

2016

Modest Retributivism

Mitchell N. Berman

University of Pennsylvania Law School, mitchberman@law.upenn.edu

Follow this and additional works at: http://scholarship.law.upenn.edu/faculty_scholarship

 Part of the [Criminal Law Commons](#), [Criminal Procedure Commons](#), [Criminology and Criminal Justice Commons](#), [Ethics and Political Philosophy Commons](#), [Jurisprudence Commons](#), [Law and Psychology Commons](#), [Legal Theory Commons](#), and the [Theory and Philosophy Commons](#)

Recommended Citation

Berman, Mitchell N., "Modest Retributivism" (2016). *Faculty Scholarship*. Paper 1496.
http://scholarship.law.upenn.edu/faculty_scholarship/1496

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.

MODEST RETRIBUTIVISM

Mitchell N. Berman*

(March 19, 2014)

INTRODUCTION

Michael Moore is a giant of legal philosophy. Over the course of an academic career that already spans more than four decades yet shows no signs of abating, Moore has made profound contributions to our collective understanding of such diverse matters as moral realism and its implications for law, the relationship between law and psychiatry, the theory of action that undergirds Anglo-American criminal jurisprudence, the nature of legal interpretation, and, most recently, the metaphysics of causation. Very probably, though, Moore has achieved his greatest fame and influence as a theorist and proponent of a retributivist justification for criminal punishment.

When Moore's first papers in punishment theory appeared in the early 1980s, retributivism was only starting to recover from what had become, by mid-century, its near-total repudiation in respectable philosophical circles. Within a generation of Moore's arrival, however, scholars were widely contending—albeit more in horror than satisfaction—that retributivism had reemerged as the dominant account of the moral justifiability of the infliction of criminal punishment.¹ Regardless of whether that is an accurate description of the state of theory in the early twenty-first century (and I confess to some skepticism), that retributivism has enjoyed a remarkable revival cannot reasonably be questioned. Nor

* Richard Dale Endowed Chair in Law, and Professor of Philosophy, the University of Texas at Austin. A distant predecessor to this paper was presented at the Colloquium on Moral versus Political Perspectives on the Criminal Law, held at Berkeley Law in October, 2013. I am very grateful to the organizers and participants at that event for helpful reactions, and especially wish to acknowledge debts to Antony Duff, Alon Harel, Leora Katz, Chris Kutz, Sandra Marshall, Seana Shiffrin, Victor Tadros, and Arthur Yates. I am indebted as well to: David Dolinko for excellent written comments on a draft; the editors of this volume, Kim Ferzan and Stephen Morse, for inviting me to contribute; and our honoree, Michael Moore, for all that he has taught me.

¹ See, e.g., Mirko Bagaric & Kumar Amarasekara, "The Errors of Retributivism," 24 *Melb. U. L. Rev.* 124 (2000), 126.

can it be seriously doubted that Moore himself bears greater responsibility for the retributivist renaissance than does any other scholar. Moore is the foremost living retributivist among Anglophone moral theorists and philosophers of law.

My goal in this paper is to critically assess the form of retributivism that Moore has advocated, with rigor and acuity, and to great success. Plausibly, and only to a first approximation, the core retributivist claim—the claim that distinguishes retributivist views from their nonretributivist alternatives—holds that it is intrinsically valuable or right to furnish wrongdoers with the negative consequences that they deserve. But Moore claims much more for a wrongdoer’s negative desert than this. That is, he attributes significantly greater normative force to the non-instrumental value or rightness of furnishing a wrongdoer with his negative desert than is necessary for a view to qualify as retributivist. I will call Moore’s retributivism “robust retributivism.” An account that resides in a nearly polar neighborhood of retributivist logical space may be labeled “modest retributivism.” In this paper, I sketch robust and modest retributivism and aim to make the latter more eligible, principally by raising doubts about Moore’s arguments for each of the components that jointly comprise the former. My ambition is neither to defend modest retributivism nor to defeat robust retributivism. It is to establish that retributivism embraces a greater diversity of possible and even plausible views than Moore himself allows—including views that, in a fairly straightforward sense, are considerably more moderate than the version that Moore has defended.

I. THE CORE OF RETRIBUTIVISM

Any simple definition of retributivism will be contestable. That acknowledged, it is widely agreed that, for a justification of punishment to qualify as retributivist, a wrongdoer’s negative desert

must play some significant role in the justificatory account.² Very probably, little can be said about the particular nature of desert's role, or about how it figures into the logic of the justificatory argument, without taking sides in intra-retributivist disagreements. Accordingly, if any claim can be fairly identified as core or central to retributivism, let alone definitional of it, that claim must be formulated in quite general or vague terms.

