppose torture in punishment. We have to look for that argument elsewhere.

B. Why Retribution Fails, Practically Speaking

Why does retribution seem to say so little on so many pressing issues of criminal justice, and when it does have something to say, it seems to say the wrong thing? These exemplary examples represent the general failings of retribution. I think these failings can be attributed to two features of retribution, which we can call, for convenience, *closure* and *compartmentalization*.

By *closure*, I mean that retribution, in most of its versions, gives us a theory that is mostly detached from any empirical truths about the world. It disdains (sometimes explicitly, as in Kant, but mostly implicitly) any reliance on the “facts on the ground” about resource distribution or about moral responsibility. It tends, as with Kant, towards a reliance on transcendental truths about human beings, or speaks in generalities about equality and dignity. It does not come to grips with the facts on the ground about crime and punishment and instead attempts—but largely fails—to provide an ideal theory of punishment.

By *compartmentalization*, I mean that retribution focuses solely on the moment of punishment and resists any discussion about criminal justice in general. It sets as its task only the question of what would justify punishment for crimes, and leaves to another day or another theory subjects that would otherwise be essential to a complete theory of the criminal law.

1. Closure

We can take the theories of Morris and Hampton as providing good instances of *closure* to the real world. Morris and Hampton do not consider what actually may be true of equality in the real world (except, perhaps, by simply assuming that all people are in fact

equal). Nor do they spend much time considering whether people are responsible in the way that their theories would require. But both of these things would seem to be essential to the plausibility of their theories. If we are not equal in many respects, materially or legally, or if our social conditions mean our acts do not give us an advantage, or express our superiority, then the theories of Morris and Hampton are *prima facie* implausible, however elegant and stirring they may be *qua* theories.

The retributivist commitment to formal equality shows up clearly in its support of mandatory sentencing. For retributivists, mandatory minimums would seem not to go far enough, for there really should be *no* variations in sentences for the same or similar crimes. Discretion is the enemy of equality, which is here understood as *formal* equality: treat everyone the same by punishing the same, regardless of differences in background circumstances. Even if the people committing the crime may be very different, their crime is the same, and (formal) like must be punished with (formal) like. On a purely retributive picture, it is hard to make sense of any differences in punishments for the same crime. For on what grounds could such variations be justified on their picture, if all that matters is the individual's decision to commit a certain crime?¹²⁸

Kant, at least, was more honest that his theory was not about the empirical world.¹²⁹ For Kant, this transcendence from the empirical was partly the point: we could still think of justice being done above the messiness of the real world, even if this justice took place on another plane. My sense is that most contemporary retributivists would reject Kant's metaphysical transcendentalism, but they do not engage with facts about the real world in a way that such a rejection would imply. They keep Kant's disdain for the world while discarding his overall world view, and end up at best with a theory that applies only in an ideal world.

The retributive closure to empirical facts comes across clearest when it comes to the reality that retributive justice costs money. Retribution tends to put things in terms of justice, which by its nature—by being an intrinsic good—resists assimilation to cost-benefit analysis. How do we weigh justice against considerations of cost? Again, we might say that the resistance to cost considerations is

128. For an excellent historical analysis of this idea of equality, see Whitman, *supra* note 5.

129. *Supra* Part I. Hegel seemed to follow Kant in this “unworldliness.”

part of the *point* of traditional retributive justice: justice must be done, even if the heavens fall, as Kant said. But this very resistance, which some retributivists may see as a feature of their theory, is actually a bug when it comes to the application of the theory. Either we have to say that cost is irrelevant when it comes to punishment, that it would be wrong even to consider the fact that criminal justice costs money, or we simply have to reject retribution as a plausible theory of criminal justice in the real world.¹³⁰ The former option means just ignoring the real world, and the expenses of justice; the latter option means leaving pure retributive theory behind.

2. *Compartmentalization*

The problem of closure is compounded when combined with what we can call the problem of compartmentalization. Retribution presents itself as a theory of punishment; no more, no less. It does not, or does not necessarily, say *what* things should be crimes. It does not say *how* people should be punished (this is the source of retribution's frustrating indeterminacy). It just says *why* people should be punished: they should be punished because they deserve it. Sometimes retributivists are explicit about the limits of their theorizing.

