2002

Deterring Retributivism: The Injustice of 'Just' Punishment

Russell Christopher

Follow this and additional works at: http://digitalcommons.law.utulsa.edu/fac_pub

Part of the Criminal Law Commons

Reprinted by special permission of Northwestern University School of Law, Northwestern University Law Review.

Recommended Citation

This Article is brought to you for free and open access by TU Law Digital Commons. It has been accepted for inclusion in Articles, Chapters in Books and Other Contributions to Scholarly Works by an authorized administrator of TU Law Digital Commons. For more information, please contact daniel-bell@utulsa.edu.
Articles

DETERRING RETRIBUTIVISM: THE INJUSTICE OF "JUST" PUNISHMENT

Russell L. Christopher*

Introduction ........................................................................................................... 845
I. Fundamentals of Punishment Theory .............................................................. 852
   A. Retributive Versus Consequential Theories of Punishment .................... 855
   B. Varieties of Retributivism ..................................................................... 865
   C. The Mixed Theory of Punishment ......................................................... 867
II. Retributivism’s Intentional Punishment of the Innocent ............................. 869
   A. The Debate over Intentional Punishment of the Innocent .................. 870
      1. The Retributivist Charge: Consequentialism Justifies the Intentional Punishment of the Innocent ................................................. 870
      2. Consequentialist Responses .............................................................. 871
         a. Denial ............................................................................................. 871
         b. Acceptance with a Defense .............................................................. 873
         c. Retributivism Does It Too ............................................................... 877
   B. The Moralistic Version of Retributivism ................................................ 880
   C. The Legalistic Version of Retributivism ................................................ 885
   D. Objection .............................................................................................. 886
   E. Conclusion .............................................................................................. 887
III. Retributivism’s Inability to Provide a Principled Account of
     the Degree of Desert ................................................................................... 888
   A. The Moralistic and Legalistic Combined Version of Retributivism ...... 889
   B. The Retributivist Claim That Only Retributivism Can Provide
      a Principled Determination of the Degree of Desert ........................... 890
   C. Combined Retributivism Supplies Contradictory Determinations of the
      Degree of Desert .................................................................................... 894
   D. Conclusion .............................................................................................. 897
IV. Retributivism’s Indirect Punishment of the Innocent ................................. 898
   A. The Debate over Mistaken Punishment of the Innocent and the
      Doctrine of Double Effect ..................................................................... 899

* Visiting Assistant Professor of Law, Florida State University, College of Law (Spring 2002); Assistant Professor of Law, University of Tulsa, College of Law (Fall 2002). I am indebted to George Fletcher, John Gardner, Kent Greenawalt, Ken Levy, and Herb Morris for their criticisms of earlier drafts of this Article. I also benefited from discussions with Mike Dorf, Maria Pagano, and Craig Williams. Any errors are, of course, my own.
1. The Consequentialist Charge: Retributivism Is Unjust by Mistakenly Punishing the Innocent ................................................................. 899

2. The Retributivist Response: Invocation of the Doctrine of Double Effect ................................................................. 900

B. Retributivism's Punishment of the Innocent by Substantive Means .......... 904
   1. Retributivism’s Justification of Punishment Under Strict Liability .......... 904
   2. Retributivism’s Justification of Punishment Under an Extreme Form of Absolute Liability ......................................................... 908

C. Retributivism’s Punishment of the Innocent by Procedural Means ....... 909
   1. Retributivism’s Justification of Punishment Under a Presumption of Guilt Standard of Proof ......................................................... 910
   2. Retributivism’s Justification of Punishment By Guilty Verdicts from a Non-Majority of Jurors ......................................................... 915

D. Objections .................................................................................................................. 917
   1. Efforts to Avoid Punishment of the Innocent ........................................... 918
   2. Reasonable Efforts to Avoid Punishment of the Innocent ..................... 919
   3. Theory Versus Application .................................................................... 921

E. Conclusion .................................................................................................................. 923

V. Retributivism Uses Persons Not as Ends but as Mere Means ............... 924

A. Means and Ends ....................................................................................................... 925
   1. The Retributivist Charge: Consequentialism Uses Offenders as Mere Means .................................................................................... 926
   2. Accounts of Using Persons as Mere Means ............................................. 927
   3. Consequentialist Responses .................................................................... 930

B. Retributivism Uses Crime Victims as Mere Means .................................... 933
   1. The Necessary Use of Victims as Means .................................................. 934
   2. The Irrelevance of the Interests of Victims as Ends ............................... 937
   3. The Use of Victims as Mere Means ......................................................... 939
   4. Can Retributivism Avoid the Claim That It Uses Victims as Mere Means? ....................................................................................... 940

C. Objections .................................................................................................................. 941
   1. Kant’s Maxim Does Not, and Should Not, Apply to Retributivism’s Use of Crime Victims ............................................................... 942
   2. Retributivism Does Not Intentionally Use Crime Victims as Mere Means ....................................................................................... 943
   3. Retributivism Uses Crime Victims as Means But Not as Mere Means .... 944
   4. Retributivism Does Not Use Crime Victims at All .................................. 947

D. A Brief Digression: The Tension Within Retributivism over Victims’ Rights .................................................................................. 950

E. Conclusion .................................................................................................................. 952

VI. Conceptually Consequential Retributivism’s Punishment of the Innocent and Use of Offenders as Mere Means ................................. 953

A. Conceptually Consequential Retributivism ............................................... 954
   1. Why Use Consequences? The Retributivism Dilemma:
INTRODUCTION

Retributivism is all the rage. Whether it is a "revival," a "resurgence," or a "renaissance," retributivism's rapid rise since the early

1 Retributivism is a theory, or justification, of punishment. Though a precise definition of retributivism has proven elusive, stated most simply, the theory holds that punishment is justified solely because the person being punished deserves it. As one commentator observed:

What we may call the retributive view is that punishment is justified on the grounds that wrongdoing merits punishment. It is morally fitting that a person who does wrong should suffer in proportion to his wrongdoing. That a criminal should be punished follows from his guilt, and the severity of the appropriate punishment depends on the depravity of his act. The state of affairs where a wrongdoer suffers punishment is morally better than the state of affairs where he does not; and it is better irrespective of any of the consequences of punishing him.


Other commentators locate the common denominator of all retributivist accounts in retributivism's relation to a past wrong or offense. See, e.g., R.A. Duff, Trials and Punishments 4 (1986) ("[A]ll retributivist theories find the sense and justification of punishment in its relation to a past offence"); George Fletcher, Rethinking Criminal Law 416-417 (1978) ("Retribution simply means that punishment is justified by virtue of its relationship to the offense that has been committed.").

