Retribution's Role

JOHN BRONSTEEN*

Two main types of principle, retributive and consequentialist, have long been identified as the main approaches to justifying criminal punishment. Retributivists deem punishment justified by the wrongdoing of the offender, whereas utilitarians deem it justified by its good consequences such as deterring future crime. Over the past fifty years, each has spent decades as the dominant theory, and many hybrid theories have also been advanced. But few, if any, of the hybrid approaches have valued heavily both retributive and consequentialist considerations while locating the particular justificatory role each category plays. This Article points in that direction by reframing the central question of punishment justification as two questions: Why does the state have a right to punish, and why does it choose to exercise that right? The first question is answered most naturally by retributive considerations, whereas the second identifies the most natural space for utilitarian values. This framing device, it is hoped, resolves some of the disputes between retributivists and utilitarians while sharpening the focus on those that remain.

INTRODUCTION

What gives a state the right to imprison a person? The simplest answer is that the person broke the law. Why does a state choose to exercise that right? The simplest answer is that if it did not, then both this individual and others would be more likely to break the law in the future.

Those simple intuitions attach, respectively, to the two main justifications of punishment: retributive and utilitarian. This Article aims to use these core intuitions to reconcile the two justifications and thereby to build a mixed theory of punishment. This theory identifies the appropriate roles for retribution and deterrence and uses each to shore up the limitations of the other.

What emerges is a model in which retributive considerations[1] determine who may be punished and also set the upper boundary of morally permissible punishment in each instance, whereas utilitarian considerations[2] direct how much punishment the state will choose to inflict on the offender beneath that upper boundary. This theory represents one of the first efforts to create a true integrated dualist account of punishment, assigning appropriate roles to retribution and utilitarianism rather than merely conjoining them or prioritizing one over the other.

---

* Assistant Professor, Loyola University Chicago School of Law. J.D., Yale Law School; A.B., Harvard University. I have profited from discussions of this topic with Michael Cahill, Owen Fiss, Brett Frischmann, Jonathan Masur, Stephanie Stern, and Spencer Waller, as well as the many faculty members at the Sandra Day O'Connor College of Law at Arizona State University who attended my presentation of the paper there in February 2008. Brian Witkowski provided outstanding research assistance.

1. In particular, these considerations involve the severity of the crime.

2. These considerations involve the increase in overall welfare that might be achieved through punishment due to deterrence and incapacitation, among other things.

I. SOME PRELIMINARIES

Since the time of Jeremy Bentham and Immanuel Kant two hundred years ago, two primary approaches to moral theory have persisted.[3] One approach, consequentialism, or its subsidiary, utilitarianism, holds that morality is defined by the consequences of one's actions or that increasing overall welfare generally equates to doing the right thing.[4] The other approach, deontology, defines morality independent of consequences and suggests that moral acts are done for their own sake rather than in order to achieve any particular end.[5]

This division of opinion pervades not only moral philosophy but also political and legal theory.[6] In recent years, the Kantian deontological position has been associated most closely with John Rawls,[7] while consequentialism has close ties with law and economics,[8] the school of thought whose ascendance represents perhaps the most noteworthy development in legal academia in the last half century.[9]

Among the many topics within legal scholarship, the one that has been affected most thoroughly by this division is substantive criminal law. Notwithstanding the very important recent work done by Tracey Meares, Dan Kahan, and others on empirical studies of crime and its prevention,[10] the main thread of scholarship in this area has

---

3. *See* Jean-Christophe Merle, *A Kantian Critique of Kant's Theory of Punishment*, 19 LAW & PHIL. 311, 311 (2000) ("Utilitarianism and deontological ethics have traditionally dominated the debate on the justification of punishment . . . .").

4. Kyron Huigens, *The Dead End of Deterrence, and Beyond*, 41 WM. & MARY L. REV. 943, 954–55 (2000) ("Under consequentialist ethics, punishment is justified by the deterrence of harm or, more broadly, by the promotion of social welfare."). Throughout this Article, I use the terms "consequentialist" and "utilitarian" more or less interchangeably, not because I fail to appreciate the possibility that a theory could be based upon consequences other than welfare maximization, but rather because the distinction is generally not relevant to the particular lines of reasoning I discuss.

5. Aaron J. Rappaport, *Rationalizing the Commission: The Philosophical Premises of the U.S. Sentencing Guidelines*, 52 EMORY L.J. 557, 568 (2003) ("The second great moral tradition is sometimes called a 'deontological' theory or, more commonly, a theory of 'retribution.' This moral tradition rejects the idea that punishment can be justified based on its consequences for society. Rather, and speaking roughly for now, punishment is justified based on some inherent moral quality of the act or actor himself.").

6. *See* Michael Philips, *The Justification of Punishment and the Justification of Political Authority*, 5 LAW & PHIL. 393, 393–94 (1986) ("[T]he theory of punishment is importantly connected with the theory of state authority.").

7. *See* William Powers, *On the Priority of Justice*, 63 TEX. L. REV. 1569, 1569 (1985) (reviewing MICHAEL J. SANDEL, LIBERALISM AND THE LIMITS OF JUSTICE (1982)) ("Deontological liberalism is a position associated with the works of Immanuel Kant and John Rawls . . . ."); *see also* JOHN RAWLS, A THEORY OF JUSTICE (1971); JOHN RAWLS, POLITICAL LIBERALISM (1993).

8. *See* Huigens, *supra* note 4, at 954–55 ("Deterrence theories of punishment have a strong affinity with the economic analysis of law, because neoclassical economics is a variety of consequentialism.").

9. *See* Geoffrey Miller, *Tribute to Judge Richard A. Posner*, 61 N.Y.U. ANN. SURV. AM. L. 13, 13 (2005) (stating that law and economics is "arguably, the most important development in legal scholarship in the last hundred years").

10. *See, e.g.*, TRACEY L. MEARES & DAN M. KAHAN, URGENT TIMES: POLICING AND RIGHTS IN INNER-CITY COMMUNITIES (1999); Tracey L. Meares, Neal Katyal & Dan M. Kahan,

long been the philosophical endeavor of justifying criminal punishment. That endeavor, in turn, has been characterized by two different approaches—retributivism and utilitarianism—that mirror those of Kant and Bentham in moral theory.[11]

Punishment has made scholars uneasy for centuries, for reasons that go beyond the danger of its misuse.[12] According to the classical social contract theory that has its roots in the work of Thomas Hobbes and John Locke,[13] a government derives its authority from the choice of its people to accept it, and the people make that choice because they will be safer and better off in civilization than in anarchy.[14] In short, government exists to improve the lives of its citizens.[15] Although a government might be aiming to do just that when it punishes (for example, by deterring future crime and thereby protecting the innocent from would-be predators), the most direct result of the act is to harm a citizen—the criminal herself.[16]

Most people believe that a state is justified in punishing those who commit crimes,[17] and virtually all legal theorists share this view. But for the theorist, explaining why this is so involves more difficulty than might be imagined, because when a state punishes, it is intentionally and directly hurting its people rather than helping them.[18] Why is this

---

Updating the Study of Punishment, 56 STAN. L. REV. 1171 (2004).

11. George P. Fletcher, *The Nature and Function of Criminal Theory*, 88 CAL. L. REV. 687, 689 ("Today, we take the debate between Kant's retributivism and Bentham's utilitarianism to be paradigmatic of the style of intellectual confrontation that has induced philosophical reflection on every factor bearing on the question whether we should punish a particular individual on a particular occasion.").

12. *See* Daniel McDermott, *The Permissibility of Punishment*, 20 LAW & PHIL. 403, 425 (2001) ("The problem of matching crime with punishment has occupied philosophers for centuries . . . .").

13. *See* THOMAS HOBBES, LEVIATHAN (Edwin Curley ed., Hackett Publ'g Co. 1994) (1668); JOHN LOCKE, SECOND TREATISE OF GOVERNMENT (C.B. Macpherson ed., Hackett Publ'g Co. 1980) (1690).

14. Laurie Stearns, Comment, *Copy Wrong: Plagiarism, Process, Property, and the Law*, 80 CAL. L. REV. 513, 540 (1992) ("Some political philosophers—most notably Hobbes, Locke, and Rousseau—have theorized that societies are based on an implicit 'social contract' in which people come together in communities to gain the benefits of safety, security, and support, and in exchange relinquish their freedom to behave exactly as they choose.").

15. *See* Kyron Huigens, *Virtue and Inculpation*, 108 HARV. L. REV. 1423, 1425 (1995) ("The law has a purpose, an end in view, which is to promote the greater good of humanity.").

16. McDermott, *supra* note 12, at 403 ("Punishment is morally troubling because it almost always causes human suffering. Indeed, some have argued that in order for a particular form of treatment to be considered punishment it *must* cause suffering." (emphasis in original)).

17. *See* Kyron Huigens, *The Jurisprudence of Punishment*, 48 WM. & MARY L. REV. 1793, 1797 (2007) ("It may be that conditioning punishment on individual desert produces optimal social welfare, simply because most people believe, rationally or not, that this is how punishment should be done, and because they would withhold necessary political support for the legal system if it were done otherwise.").

18. *See* Sharon Dolovich, *Legitimate Punishment in Liberal Democracy*, 7 BUFF. CRIM. L. REV. 307, 310 (2004) ("Any theory of state punishment in a liberal democracy must grapple with the problem of political legitimacy. The punishment of criminal offenders can involve the infliction of extended deprivations of liberty, ongoing hardship and humiliation, and even death.").

acceptable? Two different categories of answers have been given, and they divide legal scholarship to this day.[19]

The first category belongs to the utilitarians, who believe that punishment is justified because, in the aggregate, it increases people's welfare rather than diminishing it.[20] The criminal suffers, but everyone else gains from (among other things) the security produced by deterring crime and by physically preventing incarcerated offenders from recidivating.[21] The state exists to improve its citizens' lives, and it improves those lives when it punishes criminals.