A theory that limits itself like retribution does, has accordingly limited its usefulness. It cannot have a positive theory of crime control, because *controlling* crime is different than *punishing* crime. Retribution kicks in only after a crime has been committed (however we define crime, which retribution also has nothing to say about¹³¹). So measures to reduce crime cannot be justified on retributive grounds. Indeed, retributive theory on its own terms doesn't seem even able to say why more crime is worse than less; it only says that more crime would mean more punishment. It is even possible that a retributive theory would lack the resources to say why a world with

130. A third option is possible, but it has so far been underspecified: that retributive justice is a good to be maximized, but also to be balanced with other goods. See Cahill *supra* note 8, at 861, 867 (“Any further exploration of just *how* good desert is relative to other consequences is, again, beyond the scope of the present Article . . .”).

131. A point I make in my response to Gray and Huber. See Chad Flanders, *Can Retributivism Be Progressive?: A Reply to Professor Gray and Jonathan Huber*, 70 MD. L. REV. 166 (2010); see also Shafer-Landau, *supra* note 53, at 293 (“These [retributivist] views presuppose the existence of a criminal code, rather than providing assistance in constructing one.”).

more crime *but which always punishes that crime* is intrinsically worse than a world where crime never occurs. If punishment restores the status quo ante—the balance of benefits and burdens, or the economy of esteem—how are we different or worse off than before?

Similarly, retribution's narrow focus on a justification of punishment makes it blind to phenomena such as mass incarceration. On a theory of retribution, provided that those punished are guilty of some crime, that they deserve to be punished, there is no problem in principle with mass incarceration. We need to look at each crime and each punishment in isolation, and not as a larger social phenomena, or as a question of political justice. If there is a just punishment for each crime taken by itself, then the fact that a large percentage of a nation, or even most of a nation, is incarcerated is itself unobjectionable. Retribution can say that if some groups are being punished more than others, this could be a problem, but retribution by itself cannot say whether this means *reducing* the punishments for that group or *increasing* the punishment for the other groups. In general, retribution *by itself* can't say whether we need to "level down" or "level up."¹³²

We might consider this self-limitation of retribution as having the virtue of theoretical clarity. Indeed, it has clearly defined ambitions, and it does not pretend to give guidance outside of those ambitions. It gives a theory of the basis of punishment. But this presents us with the question of whether it is fruitful to look at punishment this way, as in a sort of vacuum, separated from empirical questions about the world and about topics like crime prevention or the costs of punishment. The focus on the moment of punishment distorts the role punishment plays in a larger system of criminal justice. There is a story we need to tell about criminal justice, and the purpose of punishment has to fit in that story, somehow. It cannot be that we stop our ears to what is happening in other areas of criminal justice to focus on punishment, and say that this is necessary to understand the justification of punishment correctly. In fact, doing so makes our writing on punishment impractical and empty.

This isolation means, professionally, that punishment theorists do their work apart from those who study the causes of crime and ways of preventing crime, and even apart from those who try to determine

132. On the idea of leveling up and leveling down, see Chad Flanders, *Pardons and the Theory of the "Second-Best,"* 65 FLA. L. REV. 1559 (2013).

what things should be criminalized. This cannot be good, not as a practical matter, and not as a theoretical matter either. Retribution, in word, prevents us from thinking holistically about punishment, about punishment as part of a larger system, which includes why people become eligible for punishment, why they are punished, and what our responsibilities are to them after punishment. Punishment is not a discrete entity to be justified and explained in terms unique only to it.

C. Is Retribution a Theory of Punishment at All?

Rawls and Hart, in their early essays, both suggested that to think constructively about punishment, we need to start at the institutional level: why would a society feel the need to punish people? Punishment is one solution to a problem, they suggested, and that problem is crime. The problem is not, *pace* the retributivist, that the guilty need to be punished. It is not *functionally* useful for a society to punish only for the sake of fulfilling some transcendental moral balancing,¹³³ whether that balancing is of rights, or benefits or burdens, or of esteem. The idea that punishment is necessary for justice comes later, as a gloss on the institution of punishment. Certainly, if retribution ended up creating *more* crime, it would not long endure as a social institution.¹³⁴ In any event, Rawls and Hart say it is bad *now* to think of retribution as being a foundational purpose of the institution of punishment.