Igor Primoratz notes that "in its most complete form," retributivism contains five tenets:

1. The moral right to punish is based solely on the offense committed.
2. The moral duty to punish is also grounded exclusively on the offense committed.
3. Punishment ought to be proportionate to the offense (the lex talionis).
4. Punishment is the "annulment" of the offense.
5. Punishment is a right of the offender.


2 See, e.g., David Dolinko, Three Mistakes of Retributivism, 39 UCLA L. Rev. 1623, 1623 (1992) ("It is widely acknowledged that retributivism, once treated as an irrational vestige of benighted times, has enjoyed in recent years so vigorous a revival.""); R.A. Duff, In Defence of One Type of Retributivism: A Reply to Bagaric and Amarasekara, 24 MELb. U. L. Rev. 411, 411 (2000) ("A striking feature of penal philosophising during the last thirty years has been the revival of retributivism.").

1970s\(^6\) has been remarkable. The U.S. Supreme Court,\(^7\) state courts,\(^8\) state legislatures,\(^9\) philosophers, and legal scholars\(^10\) alike are increasingly ac-

---


\(^{4}\) Martin R. Gardner, *The Renaissance of Retribution: An Examination of Doing Justice*, 1976 Wis. L. Rev. 781, 784 (footnote omitted) (noting that "retribution is suddenly being seen by thinkers of all political persuasions as perhaps the strongest ground, after all, upon which to base a system of punishment").

\(^{5}\) Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. Rev. 1659, 1659 (1992) ("There has been a steady rise in the popularity of retributivism over the last decade, which is surprising given its near death in the 1950s and 1960s."). For the contrary view—that retributivism is not the currently favored theory of punishment—see Huigens, *supra* note 3, who comments: "The deterrence theory of punishment is probably dominant over retributivism in criminal law scholarship, though not as much as it used to be, because consequentialism, and economics in particular, has greater credibility in our society than any comprehensive morality that might undergird a retributive theory." *Id.* at 955 (footnote omitted).


\(^{7}\) See, e.g., Spaziano v. Florida, 468 U.S. 447, 461–62 (1984) (declaring that retributivism "is an element of all punishments society imposes" and enthroning it as the "primary justification of the death penalty"); Gregg v. Georgia, 428 U.S. 153, 183–84 (1976) (Stewart, Powell, and Stevens, JJ., announcing the judgement of the Court) (acknowledging that deterrence was a permissible partial purpose of punishment, but explaining, in effect, that retributivism was an important purpose of capital punishment).


\(^{9}\) See, e.g., *Cal. Penal Code* § 1170(a)(1) (West 1985) ("The legislature finds and declares that the purpose of imprisonment for crime is punishment."). 204 PA. CODE § 303.11 (2001) ("The sentencing guidelines provide sanctions proportionate to the severity of the crime [which] ... establishes a sentenc-
knowing retributivism as the dominant theory of punishment.\textsuperscript{11} Even its critics acknowledge that retributivism "can fairly be regarded as the leading philosophical justification for the institution of criminal punishment."\textsuperscript{12}

Putting the rhetorical excesses of retributivism’s ascendancy aside, what is truly striking is that a theory of punishment born out of the harsh and rigid justice\textsuperscript{13} of the Old Testament\textsuperscript{14} has been transformed into a theory that stakes out the higher moral ground.\textsuperscript{15} Simply put, retributivism justifies punishment, or the suffering by the punished, not on any actual good
consequences that might be attained, but solely because the punished deserve it.\footnote{See supra note 1.} From its roots in vengeance, bloodlust, revenge, retaliation, and an eye for an eye,\footnote{The biblical injunction of taking an eye for an eye, which appears to be harsh and severe justice from our modern perspective, was comparatively quite mild in its time. The lex talionis “did not encourage, but rather restrained the vengefulness of the wrong.” PRIMORATZ, supra note 1, at 87; see also Andrew Oldenquist, Retributive Rationale, in \textit{The Philosophy of Law: An Encyclopedia} 749 (Christopher Berry Gray ed., 1999) (“There is reason to think this was the voice of the soft-hearted in biblical times: one can take only one eye for an eye, only one life for a life (and not the criminal’s family).”) (emphasis added).} retributivism is pitched as the only theory which, in justifying punishment, does justice by (i) not justifying punishment of the innocent, (ii) not using persons as mere means to attain other goals, (iii) giving a principled account of the requisite degree of desert and punishment, and (iv) justifying punishment of only those who are morally culpable.\footnote{See Duff, supra note 6, at 25 (retributivists’ central claim is that only a retributivist conception of punishment can do justice to the guilty, as well as to the innocent”); see also infra note 23.}

The principal alternative conception of punishment—consequentialism—is claimed by retributivists to be morally inferior to retributivism on these four counts. 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. Until the retributivist revival, consequentialist theories of punishment had enjoyed the mantle of being humane, merciful, rational, enlightened, and morally sensitive from Plato onward. Retributivism was dismissed as irrational, “taking blind vengeance like a beast.”\footnote{PLATO, \textit{Protagoras}, in \textit{The Collected Dialogues of Plato} 309, 321 (Edith Hamilton & Huntington Cairns eds., 1973). The quote appears in the following passage:}

\begin{quote}
In punishing wrongdoers, no one concentrates on the fact that a man has done wrong in the past, or punishes him on that account, unless taking blind vengeance like a beast. No, punishment is not inflicted by a rational man for the sake of the crime that has been committed—after all one cannot undo what is past . . . .
\end{quote}

\textit{Id.}

\footnote{For St. Thomas Aquinas, punishment, or as he termed it vengeance, without any intention that good consequences would result, which is in essence retributivism, is unlawful and immoral:}

\begin{quote}
Vengeance consists in the infliction of a penal evil on one who has sinned. Accordingly, in the matter of vengeance, we must consider the mind of the avenger. For if his intention is directed chiefly to the evil of the person on whom he takes vengeance, and rests there, then his vengeance is entirely unlawful: because to take pleasure in another’s evil belongs to hatred, which is contrary to the charity whereby we are bound to love all men. Nor is it an excuse that he intends the evil of one who has unjustly inflicted evil upon him, as neither is a man excused for hating one who hates him: for a man may not sin against another just because the latter has already sinned against him, since this is to be overcome by evil, which was forbidden by the Apostle, who says (Rom. Xii. 21): \textit{Be not overcome by evil, but overcome evil by good.} \end{quote}

2 \textit{ST. THOMAS AQUINAS, Summa Theologica} 1656 (Fathers of the English Dominican Prince trans.,
as nothing more than "a triumph, or glorying in the hurt of another . . . [which is] contrary to reason[,] . . . against the Law of Nature and is commonly stiled by the name of Cruelty." But the tables have now turned; it is retributivism which now claims to reflect our aspiration to higher moral principles and it is consequentialism that is widely considered morally unpalatable.