The second category belongs to the retributivists, who believe that punishment is justified not because of its good consequences, but rather for its own sake.[22] To a retributivist, punishing an offender means righting a wrong, or at least refusing to stand idly by and accept that wrong.[23] Not only is it morally permissible for the state to punish, but it might well be morally required.[24] And this requirement is so regardless of whether the punishment will produce any deterrence or other tangible improvement in people's lives.[25] To imprison or otherwise punish a brutal rapist and murderer, for

---

19. *See* Steven Sverdlik, *Punishment*, 7 LAW & PHIL. 179, 179 (1988) ("The debate about the justification of punishment appears to have no end.").

20. Steven Eisenstat, *Revenge, Justice and Law: Recognizing the Victim's Desire for Vengeance as a Justification for Punishment*, 50 WAYNE L. REV. 1115, 1162 (2004) ("Utilitarians are forward looking; they countenance punishment only if a social good will come from it."); *see also* Erik Luna, *Punishment Theory, Holism, and the Procedural Conception of Restorative Justice*, 2003 UTAH L. REV. 205, 209 ("Whatever goal is espoused, utilitarian-based punishment is always forward-looking, seeking to reduce the intensity and gravity of crime in society. In other words, utilitarianism takes the position that 'bygones are bygones' and that future consequences should be the sole guide for sanctioning decisions." (quoting John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3, 5 (1955)).

21. Youngjae Lee, *The Constitutional Right Against Excessive Punishment*, 91 VA. L. REV. 677, 737–38 (2005) ("As is well-known, the utilitarian theory of punishment subscribes to the notion that the issue of whether someone should be punished, and in what way, should be considered in terms of the consequences that the punishment would bring for the overall good of the society. The purpose of punishment, under this view, is not to give each criminal what he or she deserves, but to deter future crimes, to incapacitate criminals by keeping them 'off the streets,' or to rehabilitate criminals so they would become better citizens.").

22. Mary Sigler, *Just Deserts, Prison Rape, and the Pleasing Fiction of Guideline Sentencing*, 38 ARIZ. ST. L.J. 561, 563 (2006) ("Retributivism, by contrast, is centrally concerned with the imposition of suffering in proportion to an offender's moral desert. On this view, punishment of the deserving is intrinsically good; its justification does not depend on any further positive consequences that punishment might be expected to produce.").

23. Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659, 1663 (1992) ("[R]etributive justice is concerned with wrongful actions from which such harms result. Although a punishment may sometimes involve the wrongdoer compensating her victim in some way, the purpose of punishment is not to compensate the person for the harm suffered, but 'to right the wrong.'" (emphasis omitted)).

24. *See* Huigens, *supra* note 4, at 953 ("Retributivism takes punishment to be required by a kind of fitness in retrospect—a matter of balance or proportion between past and future.").

25. Sigler, *supra* note 22, at 563 ("Although a retributivist will welcome the positive consequences that punishment may incidentally yield—for example, crime prevention or character reformation—they are not part of the justification for punishment. Thus, a 'retributivist punishes because, and only because, the offender deserves it.'" (quoting Michael S.

example, is the right thing to do, not because it will increase overall utility, but simply in and of itself.[26]

These alternatives have divided scholars.[27] Utilitarians are frustrated by the difficulty deontologists sometimes have in explaining a source of morality independent of consequences.[28] Moreover, utilitarians point to the centrality of deterrence in civilized life: If there were no punishment, then there would likely be far more crime, and society itself might collapse under the weight of that crime.[29] To ignore this fact or to accord it insufficient value is, in the view of the utilitarian, a decisive mistake.[30]

Retributivists, for their part, emphasize that good consequences or increased utility do not seem to capture the true reasons we find punishment morally acceptable or even required.[31] We tend to cringe at punishing the innocent even if we know that doing so under certain circumstances could increase utility.[32] And we typically feel that the decrease in happiness that punishment creates for a sadistic murderer is not a morally regrettable feature that we accept in order to improve the welfare of others, but rather something inextricably linked to the moral rightness of the punishment.[33]

Moore, *The Moral Worth of Retribution*, *in* PUNISHMENT AND REHABILITATION 94 (Jeffrie G. Murphy ed., 1995)) (emphasis omitted)).

26. *See* Paul G. Cassell, *Too Severe?: A Defense of the Federal Sentencing Guidelines (and a Critique of Federal Mandatory Minimums)*, 56 STAN. L. REV. 1017, 1021 (2004) ("Just punishment is a justification that is anchored in the past. As articulated by two leading proponents, a just punishment 'gives an offender what he or she deserves for a past crimes [sic] [as] a valuable end in itself and needs no further justification.'" (quoting Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 454 (1997)) (alteration in original)).

27. *See* Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 454 (1997) ("The arguments of [the retributivist and utilitarian views] are generally considered irreconcilable . . . .").

28. *Cf.* Merle, *supra* note 3, at 311 ("[I]t is with good reason that such a [deontological] theory has been suspected of relying more on private morality than on principles of right.").

29. *Cf.* Luna, *supra* note 20, at 221 ("More importantly, critics chastise retributivism as being wholly indifferent to the causes of crime and the consequences of punishment. Retribution largely ignores the causal roots of offending, preferring to place the full weight of condemnation on the offender—although most scholars recognize that 'the problem of crime cannot be simplified to the problem of the criminal.'"(quoting LESLIE T. WILKINS, CRIME AND MARKET FORCES 312 (1991))).

30. *See id.*

31. *See* Robinson & Darley, *supra* note 27, at 454 ("Punishment that gives an offender what he or she deserves for a past crime is a valuable end in itself and needs no further justification (such as a showing of a future benefit). This is typically referred to as the 'retributivist' or 'just deserts' view.").

32. Norval Morris, *The Future of Imprisonment: Toward a Punitive Philosophy*, 72 MICH. L. REV. 1161, 1173 (1974) ("To use the innocent as a vehicle for general deterrence would be seen by all as unjust, although it need not be ineffective if the innocence of the punished is concealed from the threatened group.").

33. *See* John M. Darley, Kevin M. Carlsmith & Paul H. Robinson, *Incapacitation and Just Deserts as Motives for Punishment*, 24 LAW & HUM. BEHAV. 659, 660 (2000) ("In the late 18th century, Immanual [sic] Kant argued that . . . punishment ought to be 'pronounced over all criminals proportionate to their internal wickedness.'" (quoting Immanuel Kant, *The Science of Right*, *reprinted in* 42 GREAT BOOKS OF THE WESTERN WORLD 402, 447 (W. Hastie trans., 1st ed. 1952))).

In response to this divide between competing justifications, many scholars have advocated mixed or hybrid theories that find a place for both retributive and utilitarian concerns.[34] These attempts themselves have generally fallen into two categories: "hierarchical accounts" that are barely mixed at all in that they assign near-absolute primacy to one of the two values, and the "mere conjunction approach" holding that both values are necessary to justify punishment.[35] The hierarchical accounts have, understandably, failed to satisfy adherents of the opposing camp;[36] whereas the conjunction approach has received even less support, recently being deemed "ad hoc."[37]

Some view these failures as evidence that only a pure theory can ultimately make a claim to truth,[38] or at least that the current scholarly environment will be unreceptive to compromise.[39] But a less pessimistic and more plausible view is that a refined approach to the hybrid theory is needed.[40] The holy grail would seem to be, in the words of Mitchell Berman, an "'integrated dualist' account" that "assigns distinct and coequal roles to retributivist and consequentialist reasoning within the theory's structural logic."[41] But apart from Berman's own attempt, discussed in Part II.A.3, "no such account has yet appeared, and the possibility of such an appearance has been doubted."[42]

This Article suggests a new path toward an integrated dualist account. It first explores the reasons that the two major ideologies have each attracted such ardent support,[43] and then it goes beyond that exploration to propose a new way of looking at the problem.[44] Specifically, it suggests that the standard question of what justifies

---

34. *See, e.g.*, Lawrence Crocker, *The Upper Limit of Just Punishment*, 41 EMORY L.J. 1059, 1062 & n.8 (1992).

35. Mitchell N. Berman, *Punishment and Justification*, 118 ETHICS 258, 258–59, 287 (2008).

36. *Id.* at 259.

37. *Id.*

38. *See, e.g.*, Luna, *supra* note 20, at 242 ("In a winner-take-all philosophical world, it is either utilitarianism or retributivism—there is no space for compromise on the theoretical high-ground. Mixed theories are mere pretenders to the throne, with a close analysis revealing a half-hearted allegiance to one or the other theory, thereby exposing the hybrid to standard criticism by the opposing camp and ostracism from the pure theory as an ill-conceived child.").

39. *See id.* at 205 ("Punishment theories brutalize one another, staking out turf on principle and refusing to budge from their respective positions. As a result, the various theoretical camps spend most of their time on three endeavors: demonstrating the superiority of their approach to criminal sanctioning, subjecting all other theories to harsh criticism, and repairing the damage done to their own theory from equally severe attacks. The upshot is an unwinnable war of critiques within an ethos of mutual exclusivity: It is either one theory or another, but certainly not both.").

40. *See* Philips, *supra* note 6, at 394 ("[A] good deal of the discussion of punishment consists in exchanges of intuitions between deterrence theorists and retributive theorists that leave each side unconvinced.").

41. Berman, *supra* note 35, at 259.

42. *Id.*

43. *See infra* Part II.

44. *See infra* Part III.

criminal punishment has itself created confusion.[45] That question may profitably be unpacked to reveal two separate internal points that are typically conflated by punishment theorists:

(1) Why does the state have the right to punish?
(2) Why does the state choose to exercise that right?