Commentators reflect this point when describing retributivism as the theory, say, that punishment must be justified "in terms of" a wrongdoer's negative desert.³ Yet that is perhaps a little *too* general or vague. Consider the most commonly espoused version of a "mixed theory"—the view (to a first approximation) that punishment is justified if and only if it produces net good consequences and is not inflicted on persons who lack negative desert.⁴ Such a view justifies punishment by reference, in part, to an offender's negative desert and therefore would count as retributivist if mere reference to desert were sufficient to render a justification of punishment retributivist. And sure enough, John Mackie's name for this view—"negative retributivism"⁵—held for some batch of years. Increasingly, however, philosophers of punishment find it more accurate or illuminating to locate this position outside of retributivism, many preferring Antony Duff's apt label for this view, "side-constrained consequentialism."⁶

Here, then, is a stab (subject to refinement) at the core retributivist thesis:

² Doug Husak makes the same observation in his contribution to this volume. See Douglas Husak, *What do Criminals Deserve?* (this volume).

³ C.L. Ten, *Crime, Guilt, and Punishment* (New York: Clarendon, 1987), 46.

⁴ I say this is only a first approximation of the view because those who reject wholly objective conceptions of moral justification would likely add some sort of epistemic qualifiers, possibly at two places. A partially subjectivized variant of this version of the mixed theory might maintain, for example, that punishment is justified if and only if it is *reasonably expected* to produce net good consequences and is not *knowingly* inflicted on persons who lack negative desert, or is inflicted only on persons whom the punishing authorities *believe in good faith* to deserve it, or something along these lines.

⁵ J.L. Mackie, "Morality and the Retributive Emotions," 1 *Crim. J. Ethics* 3 (1982), 3.

⁶ R.A. Duff, *Punishment, Communication, and Community* (New York: Oxford, 2001), 11.

- (R) The goodness or rightness of satisfying a wrongdoer's negative desert morally justifies the infliction of criminal punishment, without regard for any further good consequences that might be realized as a contingent result of satisfying the wrongdoer's desert.⁷

I do not claim that acceptance of (R) is definitional of any thesis properly denominated "retributivist." Still, it is highly likely that (R) is accepted by the overwhelming majority of extant views about punishment that are considered retributivist by their adherents and critics alike. Importantly, (R) takes no position on such contested matters within the retributivist camp as what a wrongdoer's negative desert consists in (e.g., suffering, or being punished),⁸ or whether the logical structure of the justification that a wrongdoer's negative desert provides is consequentialist or deontological.

II. MOOREAN RETRIBUTIVISM

In a well-known portion of his 1984 book, *Law and Psychiatry: Rethinking the Relationship*, reprinted and revised as "Closet Retributivism" thirteen years later in his *Placing Blame: A General Theory of the Criminal Law*, Moore explains what retributivism is, and is not. It is not, says Moore, any of several views with which it has frequently been confused. For example, it is not *lex talionis*, for retributivists qua retributivists need not endorse any single view of the proper measure of punishment. It is also not "negative retributivism," for whilst retributivists will endorse this position, one need not be a retributivist to do so. "The distinctive aspect of retributivism is that the moral desert of an offender is

⁷ (R) should not be taken to claim that the goodness or rightness of satisfying a wrongdoer's negative desert justifies the infliction of punishment "conclusively" or "all things considered," without regard for other considerations. The form of moral justification at issue in (R) is what I call "tailored." I elaborate on the distinction in Section III.B, below.

⁸ On Feinberg's influential triadic analysis of desert, an agent, A, deserves some treatment or state of affairs, O, on account of some action or property, B. Feinberg called that which grounds a deserved treatment (B, in the statement above) the "desert basis." Joel Feinberg, *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton, NJ: Princeton University Press, 1970). I call that which is deserved (O, in the statement above) the "desert object." To suffer and to be punished are different desert objects if, as is widely accepted, punishment itself is defined in partly functional or intentional terms, say as a treatment intended to be experienced as disagreeable by the subject upon whom it is imposed. See, e.g., H.L.A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (New York: Oxford, 1968), 4-5. If what a wrongdoer deserves is *to suffer*, then his negative desert can be realized even if he is not *punished*. What about the converse? Is it punishment if the person subjected to it does not suffer? Or, better, must punishment be designed to inflict suffering (or something close to it)? People hold different views about this.

a *sufficient* reason to punish him,” and not only, as negative retributivism maintains, that it is a necessary condition.⁹ Similarly, retributivism is not a denunciatory theory of punishment, “should not be confused with a theory of formal justice (the treating of like cases alike),” and is not the view “that punishment of offenders satisfies the desires for vengeance of their victims, . . . [or] that the preferences of all citizens (not just crime victims) should be satisfied, . . . [or] that punishment is justified because without it vengeful citizens would take the law into their own hands.”