My project is not to defend a consequentialist theory of punishment, but to displace retributivism's place on the high moral ground. Just as retributivism is said to turn back the crime on the criminal, I will seek to turn back retributivism's principal criticisms of the consequentialist theory onto itself. As a result, I conclude that retributivism is no better than consequentialism as a theory of punishment.

And it is perhaps worse. By presenting new arguments demonstrating that retributivism fails to satisfy its own criteria of just punishment, the injustice of the "theory of just punishment" is revealed. By showing that retributivism falls victim to its own withering critique of other theories of punishment, retributivism succumbs to its own "Ishmael effect." That is,

---

1947) (located in part 2-2, question 108, 1st article). But punishment may be lawful if it is imposed with the intention that good consequences would result:

If, however, the avenger's intention be directed chiefly to some good, to be obtained by means of the punishment of the person who has sinned (for instance that the sinner may amend, or at least that he may be restrained and others be not disturbed, that justice may be upheld, and God honored), then vengeance may be lawful, provided other due circumstances be observed.

Id.

21 THOMAS HOBBES, LEVIATHAN 79 (Ch. 15) (J.M. Dent & Sons 1976) (1651). The quoted language is drawn from the following passage in which Hobbes sets forth the Seventh Law of Nature affirming that the purpose of punishment is its good consequences and rejecting retributivism as revenge:

Whereby we are forbidden to inflict punishment with any other designe, than for correction of the offender, or direction of others . . . . Besides, Revenge without respect to the Example, and profit to come, is a triumph, or glorying in the hurt of another, tending to no end; (for the End is always somewhat to Come) and glorying to no end, is vain-glory, and contrary to reason; and to hurt without reason, tendeth to the introduction of Warre; which is against the Law of Nature; and is commonly stiled by the name of Cruelty.

Id.

22 Fletcher, supra note 14, at 517; see G.W.F. HEGEL, ELEMENTS OF THE PHILOSOPHY OF RIGHT 129 (Allen W. Wood ed., H.B. Nisbet trans., 1991) (1821) (punishment is "the crime turned round against itself"); id. at 127 (through punishment "what the criminal has done should also happen to him"); IMMANUEL KANT, THE METAPHYSICS OF MORALS 169 (Mary Gregor trans., 1991) (1797) ("For the only time a criminal cannot complain that a wrong is done to him [by punishment] is when he brings his evil deed back upon himself, and what is done to him in accordance with penal law is what he has perpetrated on others . . . .").

23 The "Ishmael effect" is a philosophical term meaning "[t]he claimed ability of some philosophical theory to escape from the fate to which it condemns all other discourse." SIMON BLACKBURN, OXFORD DICTIONARY OF PHILOSOPHY 200 (1996). The term is named after Ishmael, the character and narrator of Herman Melville's Moby Dick. In an epilogue Ishmael informs the reader, after telling the tale, "‘‘and I only am escaped alone to tell thee.’’" Id.

Retributivists typically condemn all theories of punishment for failing to do justice in punishing wrongdoers; all, that is, except for retributivism. See, e.g., Murphy, supra note 6, at 7 ("I believe that retributivism can be formulated in such a way that it is the only morally defensible theory of punish-
retributivism is incoherent. In contrast, consequentialism, while undeniably flawed, may be true to itself.

This age-old debate over the justification of punishment is not merely of theoretical interest. The resolution of a wide-ranging spectrum of practical issues may crucially hinge on, or be substantially influenced by, the particular justification of punishment perceived to be defensible. What is perceived to be the leading theory of punishment influences legislatures’ articulation of the purpose of punishment and courts’ construction of such statutory articulations, which in turn, affect a host of other practical doctrines, policies, and procedures. Prominent among these are the viability of the victims’ rights movement, the weakening of probation and parole, the “three strikes and you’re out” legislation aimed at recidivists, and perhaps most significantly, the morality of the death penalty. Capital punishment has traditionally been supported by, and associated with, retributivism.

ment.

See Introduction, in PUNISHMENT: A PHILOSOPHY & PUBLIC AFFAIRS READER, supra note 6, at vii ("The problem of justifying legal punishment has remained at the heart of legal and social philosophy from the very earliest recorded philosophical texts to the most recent.").

Consider the following examples:

[Judges sentence according to the purposes authorized by law. Where jurors determine a sentence, they may be instructed by the court and exhorted in closing arguments by attorneys to consider particular purposes in making their decision. In voir dire, attorneys may employ the state’s articulated purposes for punishment as a means of screening and influencing jurors. A defendant might challenge a particular punishment for failing to comport with the purposes articulated in a state constitutional provision. The articulated purposes thus affect jury selection, instructions, and arguments; sentencing and parole; and even assessment of the legitimacy of particular punishments.

Cotton, supra note 8, at 1317 (footnotes omitted).

Michael Moore, Victims and Retribution: A Reply to Professor Fletcher, 3 BUFF. CRIM. L. REV. 65, 67 (1999) ("[V]ictims should and must be ignored if you are claiming to be doing retributive theory.").


Louis Kaplow & Steven Shavell, Fairness Versus Welfare, 114 HARV. L. REV. 961, 1303-04 (2001) (footnote omitted) (suggesting that “three strikes and you’re out” policies “seem to be fueled by concerns about retribution, which are particularly sharp, many believe, because multiple recidivists have so clearly rejected society’s norms and institutions”).


[S]ince the mid-seventies retribution has come back with a vengeance, enjoying today a greater prominence in public discourse over crime and punishment than at any other time in post-war America.

Nowhere is this mood better demonstrated than in the debate over the death penalty. Where once capital punishment was reluctantly defended as a necessary deterrent to the most heinous of murders, it has, in recent years, been eagerly embraced by politicians eager to adopt a “tough on
Court's enthroning retributivism as the "primary justification for the death penalty."\(^3\)

After some introductory remarks on the fundamentals of punishment theory in Part I, Part II argues that two versions of retributivism justify the intentional punishment of particular, identifiable individuals who are either legally or morally innocent. In advocating the punishment of morally culpable wrongdoers, moralistic retributivism justifies the intentional punishment of legal innocents who have extrinsic defenses unrelated to, and which fail to negate, moral culpability. In supporting punishment of offenders found legally guilty, legalistic retributivism justifies the intentional punishment of moral innocents.

Part III considers the retributivist criticism that consequentialist theories fail to supply a principled rationale for determining the degree of an offender's desert and punishment in light of yet another version of retributivism—combined legalistic and moralistic retributivism. Though this version avoids justifying the intentional punishment of particular, identifiable innocents, it does so at the cost of being unable to supply a principled account of the degree of an offender's desert and punishment. In some instances, the combined version furnishes contradictory assessments of offenders' degree of desert and punishment.

Part IV demonstrates retributivism's justification of the indirect punishment of the innocent. Retributivism will be shown to justify the adoption of, and punishment under, standards and practices in which virtually every innocent defendant would be convicted and punished. These standards include strict liability for serious criminal offenses and a presumption of guilt standard of proof.