I would like to suggest that the answer to the first question is supplied by retributive, backward-looking principles; that is, the state has the right to punish, not because of the future consequences of punishment, but because the criminal forfeited certain rights by breaking the law.[46] And I believe that the answer to the second question is supplied in significant part, at least, by utilitarian principles; that is, the state chooses to exercise its right to punish criminals because by doing so, it increases the overall welfare of its citizens in many ways (in particular, by reducing crime via deterrence and incapacitation).[47]

The distinction I propose is not meant to bring to an end all of the disputes surrounding the retributive/utilitarian divide in punishment theory. As I will discuss in Part IV, many important questions remain. However, the distinction helps make sense of the entrenched attachment of scholars to each side of the dispute, and it helps bring into focus the specific issues on which their disagreements can and cannot be reconciled.[48] Most important, it tracks our intuitions and considered judgments regarding the roles that retribution and utilitarianism play in justifying criminal punishment.

These two theories have been durable because each has much to offer in explaining punishment's place in society.[49] My hope is that the distinction suggested here will help shed light on the specific work each theory does in telling us why we accept and use punishment in response to law-breaking.

---

45. *See infra* Part II.A.

46. *See* Richard Dagger, *Playing Fair with Punishment*, 103 ETHICS 473, 476 (1993) ("Punishment is justified, ceteris paribus, because the persons who disobey the law fail to meet their obligations to the other members of society.").

47. *See* Kevin M. Carlsmith, John M. Darley & Paul H. Robinson, *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERSONALITY & SOC. PSYCHOL. 284, 284 (2002) ("[The utilitarian perspective] holds that social harmony is best served by the prevention of future harm and that the justification for punishment lies in its ability to minimize the likelihood of future transgressions.").

48. *See* H.L.A. HART, *Prolegomenon to the Principles of Punishment*, *in* PUNISHMENT AND RESPONSIBILITY 1, 3 (1968) ("[W]e should bear in mind that in this, as in most other social institutions, the pursuit of one aim may be qualified by or provide an opportunity, not to be missed, for the pursuit of others.").

49. Paul H. Robinson, *The Role of Moral Philosophers in the Competition Between Deontological and Empirical Desert*, 48 WM. & MARY L. REV. 1831, 1832 (2007) (stating that each side of the debate on punishment justification theory "has plausible arguments to make").

II. TWO APPROACHES TO JUSTIFICATION: THE NATURE AND ROOTS OF RETRIBUTION AND CONSEQUENTIALISM

The gap that this Article seeks to fill has been fueled by two sources. The first is a terminological confusion. Retributivists and utilitarians strive to demonstrate that their theories justify criminal punishment, but the question of justification must itself be divided to reveal the relevance of each theory. The second source of the problem is that each of the two main theories has points in its favor so strong as to prevent either one from being able to sway the adherents of the other.

Let us consider those two causes of the current impasse in the literature on punishment. Considering the first (that is, the terminological) problem will create the opportunity to explore several important hybrid theories, whereas considering the second problem will introduce us to the pure theories.

A. Hybrid Theories and the Terminological Problem

As noted in the Introduction, many scholars have tried to bridge the retributive/consequentialist divide by introducing hybrid theories. But these theories typically elevate one of the two values over the other to an extent that effectively denies the value of the less-favored approach. Such theories then face the problems associated with the pure theories described in Part I.B. A second type of theory acknowledges the value of both forms and conjoins them by making each a necessary, but alone an insufficient, condition for punishment. A more satisfying approach would be a third way—one that locates the appropriate role of each theory in the enterprise of justifying punishment.[50]

The following three subparts address leading examples of each category of hybrid theory. Subpart 1 considers the most famous of all such theories, that expressed by H.L.A. Hart—a theory that prioritizes consequentialism over retribution. Subpart 2 considers two major conjunction theories that bear some relationship to the integrated dualist view I will advance in Part III. And Subpart 3 considers a forthcoming attempt to achieve the integrated dualist goal via an entirely different mechanism.

In my view, solving the terminological problem that I will explain in the context of discussing Hart's theory is the linchpin to finding the relevant roles for retribution and consequentialism.

1. H.L.A. Hart and the Hierarchical Mixed Theory

The central question in punishment theory is generally styled as follows: What is the justification of criminal punishment?[51] This formulation invites wrangling between the two comprehensive theories, each claiming that it is the answer, but the most famous

---

50. Berman, *supra* note 35, at 259.

51. TED HONDERICH, PUNISHMENT: THE SUPPOSED JUSTIFICATIONS 142–43 (1969) ("Our fundamental and final concern, obviously, must be with the question of what justification can be given for the practice of punishment taken as a whole. It seems to me clear, although it is sometimes doubted, that this is a satisfactory formulation of the problem which *all* theories of punishment attempt to answer." (emphasis in original)).

attempt to resolve that problem by a terminological innovation has unfortunately resulted in more confusion than did even the general question itself.

In 1968, H.L.A. Hart published *Punishment and Responsibility*, in which he proposed a way to separate the central question of justification into subsidiary questions that would resolve the conflict between retribution and utilitarianism.[52] His proposal was as follows: one question would ask what the general justifying aim of punishment is, and the other would ask what principles guide the distribution of punishment.[53]

This methodology has been so influential that it is repeated to this day, even though it has failed to achieve its purpose of resolving the disputes between retributivists and utilitarians.[54] Hart's suggestion was motivated by his commitment to utilitarianism, which he identified as the answer to his first question—the general justifying aim of punishment.[55] Indeed, the two questions Hart proposed were designed to address the key drawback facing utilitarian theories of punishment, namely that those theories call for punishment of the innocent when that would increase overall utility.[56] By naming retribution as the principle governing "distribution"—that is, who can be punished—Hart's approach confined punishment to the guilty while still labeling utilitarianism as the "general justifying aim" of punishment.[57]

As clever a move as this was, it ultimately failed to resolve the central issue within punishment theory. In order to achieve such an advancement, a new terminology or distinction must take into account why each camp feels so strongly that it is right, and then use the strengths of each one in tandem to answer the question of justification.[58]

---

52. *See* HART, *supra* note 48, at 9.

53. *Id.*

54. *See* Sverdlik, *supra* note 19, at 180 ("In many ways the contemporary discussion of punishment centers on the well-known essay of H.L.A. Hart, 'Prolegomenon to the Principles of Punishment.'"); *see also* Dagger, *supra* note 46, at 474 ("If we want to provide a justification for legal punishment, then, we must answer two distinct questions: (1) What justifies punishment *as a social practice?* and (2) What justifies *punishing particular persons?*" (emphasis in original)).

55. *See* Russell L. Christopher, *Deterring Retributivism: The Injustice of "Just" Punishment*, 96 NW. U. L. REV. 843, 868 (2002) ("For Hart, the General Justifying Aim of our institutional practice of punishment is the beneficial consequences that punishment generates.").

56. *See* HART, *supra* note 48, at 9 ("Much confusing shadow-fighting between utilitarians and their opponents may be avoided if it is recognized that it is perfectly consistent to assert *both* that the General Justifying Aim of the practice of punishment is its beneficial consequences *and* that the pursuit of this General Aim should be qualified or restricted out of deference to principles of Distribution which require that punishment should be only of an offender for an offence." (emphasis in original)).

57. Christopher, *supra* note 55, at 868–69 ("As to who should be punished, the Distribution question should be answered by retributivism—'only an offender for an offence.' Under Hart's mixed theory, the justification for why punishment may be imposed in general is answered by the aim of deterrence and the question of who should be punished is answered partly by deterrence and partly by negative retributivism. An offender may only be punished if he is guilty and if deterrence could be promoted." (quoting HART, *supra* note 48, at 8)).

58. *See* IGOR PRIMORATZ, JUSTIFYING LEGAL PUNISHMENT 112 (1989) ("When philosophers set out to bridge a gap, to combine theories which seem to be mutually opposed and even irreconcilable, usually the first thing they do is introduce a distinction. This has been the case in

Only then will adherents of either approach find reason to accept the proffered middle ground. Hart's suggestion, though, offered no part of the "general justification" to the retributivists and merely attempted by fiat to exclude their central argument.

Put another way, Hart offered no meaningful defense of a central assumption within his theory: that principles of distribution are not inherently a part of the general justification of punishment. Whatever justifies punishment would seem, almost by definition, to dictate the circumstances under which a person may justifiably be punished. If utilitarianism cannot dictate those circumstances, as Hart's distinction suggests, then in what meaningful sense would it be a justification? It could supply the "aim" of punishment, but to equate aim with justification is to assume a utilitarian answer rather than to prove one.

More generally, the term "general justifying aim" merely replaces the original term "justification" that Hart was trying to unpack. Too much is encompassed within "general justifying aim," because rather than break down justification into component parts (some with retributive roots, and others with utilitarian ones), the term leaves justification intact and essentially defines it as being involved with aims (that is, ends or consequences) rather than with retributive considerations.

Thus, despite Hart's efforts, the debate between retributivists and utilitarians has persisted in the forty years since his book was published.[59] But this persistence cannot be attributed solely to the terminological confusion that resulted from Hart's work and from the original problem of a single-question justification-framing device that Hart set out to improve. It derives also from the unsurprising fact that both retribution and utilitarianism have much to recommend them, as we will see both in the following subsections and especially in the discussion of pure theories.

2. Lawrence Crocker, Norval Morris, and the Conjunction Theory

Dissatisfied with hierarchical approaches that make mere gestures toward including both major values, some scholars have offered mixed theories that make both retributive and consequentialist considerations necessary for the justification of punishment. Two important examples that bear strong similarity to each other,[60] and a weaker but still noteworthy similarity to my dualist account,[61] are those advanced by Lawrence Crocker and Norval Morris.[62]

---

all important attempts at a middle-of-the-road philosophy of punishment. All the attempts have proceeded from a distinction supposedly overlooked, or at least not fully appreciated, in the preceding debate . . . .").