Retributivism is none of these things. Instead, it

is a very straightforward theory of punishment: We are justified in punishing because and only because offenders deserve it. Moral responsibility (“desert”) in such a view is not only necessary for justified punishment, it is also sufficient. Such sufficiency of justification gives society more than merely a *right* to punish culpable offenders. It does this, making it not unfair to punish them, but retributivism justifies more than this. For a retributivist, the moral responsibility of an offender also gives society the *duty* to punish. Retributivism, in other words, is truly a theory of justice such that, if it is true, we have an obligation to set up institutions so that retribution is achieved.¹⁰

I accept each and every contention in Moore’s list of “not’s.” None of the views from which Moore withholds the retributivist label is equivalent to, entails, or is entailed by, (R). But nor does the rejection of these several views entail the theory that Moore ends up advancing as the definition of retributivism. Three features of Moore’s definition of retributivism warrant isolation and attention in particular.

First, for Moore, the goodness or rightness of giving wrongdoers what they deserve does not justify criminal punishment only in the sense that it renders the infliction of punishment morally permissible. Rather, the goodness or rightness of giving wrongdoers what they deserve grounds a moral *obligation* to inflict criminal punishment. In a footnote that follows the fourth sentence in the passage quoted above, Moore specifically cites K.G. Armstrong for the proposition that retributivism “only

⁹ Michael Moore, *Placing Blame: A General Theory of the Criminal Law* 88 (New York: Oxford, 1997), 88. For simplicity, all citations to Moore’s work will be to *Placing Blame* alone, and not also to predecessor books and articles.

¹⁰ Moore, *Placing Blame*, at 91 (emphases in original; citation omitted).

giv[es] the state the *right* to punish but not the *duty* to do so,” but dismisses this view as yet another “mischaracterization of retributivism.”¹¹ As he puts it elsewhere, while “[t]he desert of offenders certainly gives [state] officials *permission* to punish offenders, . . . retributivism goes further. As a theory of a kind of justice, it *obligates* us to seek retribution through the punishment of the guilty.”¹²

Second, the goodness or rightness of giving wrongdoers what they deserve does not merely justify (by making obligatory) the infliction of punishment in an individual case. Rather, it grounds an obligation to establish and maintain *institutions* devoted to the imposition of criminal punishment. The obligation to inflict retributivist punishment, says Moore, “means that officials have a duty to punish deserving offenders and that citizens have a duty to set up and support institutions that achieve such punishments.”¹³

Third, it is not merely the case that the goodness or rightness of giving wrongdoers what they deserve, without regard for any further good consequences of doing so, is sufficient to justify the infliction of criminal punishment. Rather, Moorean retributivism maintains that effectuating just deserts is the *only* goal, consideration, or rationale that can justify criminal punishment. It is “only because” offenders deserve punishment that we are justified in inflicting it.

Each of the foregoing three features stakes a position on a distinct dimension on which different accounts of the justifiability of punishment could plausibly vary. Call these the dimensions, respectively, of *force*, *scope*, and *exclusivity*. With respect to *force*, Moore maintains that, not only does the goodness or rightness of inflicting negative desert provide a reason to inflict punishment that is sometimes adequate to justify it all things considered, but it grounds a duty to do so as well. With respect to *scope*, he contends that a wrongdoer’s negative desert provides an obligation, not only to

¹¹ More precisely, Moore calls Armstrong’s conception “perhaps an eighth mischaracterization of retributivism.” Moore, *Placing Blame*, at 91 n.17 (citing K.G. Armstrong, “The Retributivist Hits Back,” in H.B. Acton (ed.), *The Philosophy of Punishment* (London: Macmillan, 1969)). It is hard to know what sense to make of that (uncharacteristically) fudgy “perhaps.”

¹² Moore, *Placing Blame*, at 154 (emphases in original).

¹³ *Id.*

impose criminal punishment in individual cases, but to establish and maintain institutions of criminal punishment as well. And with regard to *exclusivity*, Moore claims that furnishing negative consequences because they are deserved is the only morally acceptable justification for the infliction of criminal punishment.

Let us call this package of views “robust retributivism.” It is a stipulated term of art. David Dolinko termed a version of retributivism that is strong on the dimension of force alone (i.e., that views the satisfying of negative desert not only as permissible but as obligatory) “bold” retributivism.¹⁴ Robust retributivism is triply bold.

A retributivist package of views that represents a polar alternative to robust retributivism would differ on all three of these dimensions. It would maintain that the value or rightness of realizing negative desert provides reason to punish but does not ground an obligation to punish; that considerations other than desert must be relied upon to justify the creation and maintenance of an institution of criminal punishment; and that proper recognition of the moral significance of negative desert does not foreclose the possibility that the act or institution of punishment could be justified on the strength of other values or reasons. Call this package “modest retributivism.” The two positions are summarized in the table below.

¹⁴ David Dolinko, “Some Thoughts About Retributivism,” 101 *Ethics* 537 (1991), 542.

	Robust Retributivism	Modest Retributivism
Scope	Negative desert justifies the institution of punishment as well as individual instances.	Negative desert justifies the infliction of punishment on wrongdoers, but does not by itself justify creation and maintenance of the institution of criminal punishment.
Force	Negative desert grounds not only a reason to punish but a duty to punish as well.	Negative desert grounds only a reason to punish, and not also a duty.
Exclusivity	Negative desert grounds the only morally acceptable justification for the infliction of criminal punishment.	Although negative desert can justify the infliction of criminal punishment, punishment is at least potentially justifiable on other grounds as well.