Part V inverts the standard retributivist criticism that consequentialist theories violate the Kantian maxim of using offenders as mere means by contending that retributivism perversely uses crime victims as mere means in order to respect offenders as ends in themselves. Stated simply, the argument is that since the use of crime victims is necessary as a means to attaining retributive justice, and that the interests of crime victims as ends in themselves are irrelevant, retributivism violates the Kantian duty to treat crime victims not as mere means but rather as autonomous ends.

Part VI addresses a large class of retributive theories that justify punishment by resorting to consequences. To avoid the circularity stemming from the simple retributivist formula of "it's right to punish criminals because doing so is right,"\(^3\)\(^2\) some versions of retributivism seek to justify punishment in terms of the good consequences that might be promoted. In

\(^2\) Dolinko, supra note 12, at 507.
turn, to avoid their theory collapsing into consequentialism, retributivists
argue that the consequences are only conceptual, rather than the actual
sequences found in consequentialist theories. Even if the use of merely
conceptual consequences suffices to prevent the collapse into
consequentialism, the reliance on consequences renders this class of
retributivism subject to the same problems as consequentialist theories—
tentional punishment of the innocent and the use of offenders as mere
means. I conclude that retributivism does not deserve its self-anointed place on
the high moral ground. Rather, like what retributivism claims criminals de-
serve, retributivism deserves to suffer (in our esteem). Because retributiv-
ism is subject to the same criticisms it levels against consequentialist
theories, retributivism is no better a justification of punishment than conse-
quentialist theories. And it is perhaps worse. By failing to satisfy its own
principles, retributivism is incoherent.

I. FUNDAMENTALS OF PUNISHMENT THEORY

Why do we need to justify punishment? After all, crime and punish-
ment seem to go together, as the old Frank Sinatra song goes, like “love and
marriage” and a “horse and carriage.” To doubt the legitimacy of pun-
ishment would seem to cast doubt on the enterprise of criminal law itself.
Why would we bother to promulgate the prohibitions of the criminal law if
they could be violated with impunity? For violations of the norms of crimi-
inal law, punishment seems to be an obviously fitting response. But pun-
ishment does require justification, for the same reason we consider
conduct violating the core prohibitions of our criminal law to be wrong.
Punishment involves the deliberate infliction of pain, suffering, and depri-
vation, which is prima facie wrong. So too, committing homicide or caus-
ing grievous bodily damage, under ordinary circumstances, is prima facie
wrong. But just as the prima facie wrong of homicide may be justified or
negated when committed under circumstances of self-defense as a response
to a criminal attack, so also the state’s infliction of the suffering and depri-

33 See “Love and Marriage” (lyrics by Sammy Cahn; music by Jimmy Van Heusen), in S. CAHN, I
SHOULD CARE: THE SAMMY CAHN STORY 298 (1974) (“Love and marriage... go together like a horse
and carriage. ... Try, try, try to separate them, it’s an illusion... Love and marriage... you can’t
have one without the other!”).
34 DUFF, supra note 1, at 1 (“It is agreed that a system of criminal punishment stands in need of
some strenuous and persuasive justification...”); Richard Wasserstrom, Why Punish the Guilty?, 20
PRINCETON U. MAG. 14 (1964), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT 328, 337
(Gertrude Ezorsky ed., 1972) (“Punishment is an evil, an unpleasantness; it requires that someone suffer.
Its infliction demands justification.”).
35 See, e.g., PRIMORATZ, supra note 1, at 7 (“To punish is to inflict an evil. But to inflict evil on
someone is something that, at least prima facie ought not to be done.”); C.L. TEN, CRIME, GUILT, AND
PUNISHMENT: A PHILOSOPHICAL INTRODUCTION 3 (1987) (“We are not normally justified in depriving
people of the things which they value, such as their liberty or their property.”); Murphy, supra note 6, at
9, 11 (punishment is a form of coercion by the state, which is prima facie wrong).
vation constituting punishment may be susceptible to justification as a response to the commission of a crime. Punishment might be conceived in this way as an institutionalized form of self-defense by the state against crime.

Among perhaps many, there is one obvious disanalogy between self-defense and punishment. Self-defense is used to prevent criminal aggression; it is employed before the fact. Punishment is employed after the fact (after the crime has already occurred) and thus cannot prevent the crime. But though punishment of a crime cannot prevent that crime from occurring, as self-defense might, it still might be preventive in nature. Punishment, by deterrence or incapacitation or reformation, may serve to prevent future crime. Or, from a retributive perspective, punishment may serve to prevent what a crime might otherwise signify—if a crime is a negation of the crime victim’s autonomy, punishment negates the negation.

Punishment’s justification, however, should be kept distinct from its definition. There have been some attempts to solve the justification issue by definition. For example, the retributivist thesis that an offender must

---


37 Forging a middle position between self-defense preventing a present or imminent crime and punishment preventing future crime, self-defense (in the state of nature) might serve the purpose of future deterrence even though it would not qualify as punishment:

It seems right to think that if we are in a state of nature, and you are about to cut my throat, then I have a privilege of causing you more harm than is necessary to defend my throat against your current assault on it: I have a privilege of causing you as much harm as is necessary to defend my throat against not merely your current but your otherwise likely future assaults on it. But punishment strikes me as different.


38 Viewing punishment in this way, by analogizing to self-defense, would seem to suggest that punishment be understood in consequentialist terms. But there are a number of aspects of self-defense that cannot necessarily be justified on consequentialist terms. Though the purpose of self-defense may be the prevention of criminal aggression, self-defense is sometimes conceived as a right that is good in itself without regard to beneficial consequences. The right to use self-defense is not limited to those situations where more harm will be avoided than caused. For example, lethal force in self-defense may be used to prevent non-lethal aggression such as rape or grievous bodily harm. Additionally, to save the life of one person from criminal aggression, a self-defender may kill 10, 100 or 1,000 aggressors.

39 See HEGEL, supra note 22, at 123 ("punishment is merely the negation of the negation"); id. at 127 ("The cancellation ... of crime is retribution in so far as the latter, by its concept, is an infringement of an infringement ... ").

40 See, e.g., HOBBES, supra note 21, at 164, Ch. 28 (emphasis omitted) ("A Punishment, is an Evill inflicted by publique Authority, on him that hath done, or omitted that which is Judged by the same Authority to be a Transgression of the Law; to the end that the will of men may thereby the better be disposed to obedience.").

Built into Hobbes’s definition of punishment is that it has the purpose of general deterrence:
be guilty to be punished is part of the meaning of punishment and consequentialism provides the justification. Thus, if punishment is defined so as to apply only to the guilty, then the problem of punishment of the innocent for consequentialist theories is simply defined away. H.L.A. Hart termed this argument the "definitional stop" and declared it impermissible as an "abuse of definition." Similarly, against the consequentialist criticism that retributive theories, even if they may justify punishment in a particular case, cannot justify the state's institution of a system of punishment, retributivists might reply that punishment by the state is part of the definition of punishment. This too would seem to be a "definitional stop" that simply defines away the substantive problem of justification.