59. Dagger, *supra* note 46, at 473 (stating that "[p]unishment remains the center of an equally lively debate" even after Hart's work).

60. Crocker, *supra* note 34, at 1062 ("[My theory] is very similar to the theory of punishment long championed by Norval Morris.").

61. *See infra* Part III.

62. NORVAL MORRIS, MADNESS AND THE CRIMINAL LAW 149–60 (1982) [hereinafter MORRIS, MADNESS]; NORVAL MORRIS, THE FUTURE OF IMPRISONMENT 58–84 (1974) [hereinafter MORRIS, IMPRISONMENT]; Crocker, *supra* note 34; Norval Morris, *Punishment, Desert and Rehabilitation*, *in* EQUAL JUSTICE UNDER LAW: U.S. DEP'T OF JUSTICE BICENTENNIAL LECTURE SERIES 136, 139–40 (1977) [hereinafter Morris, *Punishment*].

Crocker and Morris argue that retributive considerations set an upper (and, at least in the case of Morris, lower) limit of acceptable punishment,[63] and that below that upper limit, punishment is justified if it also satisfies utilitarian requirements.[64] After an upper bound has been set according to the severity of the crime committed,

> the state must still clear additional hurdles before any penalty may be handed down. In particular, the state must determine that the penalty will serve the purpose of crime control—maximizing social benefits—after subtracting its costs, including costs to the offender, or, alternatively, that the penalty will serve the purpose of minimum adequate denunciation.[65]

Crocker and Morris differ as to the method of determining the upper limit applicable to a given crime (with Morris rejecting the emphasis on equivalence between crime and punishment that Crocker favors),[66] but that divergence is immaterial for our purposes in reviewing their theories.

As will be explained in Part III, I agree with certain core features of this approach. Specifically, I take the view that a rights-forfeiture conception of retribution should determine the upper bound for punishment, and that consequentialist considerations become relevant in choosing the appropriate punishment below that limit. Two features, however, differentiate my approach from the approaches of Crocker and Morris. First, I aim for integrated dualism rather than mere conjunction by offering a new linguistic account of why retribution and consequentialism should play these respective roles. And second, as a result of the form of integration I have chosen, I conclude that punishment can be morally permissible even if consequentialist requirements are not fulfilled.

Notwithstanding its many valuable and impressive elements, the main feature of Morris's and Crocker's work that is unsatisfying is the absence of an attempt to explain *why* retribution and utilitarianism should play the roles those authors designate for them. I seek to offer such an explanation, and the explanation leads me to a different substantive conclusion.

---

63. Morris, *Punishment*, *supra* note 62, at 140 ("No sanction should be imposed greater than that which is 'deserved' . . . . Nor should a sanction be imposed which is so lenient that it unduly depreciates the seriousness of the crime.").

64. *See, e.g.*, MORRIS, IMPRISONMENT, *supra* note 62, at 60; Crocker, *supra* note 34, at 1078.

65. Crocker, *supra* note 34, at 1078.

66. *Id.* at 1062 n.4 ("[T]his Article's treatment of the key matter of the upper limit is in considerable conflict with that of Morris."). *Compare id.* at 1065 (arguing that "the offender forfeits immunity to criminal punishment in an amount" equal to that which "the offender has imposed upon society," and that "[a]ny penalty that is too severe to be countenanced under [this equivalence approach] is unjustly severe—however laudable its consequences"), *with* MORRIS, MADNESS, *supra* note 62, at 179 ("'[A] deserved punishment' . . . does not mean the infliction on the criminal offender of a pain precisely equivalent to that which he has inflicted on his victim; it means rather a 'not undeserved punishment which bears a proportional relationship in a hierarchy of punishments to the harm for which the criminal has been convicted.'").

3. Mitchell Berman's Version of Integrated Dualism

Mitchell Berman has just published a paper that sets forth an "integrated dualist" theory of punishment—that is, a theory that makes use of both retribution and consequentialism in a non-hierarchical form and that identifies distinct roles for each to play.[67] My view is that integrated dualism is the appropriate goal for punishment theory, though I am not persuaded by the particular account Berman advances.

Berman divides punishment into two categories: that which is inflicted on actual offenders, and that which the state mistakenly inflicts upon those it believes to be guilty, but who in fact are not guilty.[68] Cases within the first category are the "core" of punishment, whereas those within the second are the "periphery."[69] Berman does not address a third group of cases, those involving knowing infliction of punishment on non-offenders, because he assumes that those instances of punishment are not justified.[70]

For Berman, retribution justifies punishment in core cases, and consequentialism justifies it in peripheral cases.[71] His reasoning is that retribution involves the cancellation of rights: one's right not to be punished is cancelled by the crime one has committed. By contrast, consequentialism involves the override of rights: one's right not to be punished is overridden by the value to be achieved by the punishment.[72]

On this view, punishing actual offenders is retributive because their wrongdoing reshapes what is required for them to be treated with respect by the state.[73] It cancels their right to be free from punishment. By contrast, erroneously punishing those thought to be guilty is consequentialist because although the offenders retain their full right to be free from punishment, that right is overridden by the utilitarian considerations to be achieved. The argument depends on the further contention that "whether conduct infringes a right . . . can be belief dependent. . . . In contrast, whether a state of affairs is intrinsically bad or intrinsically good (or intrinsically neutral) is belief independent."[74]

Without giving a detailed response to this line of reasoning, I will merely express some general reservations. First, it is not clear to me that punishing the innocent, albeit in good faith, is justified. Second, if it is justified, then I would not characterize the reason for its justification as consequentialist. Here Berman's choice not to discuss "degenerate" cases (knowing punishment of the innocent) becomes relevant because if such cases are unjustified whereas good-faith errors *are* justified, the distinction would seemingly have to be based on non-consequentialist reasons. That is, knowingly punishing the innocent could have consequences equally good as mistakenly punishing

---

67. Berman, *supra* note 35, at 259–60.

68. *Id.* at 261–62.

69. *Id.*

70. *Id.* at 262 & n.14 ("Eliminating degenerate cases from the scope of 'punishment' is a stipulation designed to simplify exposition, on the assumption that most commentators are agreed that [punishment in these cases] is not justifiable.").

71. *Id.* at 281–84.

72. *Id.*

73. This reshaping is conceived as distinct from traditional rights-forfeiture accounts. *See id.*

74. *Id.* at 283.

the innocent. The meaningful difference would thus have to be retributive, not consequentialist.

In short, Berman should be credited for his taxonomy of punishment theories and for his attempt to assemble an integrated dualist theory. But the particular method of integration he chooses is, in my view, unavailing.

B. The Persuasive Force of Both Theories and Their Parent Doctrines in Moral Philosophy

As important as the hybrid theories have been, they emerged mainly as a reaction to the entrenched attachment of many scholars to each of the two pure theories. Pure retributivism and consequentialism each carry considerable explanatory power, which explains their durability. They also find themselves situated within larger philosophical traditions that have been even more durable. Let us consider the reasons for the success of these theories within the context of punishment, as well as in the larger realm of moral theory.

1. Punishment Theory

i. Utilitarianism

The utilitarian theory of punishment has considerable persuasive force. What could more simply or elegantly explain why the state is warranted in inflicting harm than theorizing that by doing so the state decreases harm in the aggregate?[75] The heft of this viewpoint is demonstrated further by the fact that it taps into perhaps the primary reason for forming a society at all: security.[76] Without government or law, one's de jure unchecked freedom is outweighed by the de facto unchecked dangers. Society creates power within the government to protect citizens from their neighbors, mainly by punishing those who harm others and thereby deterring future wrongdoing.[77]

Not everyone agrees about what society would look like if the state did not punish criminals, but it would be difficult to dispute that at a minimum, crime would be committed more often due to the decreased deterrent. The vacuum might well be filled by retaliatory acts of vigilantism, with escalating violence back and forth. Few would choose to live that way if given a reasonable alternative, and state-sanctioned punishment, despite its many problems, has done a decent job as that alternative.

---

75. *See* Carl Emigholz, Note, *Utilitarianism, Retributivism and the White Collar-Drug Crime Sentencing Disparity: Toward a Unified Theory of Enforcement*, 58 RUTGERS L. REV. 583, 599 (2006) ("[U]tilitarian punishment theory seeks to minimize pain (societal costs) while maximizing pleasure (societal benefits), and rationalizes the behavior of would-be criminals.").

76. *See* Michael Tonry, *The Functions of Sentencing and Sentencing Reform*, 58 STAN. L. REV. 37, 52 (2005) ("The preventive functions center on the collective interest in domestic tranquility, on enabling citizens to get on with their lives in security. These are core functions of the state and basic goals of the criminal law and punishment.").

77. Dagger, *supra* note 46, at 475 ("With the aid of the institution of punishment, however, we can provide a guarantee that 'those who would voluntarily obey shall not be sacrificed to those who would not.'" (quoting H.L.A. HART, THE CONCEPT OF LAW 198 (1961))).

These facts are a cornerstone of the utilitarian theory of punishment. A utilitarian asks how the primary justification of punishment could be anything other than the need to protect the innocent and discourage lawbreaking, when it seems so plain that these are central reasons we give the state the power to punish. Any theory that seeks to justify punishment will fail to garner support among utilitarians unless it accounts for this point.[78]

ii. Retribution

Like the utilitarian theory, the retributive theory has much to recommend it. The retributive approach makes sense of our desire to punish only those who have committed crimes, and it also identifies what seems to be a deep-seated human desire to right wrongs and to stand up for victims.