In part, Rawls and Hart thought this was just descriptively true. They looked at the purposes of punishment, what was doing the work in *why* we had punishment, and they saw utilitarian purposes, not retributive ones. But the burden of Parts I and II of my essay was to show that retribution can't do the work expected of it as a theory of punishment. That is, it is not just descriptively true that retribution is not foundational to punishment but normatively true as well. Retribution promises to tell us why people deserve to be punished, and how much punishment they deserve. But it falters. On its own terms, retribution cannot explain how we are to fix punishments and crimes; it cannot, that is, say *why* the rapist deserves *this* much punishment, and the robber *that* much. Even to explain

133. Hart famously referred to the “mysterious piece of moral alchemy” of many forms of retributivism. See H.L.A. HART, PUNISHMENT & RESPONSIBILITY 234 (1968).

134. A point helpfully made by Shafer-Landau, *supra* note 53, at 296–97.

why people deserve their punishments, either retribution resorts to intuition (“they just deserve it”) or makes implausible assumptions about our social world (that we are equal in benefits and burdens, or in esteem, or in rights and liberties) or simply smuggles in consequentialist concerns.

What has gone wrong here? The problem is that a theory that serves best to constrain utilitarianism is instead being offered as a stand-alone theory. The intuition that the innocent should not be punished is not sufficient to show that the guilty deserve to be, nor will the intuition that insists unequal sentences are wrong necessarily have any resources to say whether a punishment is “too much” or “too little.” There may be things that we are entitled to as persons, and these might be explained by retribution: that we cannot be punished if we are innocent or that we cannot be punished unfairly. But retribution cannot turn around and use these normative truths about us to stand in for a theory of punishment. When it does, things go wrong: retribution comes up with unappealing explanations or no explanations at all.

Either retribution is not a theory of punishment, or it is a weak one, a theory that cannot help us think meaningfully about our criminal justice system. It is, at best, a theory of why some sentences or some punishments are unjust. Or, more likely, it is simply a stand-in for other principles that should be considered not as part of a theory of punishment, but as justified in their own right: don’t punish the innocent, or don’t punish too harshly. These are often, collectively, referred to as ideas of “legality,” and we can endorse them without endorsing retribution: we can say that people ought to be given fair notice, that they should be punished equally *vis-à-vis* each other (and not against some imagined scale of punishments), that they should have due process before they are punished, etc. We can uphold all of these values, without pretending to build a theory of punishment around them.

The failure of retribution *qua* theory comes across clearest when we ask it to do any practical work, as we should expect any theory of punishment worth its salt to do. The fact is, it cannot. The “solutions” retributive theory provides are so indeterminate, so empty, that it becomes clear that retribution as a theory is trying to do a work it was never meant to do.

V. CONCLUSION: REVENGE, FAIRNESS, AND DESERT

Retributivists have put much labor into refuting the idea that retribution is the same as revenge.¹³⁵ Revenge, retributivists argue, can be disproportionate, can focus on others instead of the offender, and is a matter of feeling and not reason. But the fact that retribution does not share *all* of the features of revenge does not compel the conclusion that there is not at least a family resemblance between retribution and revenge that would justify putting the two in a similar category. For retribution and revenge do share one key, defining trait: they tend to be justified *for their own sakes* (as “intrinsic goods” as it were). Revenge needs no other warrant than the fact that a wrong has been done to me (or my family) and it must be responded to; to leave the wrong standing would be shameful, or humiliating, or embarrassing. Retribution, too, rests on nothing but itself for its warrant: at the end of the day, the wrong must be answered. To not answer the wrong would be unjust, or not fitting, or not fair. As Kant again so eloquently expressed, part of the point of retribution is that no utilitarian good needs to come out of it for it to be justified.¹³⁶ Even the last murderer on the island must be killed, because justice demands it. Nothing more but also nothing less.