There are possible views intermediate between robust and modest retributivism. For example, one could maintain that negative desert justifies the institution of punishment but is not the only potentially available ground for a satisfactory justification. In focusing on the especially minimal account that I am calling modest retributivism, I do not intend to deny or disparage this or any other intermediate options. Similarly, the dimensions of force, scope, and exclusivity are not the only dimensions on which versions of retributivism could differ. I do not even claim that they are the most significant ones. I claim only that they are significant enough to warrant our attention.

III. LESS THAN MOORE

To reiterate a point made earlier, I am not setting out to show that Moorean robust retributivism is wrong or defective. I aim only to show that several of the elements that comprise it are optional from a retributivist perspective. Moore himself equates his version of retributivism (what I'm terming "robust retributivism") with retributivism, *simpliciter*. It is my impression that many people view things just that way. My aim here is to raise doubts about that equation and thus to make more apparent other retributivist possibilities, in particular possibilities that require negative desert to do significantly less work than Moore demands of it.

A. Force

In his 1993 article, *Justifying Retributivism*, Moore acknowledged (for the first time in print, as far as I'm aware) that retributivism could be a consequentialist approach to justifying punishment. "One can, of course, *define* retributivism so that a retributivist must be a deontologist," he remarked. But that would be a mistake. The "distinctively retributivist" thesis, he insisted—broadly in accord with (R)—is "that the guilty receiving their just deserts is an intrinsic good."¹⁵

Yet, as we have seen, Moore also contends that we have a duty or obligation both to inflict punishment on wrongdoers and to establish institutions whose purpose or function is to mete out retributive punishment. It seems plain on the face of things that these claims go far beyond (R) or the view that Moore has described as "distinctively retributivist." Not plain, in my estimation, are the arguments that Moore advances for this bolder position.

Let us start with the dimension of force. The usual arguments for retributivism, including those advanced by Moore, depend heavily upon thought experiments.¹⁶ Common are two-world hypotheticals: Worlds A and B are identical in all respects except that, in A, some vicious wrongdoer suffers or is punished on account of his wrongs, whereas in B he lives happily ever after. If we stipulate that the punishment or suffering of the wrongdoer in World A does not generate any downstream consequences that further distinguish the two worlds, and if we think that World A is nonetheless better, then we have grounds to accept something like (R). And, as Moore has further observed, your attachment to (R) might be strengthened by imagining that you are the wrongdoer and by then reflecting both on the guilt that you would yourself experience, and the judgment that you might reach that you too should be punished.¹⁷ If we believe that there is value in the satisfaction of negative desert, and if we further believe that, presumptively, we have some reason to bring about valuable

¹⁵ Moore, *Placing Blame*, at 157.

¹⁶ See, e.g., John Kleinig, *Punishment and Desert* (The Hague: Martinus Nijhoff, 1973), 67.

¹⁷ Moore, *Placing Blame*, at 145-49.

states of affairs, then we should conclude that we have reason to promote the satisfaction of negative desert.

Notice that I have spoken so far only in terms of value and reason. I have said nothing about duties, obligations, or justice, for the simple reason that nothing about hypotheticals of this kind naturally invites conclusions formulated in such terms. Moore, however, maintains that we have a “duty” or “obligation” to satisfy negative deserts, and that such a duty or obligation is one of “justice.”

Frankly, I am uncertain precisely what Moore means by these terms or what their normative upshot is. I assume only (thanks in part to the frequent use of italics) that Moore understands and intends to signal a normative escalation. On most accounts, duties and obligations are not only moral reasons but reasons that are particularly weighty or have second-order exclusionary characteristics, or are in other respects especially stringent (even if not “absolute” or “conclusive”). I therefore interpret the move from normative concepts like “reason,” “right,” “permissible” and “justified” to “duty,” “obligation,” and “justice” to raise the stakes, so to speak. Yet I do not see what about the thought experiments themselves warrants any normative escalation from claims about value and reason to claims about duties of justice, nor am I aware of distinct arguments that Moore provides for this move. We must be careful not to start by assuming that the justification for which we search will necessarily supply obligations and not merely reasons and permissions, or even that retributivism is a form of justice. That would be to put the rabbit in the hat.

To be sure, Moore could maintain, perhaps following Aristotle, that duty and justice are built into the very concept of desert. On one such view, for example, it is a conceptual entailment of the proposition that A deserves O that justice demands that A get O, and we are all under a standing obligation to advance justice. There are difficulties with this account of desert, for it is common, and not clearly mistaken, to assert in some contexts both that some A deserves ϕ , and that nobody is obligated

to see that A gets ϕ .¹⁸ But even if the contention that justice and obligation are built into the concept of desert can be successfully defended, it would remain for Moore to establish that the reactions that the retributivists' two-world hypotheticals frequently pump go beyond judgments that it is good or right that the blameworthy wrongdoers at issue "suffer" (or "are punished"), and really amount to judgments that the wrongdoers "deserve" to suffer (or to be punished) in the particular sense of *desert* that entails that others are obligated to furnish such suffering (or punishment).¹⁹ This cannot be taken for granted.