Perhaps the most important point in favor of retribution—the point that has anchored the theory since its inception—is that it is far more reliable than is utilitarianism at explaining why punishment should be confined to those who break the law.[79] A utilitarian is committed to accepting whatever increases overall welfare or happiness and therefore cannot rule out punishing the innocent in cases where doing so would improve lives in the aggregate.[80] The examples are almost too familiar to repeat, so suffice it to say that in certain circumstances (especially those in which the innocence of the accused is known to the authorities but can be concealed from the public) it could increase utility to imprison or perhaps even to execute an innocent person.[81]

Several responses are available to utilitarians, but none have proven particularly satisfying. The simplest response is to accept that it is justified to punish the innocent when doing so increases overall welfare.[82] Some dyed-in-the-wool utilitarians will have

---

78. *See* James Q. Whitman, *A Plea Against Retributivism*, 7 BUFF. CRIM. L. REV. 85, 89 (2003) ("But [retributivism] does often seem weirdly blind to the nasty realities of the American world around it, with its otherworldly discussions of abstractly conceived autonomous actors.").

79. Gerard V. Bradley, *Retribution: The Central Aim of Punishment*, 27 HARV. J.L. & PUB. POL'Y 19, 27–28 (2003) ("Where retribution forms the moral justification for punishment, the problem of punishing the innocent is resolved simply and satisfactorily. Other theories cannot answer this conundrum so readily.").

80. *See* John Steele, *A Seal Pressed in the Hot Wax of Vengeance: A Girardian Understanding of Expressive Punishment*, 16 J.L. & RELIGION 35, 36 (2001) ("Consequentialist accounts have their own problems. If we abandon the search for a moral fit between crime and punishment, and instead measure out hard treatment by calculating the positive effects flowing from a particular punishment, we may entertain the punishment of the innocent and the cruel and excessive punishment of the guilty."); *see also* Huigens, *supra* note 17, at 1797 ("'The question is whether utilitarian arguments may be found to justify institutions . . . such as one would find cruel and arbitrary.'" (quoting John Rawls, *Two Concepts of Punishment*, 64 PHIL. REV. 3, 10 (1955))).

81. Luna, *supra* note 20, at 215 ("Utilitarianism does not reject outright the concept of punishing the blameless. If imprisoning or even executing an innocent person increases social utility through heightened feelings of public security, for instance, there is nothing about utilitarianism that would prevent the creation of a scapegoat." (citation omitted)).

82. Sverdlik, *supra* note 19, at 193 ("Utilitarians sometimes offer a counterthrust to the criticism that their theory would (or could) lead to the punishment of the innocent. They ask

no problem with this and indeed will hold that to do so is not merely warranted but also morally obligatory. But most of us feel differently. We have a powerful aversion to punishing the innocent and are generally willing to accept a decrease in utility in order to avoid such punishment.[83] A committed utilitarian might find this view frustrating and insist that it is soft-headed or otherwise misguided—that if we thought about it hard enough and rigorously enough, we would come to understand our mistake. The problem is that many intelligent people have been thinking about this in sophisticated ways for decades and are still unwilling to let go of the claim that punishing the innocent is wrong. Waiting for the majority to change its mind on this question does not seem promising.

A more sophisticated utilitarian response is that having a rule against punishing the innocent is valid for the very reason that it increases utility.[84] Deciding whether to punish an innocent in each instance could lead to instability, abuse of power, and diminished faith in the legal system. Because we cannot be confident in our ability to confine such punishment to the rare cases in which it would actually be beneficial, and because an absolute bar would make people accept the rule of law, utility dictates that we adopt such a bar. This sort of approach, with its lineage traceable to the work of John Stuart Mill,[85] is more viscerally appealing than is the direct acceptance of punishing the innocent when convenient.

But the rule-utilitarian approach is not a panacea. A rule-utilitarian must say that although it would be desirable if we could punish the innocent in those individual instances in which doing so would improve the aggregate welfare, we choose to prevent ourselves from having that choice for the long-term greater good.[86] This line of reasoning, though internally consistent, is diametrically opposed to the view most of us take of punishing the innocent in such an individual case. We do not view our inability to do so as a regrettable missed opportunity, nor do we lament that we must accept a second-best solution in choosing a blanket prohibition against this sort of punishment. We oppose punishing the innocent not out of a desire to increase long-term utility, but because we believe such punishment is unfair.[87]

---

what is so wrong, after all, in such punishment? Life in our society often involves placing burdens on innocent people. . . . The utilitarian, in other words, can ask his or her critics to specify what is wrong with punishing the innocent that would not also make, for example, taxation for public purposes wrong.").

83. HYMAN GROSS, A THEORY OF CRIMINAL JUSTICE 414 (1979) ("Condemning one who is blameless is universally abhorred as an injustice . . . .").

84. *Cf.* Lloyd L. Weinreb, *Desert, Punishment, and Criminal Responsibility*, 49 LAW & CONTEMP. PROBS. 47, 47 (1986) ("[U]tilitarians proffer elaborate arguments to show the instrumental value of punishing the guilty and only them; but tacitly or even explicitly, the real proof of the correctness of their arguments is that the results are consistent with considerations of desert, which is enough to undermine the conclusion the arguments are intended to support.").

85. *See* JOHN STUART MILL, UTILITARIANISM (George Sher ed., Hackett Publ'g 2001) (1863).

86. *Cf.* Daniel R. Williams, Commentary, *The Futile Debate over the Morality of the Death Penalty: A Critical Commentary on the Steiker and Sunstein-Vermeule Debate*, 10 LEWIS & CLARK L. REV. 625, 627 (2006) (defining a rule-utilitarian as "a person who focuses on the utility of adhering consistently to a rule rather than evaluating a particular act through that act's perceived utility in maximizing the aggregate good").

87. Eric Blumenson, *The Challenge of a Global Standard of Justice: Peace, Pluralism, and*

This last point is also the reason that utilitarians fail to persuade when they note that most of us accept punishment of the innocent at the margins. If one innocent person needed to be punished to save ten million lives, then we would accept the unfairness in return for the enormous utility advantage. It is possible to view this as evidence that we are all really utilitarians—that we reject punishing the innocent in most cases simply because in most cases it really would not improve overall welfare. When it is clear that it would, we accept it.

But once again, this sort of argument ignores the true power of the retributive position. It is not just that most people are against punishing the blameless, but also that *the reason they are against it has nothing to do with utility*.[88] Although we accept that in extreme cases, utilitarian considerations can trump fairness, it does not follow that we care only about utility and not about fairness. Even if we were willing to punish an innocent for the sake of saving millions of lives, we would still find it regrettable and disturbing that the act was done. By contrast, we would have no regrets if the person being punished were guilty,[89] even though the harm to that person of being punished would be similar to the harm imposed on an innocent target of punishment.[90] The difference is that we feel the guilty person deserves punishment, or at least that punishing him is not unfair in the way that punishing a blameless person would be.

This point is, in my view, a most serious problem for the utilitarian punishment theorist who denies that retributive considerations have any meaningful value independent of utility. It seems simply to be a fact that the vast majority of people find it repugnant to punish the innocent, and that their repugnance owes itself principally to considerations of fairness rather than of aggregate welfare. A utilitarian might regret that the rest of us take this view, but the prospect of convincing us to change our minds is bleak.

iii. Attachment to Each Theory

In light of these arguments for utilitarianism and retribution, it is unsurprising that each has such unwavering adherents. It is unreasonable to expect a utilitarian to relinquish the view that punishment is used, more than for any other reason, to enforce the rule of law and thereby improve people's lives.[91] A utilitarian might well hold the opinion that other reasons for punishment such as righting a wrong or standing up for a

---

Punishment at the International Criminal Court, 44 COLUM. J. TRANSNAT'L L. 801, 835 (2006) (stating that there is an "overwhelming moral truth that it is unjust to punish the innocent").

88. *See* Huigens, *supra* note 4, at 945 ("Fault—also known as desert, culpability, or blameworthiness—is the distinctive feature of the criminal law, but consequentialism has no independently viable conception of fault." (citations omitted)).

89. We may assume for the sake of the example that the punishment does not seem unfairly severe or disproportionate to the crime.

90. It is possible that the harm to the innocent would be greater because her innocence itself would make the punishment harder to bear. But this does not undermine the point. The main reason we would rather punish the guilty than the innocent is not that the innocent might experience more disutility from the punishment, but rather that the innocent does not deserve to be punished.

91. *See* Christopher, *supra* note 55, at 848 ("Consequentialist theories justify punishment not on the desert due the offender but on the actual, good consequences that are attained, for example, deterrence of crime, incapacitation of the offender, and rehabilitation of the offender.").

victim are thin veneers for simple vengeance.[92] Along these lines, it is humane and even noble to resist punishment for any reason other than to improve people's lives, and to deem such improvement as the source of punishment's legitimacy.

Similarly, it is unreasonable to expect a retributivist to give up the idea that punishing the innocent is wrong for reasons independent of utility.[93] Furthermore, the retributivist would be on solid ground in emphasizing that the main reason we feel justified in punishing the guilty is that they committed crimes, not that we can achieve some ulterior social purpose by doing so.[94]

Pure punishment theories in modern legal academia have been marked by a protracted struggle to accomplish the impossible: (a) to persuade the utilitarian to let go of the claim that punishment is principally about maintaining order via deterrence, incapacitation, and other welfare-increasing qualities of the practice, or (b) to persuade the retributivist to let go of the claim that punishment's principal legitimating feature is the simple, backward-looking point that the accused committed a crime and thereby forfeited her right not to be punished.

I think that neither enterprise has much chance of success, so I propose in Part III a way to break the stalemate and accommodate both central intuitions, to one degree or another. It will satisfy neither the wholly committed, thoroughgoing retributivist nor her utilitarian counterpart. But my hope is that it might be persuasive to the less entrenched, and that it might help to move the dialogue forward. Before considering that proposal, it is useful to consider briefly the larger philosophical context within which the two major punishment theories are situated.

2. Deontology and Consequentialism

The entrenched disagreement in punishment theory between retributivists and utilitarians seems particularly normal and expected when one considers that it mirrors a disagreement at a higher level of generality. No dispute in modern moral and political theory runs deeper than that between deontologists and consequentialists, and the philosophy of criminal punishment may be viewed as one small front on which the larger battle is being waged.