It is the justification, rather than the definition, of punishment that is considered to be of, and will be our, primary interest. The issue of justification is often distinguished between justification of the punishment (i) in the particular case and, (ii) the general system of institutions of punishment.

If the harm inflicted be less than the benefit, or contentment that naturally followeth the crime committed, that harm is not within the definition [of punishment]; and is rather the Price, or Redemption, than the Punishment of a Crime: Because it is of the nature of Punishment, to have for end, the disposing of men to obey the Law . . . .

Id. at 166.

41 A.M. Quinton, On Punishment, 14 ANALYSIS 512 (1954), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 6, 10 ("The infliction of suffering on a person is only properly described as punishment if that person is guilty. The retributivist thesis, therefore, is not a moral doctrine, but an account of the meaning of the word 'punishment.'").

42 For an example of such an attempt to define away the issue of punishing the innocent see HOBBES, supra note 21, at 168 ("All Punishments of Innocent subjects, be they great or little, are against the Law of Nature: For Punishment is only for Transgression of the Law, and therefore there can be no Punishment of the Innocent.").

43 HART, supra note 11, at 6; Igor Primoratz, Mixed Rationales, in THE PHILOSOPHY OF LAW: AN ENCYCLOPEDIA, supra note 17, at 559, 559–560.

44 HART, supra note 11, at 5.

45 C.L. Ten, supra note 35, at 48 ("But whatever the intuitive force of the claim that wrongdoers deserve to suffer, it gives no support to the much stronger claim that the suffering should be deliberately inflicted by the State on wrongdoers. And it is this stronger claim which is needed to support the retributive theory of punishment."); Stanley I. Benn, Punishment, in 7 THE ENCYCLOPEDIA OF PHILOSOPHY 29, 30 (Paul Edwards ed. 1967, reprint ed. 1972) (noting that retributivism is more persuasive in justifying a particular instance of punishment than in justifying "punishment as an institution").

46 GEORGE FLETCHER, BASIC CONCEPTS OF CRIMINAL LAW 38 (1998) (acknowledging that Kant's "rationale for punishment hardly works to justify the institution of punishment at its initial historical stages").

47 For view finding punishment to be broader than punishment imposed by the state, see, for example, C.L. Ten, supra note 35, at 2 ("Punishment is administered not just by the State but also by others such as teachers and parents."); Kurt Baier, Is Punishment Retributive?, 16 ANALYSIS 25 (1955), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 16, 16–17 (noting that punishment could be imposed by a father or teacher); Benn, supra note 45, at 29 (noting that punishment includes that punishment which a parent gives to a child).

48 HART, supra note 11, at 6; Benn, supra note 45, at 29.

49 For the view that the definition of punishment plays an important role in the justification of punishment, see DUFF, supra note 1, at 151–53.
Although both retributivism and consequentialism purport to provide a justification on both levels, it is generally considered that retributivism provides a stronger justification on the first level and consequentialism a stronger justification on the second level. A number of “mixed” theories of punishment divide up the task of justification along these lines.

One commentator has declared that an adequate justification of punishment must provide answers to the following four questions: “Is the practice of punishment ever justifiable and if so under what conditions? What kinds of punishment are justified and must they involve suffering? Whom are we entitled to punish? Who is morally entitled to inflict punishment?”

In order to appreciate how retributivists’ criticisms of consequentialist theories of punishment may be turned back onto retributivism, it may be helpful to broadly sketch the fundamentals of the debate over the justification of punishment. This Part will canvass the principal versions of retributive and consequential theories and highlight the major moves and countermoves in the modern debate. Readers familiar with this debate might well choose to skip ahead to Part II.

A. Retributive Versus Consequential Theories of Punishment

The modern debate over the justification of punishment stems largely from the impasse between retributivism and some form of consequentialism. With retributivism being less susceptible to precise definition, it is best understood in opposition to consequentialist theories of punishment.

---

50 Benn, supra note 45, at 30–32.
51 See generally HART, supra note 11; Rawls, supra note 1. For further discussion of mixed theories, see infra notes 128–136 and accompanying text.
52 A. Wesley Cragg, Punishment, in THE PHILOSOPHY OF LAW: AN ENCYCLOPEDIA, supra note 17, at 706, 707.
53 C.L. Ten, supra note 29, at 366 (retributivism and utilitarianism or consequentialism are the “two main types of theories of punishment”); PRIMORATZ, supra note 1, at 9 (same); Oldenquist, supra note 17, at 749 (same); Quinton, supra note 41, at 6 (retributivism and consequentialism “exhaust the possibilities” for justifying punishment).
54 C.L. TEN, supra note 35, at 38 (“There is no complete agreement about what sorts of theories are retributive . . . .”). David Dolinko has argued:

it is curiously difficult to articulate [retributivism] in a perspicuous fashion. The central claim of the rival deterrence theory can be summarized swiftly—“it’s right to punish criminals because doing so minimizes the net level of suffering.” But correspondingly concise efforts to characterize retributivism risk collapsing into “it’s right to punish criminals because doing so is right.”

Dolinko, supra note 12, at 507.

Rather than try to define retributivism, some commentators seek to explain retributivism by recourse to a “cluster of moral concepts: rights, desert, merit, moral responsibility, and justice.” JEFFREY MURPHY & JULIE COLEMAN, PHILOSOPHY OF LAW: AN INTRODUCTION TO JURISPRUDENCE 121 (1990).
55 Joel Feinberg, The Classic Debate, in PHILOSOPHY OF LAW 646, 646 (Joel Feinberg & Hyman Gross eds., 1991); see also DUFF, supra note 1:

We can usefully talk of the common features, and the common logical structure, of consequentialist accounts of punishment: but can we usefully talk of retributivist accounts in this way; or has the label “retributivist” been applied to such a diversity of views and principles that it now lacks any
Consequentialism is the view that the value of an action or course of conduct is to be assessed from its consequences. A consequentialist theory of punishment would justify punishment on the basis of the good consequences promoted by punishment. Versions of consequentialism vary based on the type of consequence perceived to be relevant. 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.

unambiguous or unitary meaning?

Id. at 4 (footnote omitted); Bagoric & Amarasekara, supra note 12, at 129 (footnotes omitted) (the various versions of retributivism "do have one thing in common: they are not utilitarian... It may well be that the negation of utilitarianism is the distinctive badge worn by retributive theories."); Dolinko, supra note 12, at 507-08 (footnote omitted) ("One feature on which retributivists and their critics have generally agreed is that retributivism is very much a non-consequentialist theory.").