B. Scope

So much for my doubts about the robust retributivist position on the dimension of force. Suppose, however, that this aspect of robust retributivism is correct even if not definitional of retributivism: we, or some of us, have a duty to inflict deserved suffering or deserved punishment on wrongdoers. Does it follow that we have a duty to establish and maintain institutions of criminal punishment (still assuming that duties are especially stringent in some fashion yet to be fully elucidated)? Again, it is hard to know for sure without knowing what such a duty amounts to. But it is important to note that the thought experiments upon which retributivists rely do not require that conclusion: they are generally about individual cases.

More significantly, the focus on individual cases appears nonaccidental once we reflect on what retributivist theories of punishment are (or at least can be) designed or offered to do. Some commentators, perhaps influenced by Hart's insistence on distinguishing between the "general justifying aim" of punishment and principles of distribution, assume that theories of punishment must be designed to justify the institution of criminal punishment and not just individual inflictions of

¹⁸ For elaboration, see Mitchell N. Berman, "Rehabilitating Retributivism," 32 *Law & Phil.* 83 (2013), 90-93.

¹⁹ This is an appropriate place to register my own tentative sense that prevailing views regarding the sort of thing that desert is might be mistaken. Husak describes desert as a "reason-giving property." Husak, *What Do Criminals Deserve?* (ms. at 7). Moore often says similar things, see Moore, *Placing Blame*, at 168 n.25, 178, 180, though elsewhere he describes desert as a "principle." *Id.* at 170, 171. I am disposed to doubt that desert is best conceived as a property belonging to individuals. But because this is not the place to elaborate on my doubts, I have tried in this paper to speak of desert in terms that square with common understandings and usage. Some things I say in this paper—possibly including my formulation of (R)—could require restatement to accord with the conception of desert (negative or otherwise) that I might ultimately favor.

punishment. I think this is a consequentialist bias, and that the argumentative dialectic concerning the justification of punishment does not require retributivists to have been arguing that the institution of punishment can be justified by reference to the goodness or rightness of furnishing deserved suffering.

I have argued elsewhere that it is useful to distinguish two types of justification: all-things-considered, and “tailored.”²⁰ An all-things-considered justification of an act or practice is just what it sounds like: it purports to establish that the act or practice is morally permitted in light of all considerations. A tailored justification aims to establish the permissibility of the act or practice at issue against one or more grounds for doubt, what I have termed “demand bases.” It shows that the particular grounds explicitly invoked or implicitly relied upon for thinking the act or practice unjustifiable are chimerical or overridden.

In my view, the dominant strand within punishment theory has involved the effort to provide tailored justifications against the particularly salient and forceful complaint that it is wrongful to intentionally cause people to suffer. (R) is a promising response to this demand basis. As such, it can be sensibly advanced to justify punishment in the individual case. The argument can be construed thusly: assuming that there already exists an institution with lawful authority to inflict punishments (or that the instant case can, for whatever reason, be treated as a one-off that does not require a preexisting institution dedicated to the infliction of criminal punishment),²¹ then the fact that a person has done grave wrong provides a fully sufficient reason to inflict punishment, regardless of whether punishment in this case would produce any of the goods, or serve any of the values that the institution was created to serve.

When we turn attention to the entire institution of punishment, however, other objections—additional demand bases—become salient, indeed inescapable. Here are two: the institution of punishment consumes substantial social resources that could be put to other uses, and invariably results

²⁰ Mitchell N. Berman, “Punishment and Justification,” 118 *Ethics* 258 (2008), 262-66.

²¹ Think of a Nuremberg-like situation, or a domestic equivalent.

in the infliction of suffering on wholly innocent persons. If I am right that these objections are squarely on the table in the institutional case but not necessarily in the individual case, then the contention that we have an obligation to establish institutions of criminal punishment solely on the strength of the value of realizing negative deserts seems both doubtful and unnecessary to attribute to retributivists, even though modest retributivists should surely believe that the noninstrumental value of satisfying negative deserts *contributes* to a successful justification of the institution.