The central question of ethics is what distinguishes between right and wrong, moral and immoral. It is the question of how we determine whether an act is good or bad; that is, whether it is the sort of thing we say people should do or the sort of thing we say they should refrain from doing. For Jeremy Bentham and his followers in the utilitarian movement, the answer was straightforward: An act is good if it improves people's

---

92. *Id.* (stating that retributivism has its "roots in vengeance, bloodlust, revenge, retaliation, and an eye for an eye").

93. *See* McDermott, *supra* note 12, at 407 ("Retributivists claim that punishment is morally permissible *because* wrongdoers have committed crimes—it is the fact of crime itself that changes wrongdoers' moral boundaries such that treatment that is normally morally impermissible becomes permissible." (emphasis in original)).

94. *See* Jeffrie G. Murphy, *Marxism and Retribution*, *in* A READER ON PUNISHMENT 47, 49 (R.A. Duff & David Garland eds., 1994) ("Since we do not always have the right to do what it would be good to do, this distinction is of the greatest moral importance; and missing the distinction is the Achilles heel of all forms of Utilitarianism. . . . This kind of distinction is elementary, and any theory which misses it is morally degenerate.").

lives, and it is bad if it worsens those lives.[95] The definition of morality is that which increases welfare or makes life better.[96] By contrast, for Immanuel Kant and his followers, whether an act was moral or immoral depended not on its consequences but instead on the act itself.[97]

A major advantage for utilitarians in this debate is that it is easier for them to ground morality in something tangible and concrete, and to explain what determines whether something is right or wrong.[98] A Kantian deontologist must explain what, if not utility, could determine the rightness of an act, and that question is not easy to answer.[99] Kant himself made a heroic attempt via the analytic device of the "categorical imperative," suggesting that an act meets the demands of morality if it can be universalized as a requirement under all circumstances without leading to a logical contradiction.[100] Others have tried milder formulations, emphasizing Kant's insistence on respecting people as ends in themselves rather than treating people merely as means to achieve the ends of others.[101]

Neo-Kantian approaches to morality have made progress in articulating the values of fairness and respect for the humanity of others that most of us associate with morality. But even so, these philosophers are still grounding their view of morality in a source more nebulous than that of the utilitarians. Human happiness or welfare is a phenomenon whose existence and properties are familiar to us all. Fairness, by contrast, is an incorporeal concept. It would have been understandable (albeit not necessarily correct) for Bentham to declare it, like rights, "nonsense upon stilts."[102]

---

95. Christopher, *supra* note 55, at 856 ("The most well-known version of consequentialism is Jeremy Bentham's utilitarianism in which a course of conduct is evaluated by the principle of utility or the amount of happiness and suffering that is generated by the conduct.").

96. Jason Lloyd, Note, *Let There Be Justice: A Thomistic Assessment of Utilitarianism and Libertarianism*, 8 TEX. REV. L. & POL. 229, 231 (2003) ("In Bentham's system, therefore, the only standard of justice is pleasure and the object of law is the maximization of pleasure.").

97. Virginia Dailey, *Sustainable Development: Reevaluating the Trade vs. Turtles Conflict at the WTO*, 9 J. TRANSNAT'L L. & POL'Y 331, 356 (2000) ("According to Kant, moral obligations are rooted in human nature; morality is found in the act itself, without regard to a cost-benefit analysis.").

98. *See* Huigens, *supra* note 4, at 953–55 ("The principal shortcoming of the retributive theory is that the source and scope of the duty to punish is not clear. . . . The deterrence theory of punishment arguably does not suffer from these burdens because it rests on a less abstract ethical theory: consequentialism. Under consequentialist ethics, punishment is justified by the deterrence of harm or, more broadly, by the promotion of social welfare." (citation omitted)).

99. *Cf.* McDermott, *supra* note 12, at 405 ("Demonstrating that retributive punishment is morally permissible will thus require an explanation of the relationship between acts of wrongdoing, the wrongdoers' altered moral boundaries, and the specific punishments we wish to inflict.").

100. IMMANUEL KANT, THE MORAL LAW: GROUNDWORK OF THE METAPHYSIC OF MORALS 70 (H.J. Paton ed. & trans., Hutchinson & Co. 1964) (1785).

101. *E.g.*, CHRISTINE M. KORSGAARD, CREATING THE KINGDOM OF ENDS (1996).

102. 2 JEREMY BENTHAM, *Anarchical Fallacies: Being an Examination of the Declarations of Rights Issued During the French Revolution*, *reprinted in* THE WORKS OF JEREMY BENTHAM 489, 501 (John Bowring ed., Russell & Russell, Inc. 1962) (1843).

The utilitarians, some of whom are the present-day adherents of the law and economics movement,[103] probably find their non-utilitarian counterparts confusing for this reason. Utility is something real, and making people happier or better off seems to be the most worthy thing one could hope to do.[104] Sacrificing such important, real benefits in the name of an intangible concept like fairness is, on this view, a grave mistake that is bizarre and potentially dangerous.[105] To advocate that sort of mistake might even be definitionally irrational.

This outlook has considerable appeal, but as in punishment theory, there are two sides to the story. Deontologists can give examples that reveal core human preferences for fairness over utility.[106] A famous example is that of the doctor who could murder a healthy patient in order to give each of his organs to a different dying person, sacrificing the one for the many. Most people recoil at that suggestion, regardless of whether it could be done secretly. And even those who might accept it probably are disturbed by the idea of killing the innocent, in a way that they would not be disturbed by the idea of someone's dying of natural causes. A death by natural causes could be sad, regrettable, and even tragic; but murdering someone who came in for a check-up, even to save five other people, strikes most of us as something that goes beyond sad. It seems unfair and morally problematic.[107]

A thoughtful utilitarian must grapple with this truth about the way people look at the world, rather than simply deny it or reject it with a blanket accusation of irrationality. To define rationality as welfare maximization, and then label non-utilitarians as irrational, is an unhelpful syllogism. If a facet of our psychological and even physiological composition as human beings is that we care about fairness and sometimes prioritize it over utility, then that needs to become a part of how we analyze moral choices. Treating it solely as a perversion that must be overcome might be fighting a battle against our nature that we are constituted to lose.

These broader questions merit (and have received) far longer discussion elsewhere, but for now we may return to the narrower point of reconciling the different approaches in the arena of punishment theory.

---

103. *See* Steven G. Medema, *Sidgwick's Utilitarian Analysis of Law: A Bridge from Bentham to Becker?*, 9 AM. L. & ECON. REV. 30, 31 (2007) ("Here, Bentham laid out several ideas that are reflected in the modern economic analysis of criminal law, including . . . the social benefit-cost analysis of punishment . . . .").

104. *See* David Dolinko, *Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment*, 16 LAW & PHIL. 507, 520 (1997) ("Happiness . . . is widely believed to be intrinsically good.").

105. *Cf. id.* at 519 ("[C]onsequentialists who believe that there is only one intrinsic good—whether it be happiness, utility, preference-satisfaction, or something else—would presumably assert that there *is* a duty to bring about, and indeed to maximize, the level (whether average or aggregate) of this intrinsic good." (emphasis in original)).

106. *See* Weinreb, *supra* note 84, at 52 ("There is . . . a relationship between responsibility and punishment, mediated by the idea of desert, which requires no additional argument.").

107. *See* Mirko Bagaric & Kumar Amarasekara, *The Errors of Retributivism*, 24 MELB. U. L. REV. 124, 133 (2000) ("[U]tilitarianism fails to protect basic individual rights and interests, and since it does not prohibit anything per se, [it] may lead to horrendous outcomes.").

III. A NEW FRAMING DEVICE

The main contribution of this Article is to propose a way to take a step forward in the discourse on punishment that does not depend upon showing the ultimate truth of either retributivism or utilitarianism. Instead, I hope to reorganize the discussion so as to locate the main role of each theory, making use of the strengths each one has while simultaneously using them to shore up each other's weaknesses.

A. The Two Questions

As noted in the Introduction, my suggestion could hardly be simpler. It is merely to replace the central question of punishment theory,

> Why is criminal punishment justified?

with these two questions:

> (1) Why does the state have the right to punish?
> (2) Why does the state choose to exercise that right?

This bifurcation of the question is not, of course, meant to resolve automatically the important debates in punishment theory. But neither is it a mere terminological trick that gets us nowhere beyond those same debates. It is valuable for one central and straightforward reason: *The retributive theory is at its best when answering the first question, and the utilitarian theory is at its best when answering the second question*.

If you were told that the state was planning to punish someone and then were asked whether you thought the state had the right to do so, what additional piece of information would you most want to know before answering? You would probably most want to know what the person had done—in particular, whether she had committed a crime. When deciding whether the state has the right to punish (or whether it is morally permissible to punish, or whatever related formulation you prefer), the primary, if not the only, consideration looks backward in time to the crime, rather than forward in time to the good consequences of punishment. Ordinarily, people have the right not to be punished by the state. Only when they break the law, usually by harming others in one way or another,[108] do they forfeit their right to be free from punishment.

This does not mean, however, that the mere fact of their law-breaking is the only reason, or even the primary reason, that the state chooses to exercise its right to punish them. Once someone has relinquished the right to be free from punishment, it becomes legitimate for the state to pursue its goals of maximizing welfare by imposing such punishment.[109] Obviously, the claim is not that any form or amount of punishment, no

---

108. The right of the people to make laws prohibiting harmless behavior, and to enforce those laws by punishment, is a fascinating and widely discussed topic that is well beyond the scope of this Article. *See generally* 4 JOEL FEINBERG, THE MORAL LIMITS OF THE CRIMINAL LAW: HARMLESS WRONGDOING (1988).