57 Jeremy Bentham's theory of punishment is the leading utilitarian theory of punishment and the principal influence on consequentialist views of punishment. PRIMORATZ, supra note 1, at 13 (claiming that "the most comprehensive and thoroughly developed" consequentialist account of punishment is Bentham's).

For general discussions of Bentham's account of punishment, see A.C. Ewing, THE MORALITY OF PUNISHMENT 53-54, 59 (1929); Hart, supra note 11, at 18-20, 40-41; MATT MATRAVERS, JUSTICE AND PUNISHMENT: THE RATIONALE OF COERCION 13-16 (2000); PRIMORATZ, supra note 1, at 15-31; C.L. Ten, supra note 35, at 87-89, 143-46.

For Bentham, conduct is to be judged according to the principle of utility, JEREMY BENTHAM, THE PRINCIPLES OF MORALS AND LEGISLATION 2 (Prometheus Books 1988) (1781) ("By the principle of utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party...") (footnote omitted) [hereinafter BENTHAM, THE PRINCIPLES], the maximization of pleasure or happiness and the minimization of pain, id. at 1 ("Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do.").

Bentham arrives at the principle of utility by first rejecting three competing moral principles which he conceives to be opposed to the principle of utility. First, the "principle of asceticism," which is like the principle of utility, "but in an inverse manner: approving of actions in as far as they tend to diminish his happiness; disapproving of them in as far as they tend to augment it." BENTHAM, THE PRINCIPLES, supra, at 9. Second, the "theological principle... professes to recur for the standard of right and wrong to the will of God." Id. at 21. Third, and of principal interest to us, is the "principle of sympathy and apathy," which praises or condemns actions on the basis of intuition and feelings. Id. at 16. The principle assesses actions not based on utility "but merely because a man finds himself disposed to approve or disapprove of them: holding up that approbation or disapprobation as a sufficient reason for itself, and disclaiming the necessity of looking out for any extrinsic ground." Id. This third principle Bentham locates as the source of retributivism. See PRIMORATZ, supra note 1, at 18. The principle is, for Bentham, "the real source of the retributive theory of punishment." Id. It "leads us to speak of offenses as deserving punishment." Id. (quoting JEREMY BENTHAM, THE THEORY OF LEGISLATION 76 (Richard Hildreth
versions of consequentialism judge acts based on their promotion of, for example, autonomy, achievement, or fairness. The principal consequentialist theories of punishment justify punishment based on the good consequences of rehabilitating the offender so that she will not commit future crimes, incapacitating the offender so that he cannot commit crimes during the term of imprisonment, deterring the offender from committing future crimes (specific deterrence), and deterring others in society from committing future crimes (general deterrence).

The most influential consequentialist theory justifies punishment based on the general deterrence of future crime. To achieve general deterrence, the appearance or publicity of punishment is crucial. Actual punishment, without society's awareness, generates no general deterrent effect; but apparent punishment, even if without actual punishment, does provide general deterrence. Actual punishment serves only to produce apparent punishment. As a result, retributivists have criticized deterrence-based theories

---

58 Gaut, supra note 56, at 177.
59 Influence of an offender, by punishment, is exerted on his “will, in which case it is said to operate in the way of reformation . . . .” BENTHAM, THE PRINCIPLES, supra note 57, at 170 n.1.
60 Influence of an offender, by punishment, is exerted on his “physical power, in which case it is said to operate by disablement . . . .” Id. at 170–71 n.1.
61 Regarding others in the community, punishment influences “their wills; in which case it is said to operate in the way of example.” Id.
62 Id. (punishment’s general deterrent effect on others in the community, or example, as he terms it, “is the most important end of all”); MURPHY & COLEMAN, supra note 54, at 118 (“While incapacitation and reform may sometimes figure in the [consequentialist] justification of punishment, deterrence has always been the mainstay . . . .”); PRIMORATZ, supra note 1, at 10 (“The most important consequences of punishment are its preventive effects . . . .”).
63 See BENTHAM, THE PRINCIPLES, supra note 57, at 192 (“Punishment cannot act any farther than in as far as the idea of it, and of its connection with the offence, is present in the mind. The idea of it, if not present, cannot act at all; and then the punishment itself must be inefficacious.”). Bentham has also argued:

if delinquents were constantly punished for their offences, and nobody else knew of it, it is evident that . . . there would be a great deal of mischief done, and not the least particle of good. . . . The punishment would befall [sic] every offender as an unforeseen evil. It would never have been present in his mind to deter him from the commission of crime. It would serve as an example to no one.


64 Bentham argues in support that it is the idea only of the punishment (or, in other words, the apparent punishment) that really acts upon the mind; the punishment itself (the real punishment) acts not any farther than as giving rise to that idea. It is the apparent punishment, therefore, that does all the service, I mean in the way of example, which is the principal object.

BENTHAM, supra note 57, at 193 (footnote omitted).
65 BENTHAM, supra note 63, at 398 (“Ought any real punishment to be inflicted? Most certainly. Why? For the sake of producing the appearance of it.”). And the greater the appearance or perception of punishment, the greater the deterrent effect. As Bentham explains: “[T]he real punishment ought to be as small, and the apparent punishment as great as possible. If hanging a man in effigy would produce
for being unable to justify actual punishment.\textsuperscript{66}

All of these consequentialist justifications of punishment share the goal of crime prevention.\textsuperscript{67} Both crime and punishment are evils,\textsuperscript{68} but punishment is only a qualified evil.\textsuperscript{69} The evil of punishment may be outweighed by the good consequences that it generates.\textsuperscript{70} That is, punishment is a nec-

the same salutary impression of terror upon the minds of the people, it would be folly or cruelty ever to hang a man \textit{in person}.”\textsuperscript{66}

According to this criticism, what does the real work of general deterrence is not actual punishment but the perception or publicizing of punishment. Primoratz, \textit{supra} note 1, at 42–43; J.D. Mabbot, \textit{Punishment, in PHILOSOPHY OF PUNISHMENT} 23, 23–24 (Robert M. Baird & Stuart E. Rosenbaum eds., 1988). If an offender commits a crime and is punished, of which the general public is unaware, the punishment fails to generate a general deterrent effect. In order to be deterred by punishment, awareness of the punishment is required. Let us take the opposite situation in which an offender commits a crime and is not actually punished, but the general public is misled into believing that the offender was punished. Though the offender was not actually punished, publicity of the faked punishment will have a deterrent effect. Since actual punishment that is not publicized has no deterrent effect and faked punishment that is perceived to be actual punishment does have a deterrent effect, it is the perception of punishment, and not actual punishment, which generates general deterrence. Retributivists conclude, therefore, that the general deterrence theory of punishment fails to justify actual punishment but only the faking of punishment.