Some readers may feel that the account I have just offered regarding why a successful retributivist justification for instances of criminal punishment need not extend to or entail a successful retributivist justification for the institution of criminal punishment depends too heavily on a particular contestable account of the dialectical logic of justification. I think, in contrast, that it accurately captures the truth that distinct questions are sometimes in play. Very frequently, we ask the “what justifies punishment?” question when we have in mind some particular case of actual or contemplated punishment. Especially when the familiar consequentialist justifications for punishment strike us as significantly attenuated (as when the offense charged is sufficiently unusual as to weaken the apparent need for general deterrence, and when the offender is sufficiently old as to weaken the need for specific deterrence or incapacitation), the question is posed “what reason have we to punish *him*?” In such cases, the retributivist response is: “because he deserves it.” That can be an acceptable answer to the question on the table (if perhaps a little telegraphic) even if it is not an acceptable answer to the broader question: “what justifies the creation and maintenance of a system of criminal punishment?”²²

²² Husak has recognized that these are different questions and has agreed that negative desert alone goes farther toward answering the former than the latter. See Douglas Husak, “Holistic Retributivism,” 88 *Cal. L. Rev.* 991 (2000), 995. But he seems also to believe that the latter question—concerning a “complete justification for creating an institution of punishment,” *id.* at 996—is, in some meaningful sense, the real question, and therefore that retributivism gains greater plausibility than it genuinely warrants “only if we *artificially narrow* our conception of the nature” of the problem that we are called upon to resolve. *Id.* at 993 (emphasis added). Samuel Sheffler has taken issue with this suggestion of Husak’s, insisting on the independent significance of “the question of whether and when society’s punishment of an individual is compatible with just treatment of that individual.” Samuel Sheffler, “Justice and Desert in Liberal Theory,” 88 *Cal. L. Rev.* 965 (2000), , 987 n.73. I’m with Sheffler on this point.

C. Exclusivity

According to Moore, retributivism maintains that negative desert is a necessary and sufficient condition for infliction of punishment and that, equivalently, “punishment is justified if and only if the person or persons receiving it deserve it.”²³ Thus, he says, “retributivism asserts that punishment is properly inflicted because, and only because, the person deserves it.”²⁴ I believe this is doubly mistaken.

Perhaps the most obvious objection would be that an offender’s negative desert does not constitute a sufficient condition for the justified infliction of punishment. It is always imaginable that some individual act of punishment would produce horrendous consequences overall and it seems implausible that, even if we have moral reason, or even obligation, to inflict deserved punishment, such reasons or obligation are so stringent as to require that we do so regardless of the consequences. This, however, is one objection I waive, for Moore has made clear, in response to Dolinko’s leveling of this very charge, that his invocation of sufficiency is intended to be context-sensitive:

when we say that one condition was sufficient for another—as in the counterfactual statement that a particular fire was sufficient to burn down someone’s house—we mean that within some limited set of conditions this one by itself was sufficient. Other conditions outside that set—such as the presence of oxygen in my fire example—are invariably *necessary* even while we idiomatically describe a condition within that set as ‘sufficient’.

The retributivist’s usage of ‘sufficiency’ is to be so construed. *Within the set of conditions constituting intelligible reasons to punish*, the retributivist asserts, desert is sufficient, i.e., no other *of these* conditions is necessary.²⁵

Given this helpful clarification about Moore’s sufficiency claim, my two objections are these.

First, embrace of (R) does not entail that desert is a necessary condition for the infliction of punishment.

Second, even if negative desert is both a necessary and sufficient condition for the justified infliction of punishment (so long as sufficiency is construed in the somewhat loose manner just described), that does

²³ Moore, *Placing Blame*, at 92.

²⁴ *Id.*

²⁵ *Id.* at 173.

not establish that punishment may be inflicted “because and only because” the person punished deserves it, on the most common construal of that expression.

1. *Retributivism need not be committed to the proposition that negative desert furnishes a necessary condition for the justified infliction of punishment.* Let me say more about the argumentative dialectic that births not only retributivism but consequentialism too. We hold out for consideration an actual or hypothetical act of criminal punishment and raise a moral objection: *this*, we say, is deeply morally problematic. It is ordinarily profoundly wrongful to treat people in this way. Therefore, if we are to continue to do such things, we need to explain how it can be morally justified. Our theories of punishment are efforts to meet this demand. One tradition says, in effect, we can justify doing such things by reference to the fact that wrongdoers deserve it. Another tradition offers a different account: we can justify doing such things by the good net consequences thereby produced. Abolitionists believe that both positions are wrong, and that there are no other better options waiting in the wings. But just as both accounts could be wrong, both could be right. More significantly, there is no reason why it must be part of either of these proposed answers to reject the other. To be sure, one who accepts one of the justificatory theories could, at the same time, reject the other. A retributivist could be an anti-consequentialist about punishment, and a consequentialist could be an anti-retributivist. My point is only that both types of theory are, by the nature of what they are asked to do, accounts about what does or can justify punishment; they are not accounts about what does not or cannot justify it. Therefore, the rejection of consequentialist justifications is no integral part of retributivism, and the rejection of retributivist justifications is no integral part of consequentialism.²⁶ If this is right, Moore is

²⁶ I am speaking here of consequentialism as a justification of punishment, not as a comprehensive moral theory. As a comprehensive moral theory, consequentialism does deny deontological retributivism, but because we have already seen that consequentialist retributivism is a real possibility, it is not even true that consequentialism as a comprehensive moral theory entails the falsity of all varieties of retributivism.