109. *See* Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181, 188 (2004) ("One might easily imagine that the role of a system of criminal procedure is to impose just

matter how large, would be justified by any form or amount of law-breaking, no matter how small. The claim is merely that whatever punishment is justified finds its source of legitimation in the law-breaking rather than in some future utility increase, whereas the reasons for choosing to impose that legitimate form and degree of punishment may well involve utility.

Just as it would make no sense to decide whether the state has a right to punish without knowing if the person is guilty, it would make no sense to decide whether to exercise the right to punish the guilty without considering the consequences of the decision. Punishing the guilty is a policy choice like anything else, and if the state and its citizens prefer to spend more tax dollars on roads or schools than on prisons (or vice versa), then that is their prerogative. Such decisions will be made by reference to everything that composes people's individual and collective preferences, including the desire to increase welfare as well as to live in a society that treats its citizens fairly.

B. The Roles of the Two Theories Within the New Framework

So far I have said that retribution answers the first question and utilitarianism answers the second—that is, the state has the right to punish because a criminal forfeits his claim to be free from punishment when he breaks the law; and the state chooses to exercise its right to punish by deciding that punishment will have beneficial consequences for the citizenry such as deterring future crime. But more needs to be said about the extent to which there may or may not be overlap between the theories in answering each question.

This is not the primary focus of my inquiry, and I mainly intend the answers to the questions posed by my framing device to be supplied by the retributivist and utilitarian scholarship of others. Nonetheless, a tentative suggestion or two follows about how the theories might work in answering the questions.

1. The Right to Punish: Retribution Alone

The reason the state has the right to punish is that a criminal has broken a legitimate law.[110] The source of the right comes from the past law-breaking, not the future good consequences.[111] One need not be able to answer all of the deep questions associated with justification in order to recognize that the vast majority of people share a core considered judgment that the right to punish derives from the wrongdoing of the accused. If we assume, as I do, that this judgment is so central that it must be a part of

---

punishments and that direct application of the coercive power of the state is the necessary and sufficient means to this end.").

110. I have mentioned earlier that it might be acceptable to punish the innocent in extreme cases such as the need to save millions of lives. The implications of this sort of caveat are generally beyond the scope of this Article, but I note here that my sentence in the text is not meant to rule out punishment in such a hypothetical example. I do not think this means that utilitarianism supplies part of the reason the state has a right to punish, at least in the vast majority of the cases that actually arise.

111. HONDERICH, *supra* note 51, at 152 ("We wish to protect society, reform criminals, deter others. These ends or purposes are 'morally and socially desirable' but they should not be confused with the moral justification of punishment. The moral justification of the practice is that it is deserved.").

any justification of punishment that hopes to gain broad acceptance, then retribution, rather than utilitarianism, must supply the answer to the first question.

But must it be the only answer? A utilitarian could argue that even if law-breaking and its attendant rights forfeiture are necessary, they are not sufficient—that the state has the right to punish only if the accused is guilty *and* the consequences of the punishment will increase utility.

I do not find that argument particularly convincing, though a complete assessment would take us too far afield. As discussed in Part II.B.2, what it means to be a utilitarian is to believe that the moral value of an act is defined by the act's overall effect on people's welfare (or the reasonably foreseeable effect, or some similar variation on the theme). According to this view, it would be not only unwise, but also morally impermissible, to impose punishment if the benefits seem unlikely to outweigh the costs.[112]

However, anyone who accepts that guilt is a prerequisite for punishment has already rejected this utilitarian value system. If we deem it legitimate to abstain from punishing the innocent even when that abstention decreases utility, then why would it be illegitimate to punish the guilty when such punishment decreases utility? Punishing is active whereas refraining from punishing is passive, but that distinction would be irrelevant to the utilitarian in this context. A state that chooses to punish a criminal even when doing so will not increase the overall welfare might be making a poor policy choice, or it might be honoring its citizens' wishes to promote values other than their narrowly defined self-interest. But it is not wronging the criminal, because she forfeited her right to avoid being punished when she committed the crime. It is not clear why that forfeiture would depend upon whether the punishment would be useful to society, unless one is a thoroughgoing utilitarian, in which case the crime and its attendant rights forfeiture are irrelevant in the first place.

It appears, therefore, that retribution, and retribution alone, answers the question of why the state has the right to punish.

2. The Choice to Exercise the Right: Both Utilitarianism and Retribution?

In the same way that the right to punish depends upon what happened in the past (i.e., the crime), the decision whether to exercise that right depends upon what will happen in the future—the consequences of punishing. When an act is morally permissible, the choice of whether to do it ordinarily involves a consideration of the consequences of doing it. If those consequences seem desirable, then the act will be undertaken; otherwise it will not.

This is the main sort of analysis relevant to a state when it considers whether to punish someone who has been convicted of a crime, and the analysis is consequentialist in nature. All of the possible values emphasized by utilitarian punishment theories are at work in answering the second question in my framing device. When the state decides whether to punish, it considers whether the punishment will deter others from committing crimes, whether the offender herself will be incapacitated from engaging in future bad acts, and whether the punishment might help rehabilitate the offender,

---

112. *See* Cassell, *supra* note 26, at 1031 ("[P]unishment should be 'optimal' in the sense that its benefits outweigh its costs.").

among other things. In this way, the second question creates a natural home for the utilitarian theory just as the first question did for the retributivist theory.[113]

That said, there may well be an asymmetry: Whereas retributivism alone was relevant to answering the first question, utilitarianism might not own sole title to the second. Most retributivists would say that retribution supplies not only reasons why the state has a right to punish, but also reasons why it chooses to exercise that right. And although such retributive reasons for exercising the right might not take precedence over their utilitarian counterparts, they are important enough to be considered alongside those counterparts, or so the argument would go.

What are these retributive reasons for the state to exercise its right to punish an offender? The most interesting is the idea that punishment is society's way of standing up for the victim and reaffirming her value—of making the statement that the wrongdoer had no right to do what he did.[114] Jean Hampton explored the philosophical underpinnings of this idea,[115] and today Kenworthey Bilz is doing valuable empirical work to demonstrate its importance.[116] According to Bilz's research, state punishment restores the victim's social status in the eyes of the community while appropriately lowering that of the offender.[117] Her data suggest that this may even be the primary reason that most people view punishment as necessary.[118]

It is plausible and appealing to view punishment as a device that reaffirms the victim's value and prevents criminals from being successful bullies who rule their neighborhoods by intimidation and violence. Whether this aim of punishment constitutes a large or small part of the collection of reasons we punish, it seems unreasonable to deny that it plays any part at all. A related retributive reason for choosing to punish is that doing so expresses society's condemnation of the criminal act and makes clear that the state in no way condones the act.[119] This, too, seems to be among the myriad reasons why the government might choose to punish those whom it has a right to punish.

The utilitarian might argue that these sorts of reasons for punishment are utilitarian rather than retributive. If people value highly the effect punishment has of restoring the victim's social status while diminishing that of the criminal, then those people's welfare is increased by punishment in this way. And if people are made happy by

---

113. *See* Lee, *supra* note 21, at 738 ("The purpose of punishment, under [the utilitarian] view, is not to give each criminal what he or she deserves, but to deter future crimes, to incapacitate criminals by keeping them 'off the streets,' or to rehabilitate criminals so they would become better citizens.").

114. *See* Dan M. Kahan, *The Secret Ambition of Deterrence*, 113 HARV. L. REV. 413, 420 (1999) ("The meanings of wrongdoing and punishment, moreover, are related: the condemning retort of punishment signals society's commitment to the values that the wrongdoer's act denies.").

115. *See generally* Jean Hampton, *The Moral Education Theory of Punishment*, 13 PHIL. & PUB. AFF. 208 (1984).

116. *See* Kenworthey Bilz, The Effect of Crime and Punishment on Social Standing (Sept. 2006) (unpublished Ph.D. dissertation, Princeton University) (on file with author).

117. *Id.*

118. *See id.*

119. Jens David Ohlin, *Applying the Death Penalty to Crimes of Genocide*, 99 AM. J. INT'L L. 747, 768 (2005) ("Under [the expressive] theory, the institution of punishment serves to express society's disapproval of a criminal act.").

societal denunciation of crime, then this too constitutes a utilitarian reason. The question whether to label these points retributive or utilitarian turns perhaps on whether they give the state reason to punish *because* of the increased welfare they produce or instead create a reason to punish *independent* of such increased welfare. Although these benefits of punishment have been regularly treated in the literature as belonging to the retributive category, my inclination goes in the other direction. Reaffirming the victim's status and expressing societal condemnation of crime are concrete aims or benefits of punishment—reasons why punishment improves people's lives.[120] Their value lies in that improvement and is not independent of it, or so it seems to me.

So utilitarians may well be able to lay claim to the second question as thoroughly as retributivists lay claim to the first, provided that they accept some overlap with retributivism in labeling certain reasons for choosing to punish, such as affirming the victim's value. But there is one truly retributive purpose of punishment that cannot be relabeled utilitarian or consequentialist. It is the view that the state must choose to punish the criminal because to let crime go unpunished is, at least in some cases, inherently immoral.[121]

This is, of course, characterized by retributive and deontological reasoning as thoroughly as an argument for exercising the right to punish can possibly be. It is not utilitarian, but is it right? Perhaps the most promising route toward an answer is to imagine an irredeemable serial offender who has caused extreme harm to his victims and shows no remorse and an unquenchable appetite for creating further misery.[122] If this person somehow ended up in a place removed from all other people and if, for the sake of argument, he could be punished there but no one else would ever know of it, then should he be punished? If no one else would benefit, and he would be harmed, then the example appears to be a relatively pure form of the question whether punishment is sometimes morally required independent of beneficial consequences.