Consequentialists respond that the general public would inevitably find out about the faked punishments and the deterrent effect would ultimately be undermined. Again, retributivists respond, in turn, that consequentialists would always be tempted to fake a punishment just once figuring that the general public would never find out. And regardless of the resolution of this empirical issue, the general deterrence theory cannot explain why faking punishment would be wrong. But consequentialists might respond that if the general deterrence effect could be generated without the costs and harms of punishing (without the public finding out and the effect ultimately undermined), then faking punishment would be justified.

\textsuperscript{67} For Bentham, “[t]he immediate principal end of punishment is to control action.” Bentham, \textit{The Principles}, \textit{supra} note 57, at 170 n.1. “General prevention ought to be the chief end of punishment, as it is its real justification.” Bentham, \textit{supra} note 63, at 396.

A further, “collateral” end of punishment, according to Bentham, is the “pleasure or satisfaction” punishment of the offender brings to the victim of the crime. Bentham, \textit{The Principles} \textit{supra} note 57, at 170–71 n.1. Though this consequence is beneficial, it is not in and of itself a sufficient purpose to impose punishment because “no such pleasure is ever produced by punishment as can be equivalent to the pain [of punishment experienced by the offender].” Id.

\textsuperscript{68} Id. at 170 (“All punishment is mischief: all punishment in itself is evil.”). Hegel rejects consequentialist theories precisely because of their conception of punishment as an evil which leads consequentialist theories to stray from the dictates of justice:

\textquote{This superficial character of [punishment as] an \textit{evil} is the primary assumption in the various theories of punishment as prevention, as a deterrent, a threat, a corrective, etc.; and conversely, what is supposed to result from it is just as superficially defined \ldots as a \textit{good}. But it is neither a question merely of an evil nor of this or that good; on the contrary, it is definitely \ldots a matter of \textit{wrong} and \textit{of justice}. As a result of these superficial points of view, however, the objective considerations of \textit{justice}, which is the primary and substantial point of view in relation to crime, is set aside \ldots . . . . }

Hegel, \textit{supra} note 22, at 124–125.

\textsuperscript{69} Bentham, \textit{The Principles}, \textit{supra} note 57, at 170 (footnote omitted) (punishment “ought only to be admitted in as far as it promises to exclude some greater evil”).

\textsuperscript{70} It might also be expressed that, for consequentialist theories, the good consequences of punishment are the \textit{only} goods sufficient to outweigh the evil or costs of punishment. \textit{See} Duff, \textit{supra} note 6, at 5.
Deterring Retributivism

necessary evil that may be justified by its diminution of the incidence of crime.71

Retributivism most broadly may be understood as a justification of punishment that does not rely on the good consequences stemming from punishment.72 Essentially, retributivism justifies punishment based not on its consequences but solely because an offender deserves it.73 Immanuel Kant74 declared that punishment “must always be inflicted upon . . . [a criminal] only because he has committed a crime.”75 Though this would seem to preclude both the rehabilitative and deterrent effects of punishment, Kant merely relegates these and other consequential goods to secondary status: the offender “must previously have been found punishable [deserving of punishment] before any thought can be given to drawing from his punishment something of use for himself or his fellow citizens.”76 Under

71 In Bentham's view,
when we consider that an unpunished crime leaves the path of crime open . . . to all those who may have the same motives and opportunities for entering upon it, we perceive that punishment inflicted on the individual becomes a source of security to all. That punishment which, considered in itself, appeared base and repugnant . . . is elevated to the first rank of benefits, when it is regarded not as an act of wrath or vengeance . . . but as an indispensable sacrifice to the common safety.

Bentham, supra note 63, at 396.

72 See, e.g., Hugo Bedau, Concessions to Retribution in Punishment, in JUSTICE AND PUNISHMENT 51, 53 (J.B. Cederbloom & William L. Blizek eds., 1977) (the emphasis of retributivism “is on the offense and nothing else, especially not any social cost/benefit or individual eugenics that can be calculated to result from punishments”); Gertrude Ezorsky, The Ethics of Punishment, in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at xi, xviii (“For all retributivists punishment has moral worth independently of any further desirable effects.”); id. at xi (noting that for retributivism, if punishment is justified, it is “irrespective of any further consequence” of punishment).

73 MICHAEL MOORE, PLACING BLAME: A THEORY OF CRIMINAL LAW 88 (1997) (“The distinctive aspect of retributivism is that the moral desert of an offender is a sufficient reason to punish him or her . . . .”); Bedau, supra note 72, at 52 (“[A] retributivist holds that a punishment is just if and only if the offender deserves it.”); Kent Greenawalt, Punishment, 74 J. CRM. L. & CRIMINOLOGY 343, 347 (1981) (retributivism is the view that “punishment is justified because people deserve it”); see Bedau, supra note 72, at 52 (“Probably the most widely held assumption about retribution in punishment is the idea that it makes desert the central feature of just punishment.”).


75 KANT, supra note 22, at 140.

76 Id. at 141. The following passage also suggests that there is not only room for concerns of deterrence but that it might be of primary interest:

The only question is whether it is matter of indifference to the legislator what kinds of punishment are adopted, as long as they are effective measures for eradicating crime (which violates the security a state gives each in his possession of what is his), or whether the legislator must also take into account respect for the humanity in the person of the wrongdoer (i.e., respect for the species) simply on grounds of Right.

Id. at 168.

Kant seems to be suggesting that punishment must first be efficacious in deterring crime and only
retributivism, morally culpable wrongdoing or guilt\textsuperscript{77} deserves, merits, or warrants punishment.\textsuperscript{78} It is morally fitting that an offender should suffer in proportion\textsuperscript{79} to her desert\textsuperscript{80} or culpable wrongdoing.\textsuperscript{81}

then does the question of also punishing justly come into play. Though in the following passage Kant clearly answers the question in the affirmative, it is unclear whether crime prevention is still of primary importance, or even of any importance: “I said that the \textit{ius talionis} is by its form always the principle for the right to punish since it alone is the principle determining this Idea a priori (not derived from experience of which measures would be most effective for eradicating crime).” \textit{Id.} (footnote omitted).

\textsuperscript{77} Quinton, \textit{supra} note 41, at 7 (“The essential contention of retributivism is that punishment is only justified by guilt.”).

\textsuperscript{78} In contrast, for consequentialism, morally culpable wrongdoing or the guilt of the offender does not necessarily warrant punishment. Punishment of an offender is impermissible if the cost of punishment outweighs its good consequences. According to Bentham, punishment should not be inflicted at all under any of the following four conditions, where punishment would be: (i) “groundless” because there is no crime or harm, (ii) “ineffective” because the crime cannot be deterred, (iii) “unprofitable, or too expensive” because the evil of the punishment would exceed the crime and, (iv) “needless” because the crime may be deterred by other means than punishment or does not require deterrence.