simply mistaken to define retributivism as the view that “punishment is justified if and only if the person or persons receiving it deserve it.”²⁷

Of course, even if I am right that Moore’s “necessary condition” claim attributes more content to retributivism *simpliciter* than is warranted, it could still be true that punishment is justified only when deserved.²⁸ Most criminal law theorists espouse that view and they could well be correct. If they are correct, however, it is on the strength of a substantive moral judgment to the effect that (to a first approximation) it is unjust or otherwise wrongful to inflict punishment on somebody who doesn’t deserve it—or, in a more subjectivist spirit, without a bona fide belief that the person is guilty and blameworthy. Call that substantive moral claim (Q). But just as consequentialists about punishment may help themselves to (Q) (witness “side-constrained consequentialism”), retributivists—i.e., those who accept (R)—are free to eschew it. (R) does not entail (Q).

2. *Retributivism need not be committed to the proposition that punishment is justified only when inflicted “because” the person punished deserves it.* Now let us suppose, arguendo, that (R) and (Q) are both true. If so, we might fairly conclude that negative desert constitutes both a sufficient and a necessary condition for the justified infliction of punishment (so long as sufficiency is construed in a sufficiently contextual or pragmatic way), which is also to say that punishment is justified if and only if the person punished deserves it. Nonetheless, it would not follow that persons may be punished “because and only because” they deserve it, if that locution means that punishment is justified only when inflicted *for the reason* that the person punished deserves it. Conditions need not be reasons, for the absence of a fact that, if present, would render some action ϕ impermissible, is a necessary condition for the permissibility of ϕ ing, but not itself a reason to ϕ .²⁹ That you have not promised not to

²⁷ Moore, *Placing Blame*, at 91. See also *id.*, at 92.

²⁸ Or, only when inflicted in the good faith belief that it is deserved, or something similar. See note 4 above.

²⁹ I think that the distinction between conditions and reasons lies close to the heart of the difference between what Dolinko calls “rational” and “moral” justification, see David Dolinko, “Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment,” 16 *Law & Phil.* 507 (1997), 518, and what Michael Lessnoff termed

quack like a duck might be a necessary condition for you to be justified in quacking but is not itself a reason for you to quack. And if you do quack, it would be unnatural to say that you did so “because of” the fact that you had not promised not to.

So the fact, if true, that a person’s negative desert constitutes both a necessary condition and a sufficient condition for punishing him does not itself entail that we are justified in punishing somebody because and only because he has negative desert. But, of course, the latter claim could nonetheless be true. Are there any reasons for a retributivist (i.e., somebody who endorses (R)) to resist Moore’s contention that punishment is justified only when inflicted “because” the punishee deserves it?

Here are two. First, one might believe, in a Kantian vein, that the value of an action depends upon the actor’s motivations or explanatory reasons. Indeed, I think that Moore himself believes this.³⁰ Adding this view to Moore’s claim that, in order for punishment of a wrongdoer to be morally justified, the offender’s negative desert must supply the reason that punishment is inflicted yields a conclusion that a retributivist might not want to own. Take the standard imagined case of punishment in which the person punished deserves it and the punishment would promote net good consequences (even putting aside any supposed value of realizing negative desert). If we follow Moore (as I construe him) in believing that, in order for punishment in such a case to be morally justified, the offender’s negative

“teleological” and “entitling” justification. Michael Lessnoff, “Two Justifications of Punishment,” 21 *Phil. Q.* 141 (1971). The distinction is also fundamental to rights-forfeiture theories of punishment. See, e.g., Christopher Heath Wellman, “The Rights Forfeiture Theory of Punishment,” 122 *Ethics* 371 (2012).

³⁰ At one place, Moore comments on the “tension that exists between crime-prevention and retributive goals. . . . due to retributivism’s inability to share the stage with any other punishment goal. To achieve retributive justice, the punishment must be inflicted because the offender did the offense. To the extent that someone is punished for reasons other than that he deserves to be punished, retributive justice is not achieved.” If “part of the reason motivating those who punish [a wrongdoer] is that his punishment will serve as an example to others,” he says, then “the criminal justice system is not achieving retributive justice.” Moore, *Placing Blame*, at 28.

Elsewhere, however, Moore writes that “the only difference” between the two most plausible versions of mixed theory—desert-constrained consequentialism and consequences-constrained retributivism—“is in the *motivations* of those who hold them. And while that may make a difference in our moral judgements of those who hold the different branches of the mixed theory of punishment, it does not make a difference in terms of the actual social institutions and judgements such theories will justify.” *Id.* at 94 (emphasis in original). These passages strike me as inconsistent insofar as the second appears to convert a consideration that the first treats as bearing on act evaluations into a matter relevant only to actor evaluations. Still, I would not be so foolhardy as to rule out that Moore has in mind a basis for reconciling these passages that eludes me.

desert must supply the reason for punishment, then the value—and possibly the justifiability—of punishment will depend upon the reasons that move the punishing authorities. A retributivist could avoid this conclusion by refusing to join Moore in traveling from “if and only if” to “because and only because.”