I suppose I would favor punishment in such a case, but that intuition might well arise from a failure fully to internalize and accept the hypothetical stipulation that no one's life is improved in any way by the punishment. More than anything else, this final conundrum (whether retribution can be an answer to the second question by virtue of the claim that we should sometimes choose to punish for the sake of punishment

---

120. *See* Jennifer M. Collins, *Crime and Parenthood: The Uneasy Case for Prosecuting Negligent Parents*, 100 NW. U. L. REV. 807, 838 (2006) ("Punishment is, therefore, 'inflicted to nullify the wrongdoer's message of superiority over the victim, thus placing the victim in the position she would have been in if the wrongdoer had not acted.' Thus, 'retribution is actually a form of compensation to the victim.'" (quoting Hampton, *supra* note 23, at 1698)).

121. H.L.A. HART, *Postscript: Responsibility and Retribution*, *in* PUNISHMENT AND RESPONSIBILITY 210, 232 (1968) (stating that one feature required by Kantian retributivism is that punishment of an offender is "obligatory, even on the eve of a dissolution of a society against whose laws the person to be punished has offended").

122. *See* IMMANUEL KANT, THE METAPHYSICS OF MORALS 106 (Mary Gregor ed. & trans., Cambridge Univ. Press 1996) (1785) ("Even if a civil society were to be dissolved by the consent of all of its members (e.g., if a people inhabiting an island decided to separate and disperse throughout the world), the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve and blood guilt does not cling to the people for not having insisted upon his punishment; for otherwise the people can be regarded as collaboraters [sic] in this public violation of justice.").

itself) provides a segue into the next section, in which I assess the framing device's value and limitations. The inability to answer the question of whether punishment can sometimes be morally required for its own sake is such a limitation.

IV. ASSESSING THE NEW DEVICE

As explained repeatedly above, I believe that breaking the central question of punishment theory into the particular components I have proposed has the virtue of helping to remove the literature on punishment from the quagmire of two theoretical camps that each decline to budge. To assess the proposed framing device, it is useful to consider the extent to which it should satisfy members of each camp, and also to identify its limits (i.e., the important questions of punishment theory that the device does not help to answer).

A. To What Extent Does the Framing Device Satisfy Retributive and Utilitarian Concerns?

Neither an uncompromising retributivist nor her utilitarian counterpart will have much use for the proposal advanced in this Article. For such a retributivist, the justification of punishment has nothing whatsoever to do with the good consequences of punishment such as deterrence and incapacitation.[123] Indeed, to punish someone for the sake of those good consequences would be seen by this sort of scholar as violating the Kantian imperative to treat people as ends rather than merely as means.[124] When our reason for inflicting harm on one person is that it will benefit others, we are using the criminal merely as a means.[125] Thus, in the view of a thoroughgoing retributivist, both my first and second questions must be answered by reference only to retribution, and therefore the questions are no different from or better than the old individual question of justification.

My quibble with the extreme retributivist is that her position accords inadequate importance to the fact that I answer the first question by reference to retribution. By requiring that punishment be meted out only to the guilty, and that their guilt be the only *reason* that the state has a right to punish them, my formulation insures that they are being treated with respect as ends in themselves when they are punished. Indeed, as others have noted,[126] the view that one forfeits some of her rights when she breaks the law shows respect for the autonomy of the accused. We hold her accountable for her choice and thereby treat her with respect, rather than looking down on her as someone who is sick or unable to make rational choices.

---

123. HART, *supra* note 121, at 231 (stating that for a retributivist, the justification of punishment is "that the return of suffering for moral evil voluntary [sic] done, is itself just or morally good").

124. *See* KANT, *supra* note 100, at 101 ("For rational beings all stand under the *law* that each of them should treat himself and all others, *never merely as a means*, but always *at the same time as an end in himself*." (emphasis in original)).

125. IMMANUEL KANT, THE METAPHYSICAL ELEMENTS OF JUSTICE 138 (John Ladd trans., Hackett Publ'g 2d ed. 1999) ("Judicial punishment can never be used merely as a means to promote some other good for the *criminal himself* or for civil society." (emphasis added)).

126. *E.g.*, Herbert Morris, *Persons and Punishment*, 52 MONIST 475 (1968).

In any event, Kant understood that we regularly and unproblematically treat others as a means to an end. He required only that we simultaneously treat them as ends in themselves. My proposed approach to punishment theory does just that. The state punishes the criminal in order to achieve the end of increasing welfare via deterrence and other values, but it also treats the criminal as an end by grounding its right to punish him exclusively in his crime. Moreover, apart from considering which view is more in keeping with Kantian theory, it seems clearly warranted to consider the consequences of punishment for others in society once we have confined that punishment to the guilty. To make policy decisions without allowing ourselves to consider the consequences of those decisions seems utterly unreasonable.

This still leaves us, however, to grapple with the dyed-in-the-wool utilitarian. For her, the state's right to punish can be no more independent of utility than can the reasons for choosing to exercise that right. As discussed above, the true utilitarian sees no value outside of utility. People's well-being is the only tangible value that it makes sense to promote. Concepts like rights and justice, if divorced from utility, are meaningless and without content. So to say that the state has a right to punish for reasons unrelated to increasing public welfare is silliness.

Again, this view was addressed above by the point that very few of us see the world that way, even when we think carefully about it. Vague and indistinct though fairness may be, it nonetheless has a hold on us that heavily affects our outlook on which acts we accept as legitimate and why. To put it most directly, we are willing to be less happy in order to live in a society we deem just. We seem psychologically constituted to care about justice, not merely as a surrogate for utility, but as a separate value in its own right. If this is true, then our concept of morality and of what acts may legitimately be undertaken is shaped in large part by non-utilitarian considerations.[127] We believe that punishing the innocent is wrong, and that we have the right to punish the guilty *because* they are guilty. These beliefs cannot be reconciled with the comprehensive utilitarian's approach to punishment theory.

Having given my responses to the committed retributivist and utilitarian, it should be noted that they are not my target audience. The real value of the framing device is that it might help punishment theorists who are not perfectly committed to one side to bridge the divide. I have asked the questions in the hope of making it easier for members of both camps to use the strengths of their theories to flesh out the answers. Let us therefore turn to the last part, in which we may consider the work that remains to be done by others in answering those questions.

B. Limits of the Framing Device

By far the most important task left open by my two questions is that of explaining *why* the state has a right to punish the guilty. For the purposes of this Article, it was enough to identify our conviction that commission of crime is the reason we accept state punishment. But as a philosophical matter, this will not do. A developed theory is needed to explain the specific way in which law-breaking gives rise to the right to

---

127. *See generally* Brian Leiter, *Morality Critics*, *in* OXFORD HANDBOOK OF CONTINENTAL PHILOSOPHY 711, 711 (2007) ("Popular, including religious, thinking has long proceeded on the assumption that 'morality' as a system of norms deserves our allegiance and that 'moral conduct' should earn our praise and admiration.").

punish. Perhaps such a theory would try to fuse neo-Kantian principles of mutual respect[128] with Rawlsian social contract theory,[129] and apply it to punishment by reference to Herbert Morris's view of benefits and burdens.[130] Morris suggested that the reason a lawbreaker forfeits the right to be free from punishment is that he has refused to do his fair share in bearing the burden of reduced freedom that is the hallmark of life in society.[131] Richard Dagger has written a most interesting piece along these lines.[132]

To ask the questions is thus not to answer them, but rather to construct a more accommodating vessel into which the answers might be poured. As described earlier, another important question left unanswered is whether punishing the guilty is morally required or merely permissible.[133] My formulation of the two questions might make it seem as though my answer is the latter, because I suggest that once it is established that someone may permissibly be punished, the choice whether to carry out that punishment is left to considerations of utility. But that is perhaps a misleading feature of the structure of the questions. I actually remain agnostic as to whether punishment can be morally required on pure retributive grounds. To answer it satisfactorily would require a theory of retribution and perhaps of morality itself that runs deeper than I can presently conceive.

Finally, asking the questions as I have does not of course provide a formula for the specific amounts of punishment (e.g., the number of years in prison) that are warranted in response to any given crime. Upper bounds must be set, and some principle must be adopted to link the severity of crime with the specific severity of punishment. A blunt equivalence principle is not the answer (torturing a torturer, raping a rapist), but something related to equivalence, with overlapping absolute restrictions on form and severity, might be.[134]

CONCLUSION

When we punish, we must look both backward to the crime committed and forward to the need to protect future victims. Retributive theories tap into the former truth and utilitarian theories tap into the latter one. Because adherents of each are seizing upon a

---

128. *See* KORSGAARD, *supra* note 101.

129. *See* Dolovich, *supra* note 18, at 314–15 ("Rawls argues that political power is legitimately exercised by the state over its citizens only when it is exercised on the basis of a collective agreement 'the essentials of which all citizens may reasonably be expected to endorse' under fair deliberative conditions.").

130. *See* Morris, *supra* note 126, at 476–78.

131. *Id.*

132. *See* Dagger, *supra* note 46, at 474 (arguing that the principle of fair play should play a central role in the discussions of the justification of punishment).

133. *See* Dolinko, *supra* note 104, at 509 ("A deontological retributivist . . . takes punishing the guilty to be obligatory on each particular occasion.").

134. *See id.* at 526 ("'[A] person's desert of X is always a reason for giving X to him, but not always a conclusive reason, [because] considerations irrelevant to his desert can have overriding cogency in establishing how he ought to be treated on balance.'" (quoting JOEL FEINBERG, *Justice and Political Desert*, *in* DOING AND DESERVING 55, 60 (1970)) (alteration in original)).

point that is accurate, they are convinced of their theory's overall veracity and determined to prove its status as the correct approach to punishment.

This state of affairs calls for a reformulation of the question that simultaneously accounts for both the retributive and utilitarian truths. I offer such a reformulation in this Article, not as an answer to the deep questions of punishment theory, but as a helping hand to those whose scholarship aims to provide those answers.