It is *this* shared feature that I think most people are pointing to when they combine retribution and revenge: that purposeless punishment is justified. To be sure, there still is at root a difference between them: one is rooted in the desire of the individual for satisfaction, the other in the ostensibly neutral purposes of the state. But both revenge and retribution say that the suffering of the offender is justified even if no good comes of it. In this respect, I think it is fair to, if not equate the two, say that the two motivations for punishment share something very important in common: that suffering of another can be justified just because it is suffering (or some other harm) and for no other reason.¹³⁷

135. I have discussed the relationship between retribution and revenge briefly in Chad Flanders, *In Defense of Punishment Theory, and Contra Stephen: A Reply to DeGrolami*, 10 OHIO ST. J. CRIM. L. 243 (2012).

136. Kant, *Penal Law*, at 18.

137. Cf. Shafer-Landau on retributivism and “free-floating goods.” Shafer-Landau, *supra* note 53, at 298 (retributivism allows decrease in citizen liberty in the name of the free-floating good of meting out deserved punishment, where a “free-floating” good is one that does not promote any particular individual’s interest).

But this may lead to a final defense of desert as a basis for punishment, and the suggestion that retribution may have something right, after all. For if the retributivist will say that punishment can be justified even if no benefit comes from it (no deterrence, no channeling of bad emotions, no message to the public that such behavior is not approved, etc.), then when we abandoned desert aren't we saying that if a punishment serves no public purpose, then that punishment is not justified? Suppose that it would make no difference (implausible as this may sound) if a murderer was punished or not, that his non-punishment would never be found out. Imagine, that is, the converse case of the innocent person being punished, and imagine the guilty person (the rapist, the murderer, the robber) not being punished. Wouldn't it be unjust if such a thing happened?

This is indeed the best case for what we can call "positive retribution," the idea that desert does not merely provide a side constraint on not punishing the guilty, but pronounces a *duty* to punish those who have done wrong.¹³⁸ And here, I think, the person who believes, like Hart and Rawls, that the foundational justification for punishment is utilitarian—crime control—must bite the bullet and say: there is no injustice, that is, no political injustice that happens in such a case, provided it is adequately described. The case is adequately described if the only reason for punishing a person is that he deserves it, on whatever theory of metaphysical desert we prefer (whether the offender has a right to be punished, or must have the benefit of his crime removed, or must be brought low in his own eyes). All of these reasons are inappropriate grounds for state punishment, because they rest on theories of desert that are not political theories, in Rawls's usage of that terminology.

We might be tempted to say that it would be unfair if the murderer didn't get punishment for his crime (even if no one found out he wasn't punished), but this notion of fairness is quite different from the idea of desert. Fairness happens when you have a punishment system in place, and gives rules on how to operate that system. The idea that it would be unfair for us not to punish the murderer is different from saying that desert alone gives us the reason to punish him. Fairness doesn't give you a reason in itself for punishing. Fairness is a value that supervenes on a system of

138. See Cahill, *supra* note 8 for a sensitive development of this distinction.

punishment that exists for reasons other than fairness; it does not ground that system. It cannot.

Indeed, if we appeal to desert to ground our punishment of individuals (and not to fairness), we are left appealing to our personal convictions—however philosophical they may be—and so in a very real sense *we are on our own*. We are trying to make the state a vehicle for our own comprehensive doctrine. And here, again, may be another point of resemblance between revenge and retribution. Revenge is said to be a matter of the individual pursuing his own ends; retribution is, by contrast, a matter for the state. But if we interpret retribution as *more* than fair side constraints on punishment, then retribution puts us back in the private realm, not the realm of public reasons and justifications. We are back, that is, to the pre-political. We are back to individuals pursuing their own private ends.

When we look at retribution this way, we should have no trouble with leaving it behind as a stand-alone theory. It is a philosophy ill-suited for the modern, technical, bureaucratic state¹³⁹: one that needs to approach crime holistically, and with sensitivity to the empirical world, what is actually on the ground, and what actually works in making communities safer. Moreover, we can keep what was good with retribution, without buying into the entire package. Indeed, it was only the thought that *without* retribution we would become monsters (disproportionate punishment, punishing the innocent), that retribution has had its staying power. But retribution, despite its promises, never really had this power to stop us from becoming monsters. It has always been up to us. What retribution has sometimes prevented us from doing is dealing with our problems practically and humanely. It has promised us much, but given us little.

139. See Rachel E. Barkow, *The Ascent of the Administrative State and the Demise of Mercy*, 121 HARV. L. REV. 1332 (2007).