Second, there is a cost to contending that punishment is justified only if imposed because—i.e., for the reason that—wrongdoers deserve it. Such a claim implies that if, contrary to what the retributivist believes, (R) is false, then punishment is not morally justifiable: if the only reason that supposedly favors punishment turns out to be no reason at all, then there are no morally cognizable reasons to punish. This is a steep price to pay for membership in the retributivist camp. For myself, I am more confident that morally justified systems, and instances, of criminal punishment are possible than I am that (R) is true.

IV. RETIBUTIVISM FROM THE BACK DOOR, AGAIN

In his well-known “argument for retributivism from the back door,” excerpted in several editions of the highly regarded Kadish & Schulhofer casebook,³¹ Moore defends the correctness of retributivism precisely by purporting to have shown that the only “theories of punishment truly competitive with retributivism, namely, the pure utilitarian theory, and the mixed theory, are each unacceptable to us.”³² But if retributivism supplies the only acceptable justification of criminal punishment, what does it maintain? I have argued that retributivism is properly understood as committed to something very close

³¹ See, for example, Sanford H. Kadish & Stephen J. Schulhofer, *Criminal Law and Its Processes: Cases and Materials*, 5th ed. (Boston: Little, Brown, 1989), 142-44.

³² Moore, *Placing Blame*, at 103. Some philosophers of punishment will reject the starting assumption that consequentialist, retributivist, and mixed theories exhaust the possible justifications of punishment. Additional possibilities would be deontological accounts that do not assume, and might even deny, that the suffering that a wrongdoer experiences is either good or right. Duff’s communicative theory is plausibly one such view, as is the view that we act rightfully in (temporarily or permanently) removing wrongdoers from law-abiding society. With many others, I am skeptical that the non-retributivist deontological positions on offer have the resources to justify treatment that rightly qualifies as punishment, see my “The Justification of Punishment,” in Andrei Marmor (ed.), *The Routledge Companion to Philosophy of Law* (New York: Routledge, 2012), 141, but we can safely table that question for the limited purposes of this essay.

to (R), and that the version of retributivism that Moore himself embraces goes significantly beyond this minimal content, on at least three discrete dimensions of variability.

Have I, at the same time, done anything to bolster the prospects for retributivism of any form? Not directly, for I offer no new arguments for (R) itself. Perhaps, though, my effort to identify respects in which Moorean retributivism goes beyond minimal retributivist commitments can serve as an indirect argument for modest retributivism, or something in its vicinity. Readers who had been disposed against Moorean robust retributivism, but had not appreciated the availability of accounts that claim much less for negative desert, might now be moved to reconsider their resistance to retributivism *tout court*. This argument, then, like Moore's, proceeds through the back door.

But is it a back door to the same house? Because modest retributivism is so unbold—perhaps even timid or pallid?—one might wonder whether it fairly qualifies as a version of retributivism at all.³³ It is impossible to address this concern fully without wading deeply into methodological debates concerning the demarcation and analysis of concepts. In concluding that what I am terming modest retributivism is indeed a version of retributivism, I rely heavily on my impressions of the criteria that criminal law theorists seem to invoke or have in mind when characterizing their own views, or even themselves, as retributivist or anti-retributivist. As a descriptive matter, it appears to me that the line between avowed retributivists and their avowed critics is most faithfully drawn between those who believe and those who deny that a wrongdoer morally deserves to suffer (or to experience hardship or deprivation, or the like) in any sense of desert that would provide some others, including the state, with

³³ Duff describes “the core retributivist thought” differently than I do. The core thought, in his view, is “that what gives criminal punishment its meaning and the core of its normative justification is its relationship, not to any contingent future benefits that it might bring, but to the past crime for which it is imposed.” R.A. Duff, “Retrieving Retributivism,” in Mark D. White (ed.), *Retributivism: Essays on Theory and Policy* (New York: Oxford, 2011), 3. I do not think that an account need view desert as supplying the “core normative justification” for punishment for it to qualify as retributivist.

reasons—nontrivial reasons, or reasons that can at least sometimes make a difference—to provide that suffering or deprivation.³⁴

(R) aims to capture the proposition that self-described retributivists overwhelmingly accept and that their self-described opponents overwhelmingly reject. But the pith of retributivism could perhaps be put in simpler and more colloquial terms: *Ceteris paribus*, the world goes better when wrongdoers receive their comeuppance.

³⁴ Moore once opined that “[t]he true worry that most people have about retributivism” is that retributivist emotions, urges, and judgments are always pathological. Moore, *Placing Blame*, at 119. I take this to be offered as an explanation for many people’s rejection of the central retributivist thesis. My experience jibes with Moore’s observation.