\textsuperscript{79} The proportionality or fit between the crime and the punishment is a central tenet of retributivism. For Kant, the type and degree of punishment is to be determined by the principle of equality between the crime and the punishment, \textit{KANT, supra} note 22, at 141, “what is done to [the offender] in accordance with penal law is what he has perpetrated on others,” \textit{id.} at 169. As Kant famously explains the principle of the \textit{lex talionis}:

\begin{quote}
[W]hatever undeserved evil you inflict upon another within the people, that you inflict upon yourself. If you insult him, you insult yourself; if you steal from him, you steal from yourself; if you strike him, you strike yourself; if you kill him, you kill yourself. But only the law of retribution (\textit{ius talionis}) \ldots can specify definitely the quality and the quantity of punishment; all other principles are fluctuating and unsuited for a sentence of pure and strict justice because extraneous considerations are mixed into them.
\end{quote}

\textit{Id.} at 141.

Georg W.F. Hegel adopts a more sophisticated version of Kant’s \textit{lex talionis}. Hegel acknowledges that the simple version of the \textit{lex talionis} is embodied in “the universal feeling of peoples and individuals towards crime [which] is \ldots that \textit{what the criminal has done should also happen to him}.” \textit{HEGEL, supra} note 22, at 127. But this simple version, Hegel recognizes, may be reduced to the absurd:

\begin{quote}
[It] is easy to portray the retributive aspect of punishment as an absurdity (theft as retribution for theft, robbery for robbery, an eye for an eye, and a tooth for a tooth, so that one can even imagine the miscreant as one-eyed or toothless); but the concept has nothing to do with this absurdity, for which the introduction of that [idea of] \textit{specific equality} is alone to blame.
\end{quote}

\textit{Id.} at 128 (footnote omitted). Rather than the specific equality in the simple version of the \textit{lex talionis}, the necessary identity or connection between crime and its retribution lies in its character and value. Whereas the simple version of the \textit{lex talionis} requires a specific equality between the crime and the punishment (e.g., theft as retribution for theft), Hegel’s more sophisticated version requires merely that the crime and punishment be comparable in character or value. Hegel defines value as “the \textit{inner equality} of things which, in their existence \ldots are specifically quite different \ldots .” \textit{Id.} at 128. That is, the two must be generally equal, or comparable:

\begin{quote}
Equality remains merely the basic measure of the crime’s \textit{essential} deserts, but not of the specific external shape which the retribution should take. It is only in terms of this specific shape that theft and robbery [on the one hand] and fines and imprisonment etc. [on the other] are completely unequal, whereas in terms of their value, i.e. their universal character as injuries \ldots they are comparable.
\end{quote}

\textit{Id.} at 129. In this way, the absurdities of Kant’s simple \textit{lex talionis} are avoided. For example, a robber who has nothing which can be robbed in retribution may still be properly punished by a punishment, e.g.
Retributivism’s justification of “it’s right to punish criminals because doing so is right” is both its strength and its weakness. Consequentialists argue that the justification, as such, is more intuition than justification, is circular or empty, and constitutes a denial of the need to supply a justification. Moreover, even if the retributive intuition that deserving wrongdoers deserve to suffer is accepted, retributivism cannot explain why the state has the right or the duty to give them the suffering that they deserve. Though imprisonment, so long as the robbery and the term of imprisonment are comparable in character and value. Hegel, however, fails to explain precisely how we are to determine whether a crime and its corresponding punishment, even if not of specific equality, are nonetheless of the requisite comparability in terms of value and character. Hegel does warn us, however, of the consequences of the failure to see the connection:

If we do not grasp either the connection, as it is in itself, between crime and its nullification, or the thought of value and the comparability of crime and punishment in term of value, we may reach the point of regarding a proper punishment as a purely arbitrary association of an evil . . . with an illicit action.

Id. at 129 (citation omitted).

For the view that the term desert is susceptible to different meanings, see Jeffrie G. Murphy, Legal Moralism and Liberalism, 37 Ariz. L. Rev. 73, 91 (1995) (identifying five different types of desert).

Benn, supra note 45, at 30 (retributivism maintains that “the punishment of crime is right in itself, that it is fitting that the guilty should suffer, and that justice, or the moral order, requires the institution of punishment”).

Wasserstrom notes that the retributivist justification for punishment—“the guilty ought to be punished because they deserve it”—is “not an argument but merely an assertion.” Wasserstrom, supra note 34, at 337.

See, e.g., Hugo Bedau, Retribution and the Theory of Punishment, 75 J. Phil. 601, 603 (1978); Dolinko, supra note 12, at 517. David Dolinko has argued:

The retributivist might claim that what is “fitting” and “proper” and “deserved” is not, after all, that the criminal suffer or receive punishment, but precisely the act of inflicting the punishment on the criminal. This, however, would seem to make the retributivist “theory” of why punishment is morally justified nothing more than a conclusory “Because it is!” Foes of retributivism have indeed faulted it as just such an unenlightening ipse dixit . . . .

Dolinko, supra note 2, at 1629–30 n.25.

Benn, supra note 45, at 30:

The most thoroughgoing retributivists, exemplified by Kant, maintain that the punishment of crime is right in itself, that it is fitting that the guilty should suffer, and that justice, or the moral order, requires the institution of punishment. This, however, is not to justify punishment but, rather, to deny that it needs any justification. To say that something is right or good in itself means that it does not need to be justified in terms of the value or rightness of anything else. Its intrinsic value is appreciated immediately or intuitively. But since at least some people do doubt that punishment is right, an appeal to intuition is necessarily unsatisfactory. Again, to say “it is fitting” or “justice demands” that the guilty should suffer is only to reaffirm that punishment is right, not to give grounds for thinking so.

Id.

Bagorici & Amarasekara, supra note 12, at 164–65 (footnote omitted) (“the claim that punishment should be administered by the state does not follow from intrinsic retributivism: even if it shows that the guilty deserve to suffer, it cannot support the claim that the suffering should be deliberately inflicted on wrongdoers by the state”); Benn, supra note 45, at 32; Dolinko, supra note 12, at 518–22; Douglas Husak, Why Punish the Deserving?, 26 Nous 447 (1992) (even if we accept the retributivist premise that deserving wrongdoers deserve to be punished, what justifies the state in giving such persons what they deserve without some cost/benefit analysis that makes the state doling out desert, in the form of punishment, worthwhile?); C.L. Ten, supra note 35, at 48 (same); C.L. Ten, supra note 29, at 368.
perhaps circular or empty, the retributivist justification would seem to preclude punishment of the innocent. This is because only those who deserve punishment may be punished and innocents do not deserve punishment. In contrast, though consequentialism provides a clear justification—punishment’s good consequences. Such a justification, however, is thought to make the theory vulnerable to the charge of intentionally punishing the innocent. This is because the good consequences—deterrence, for example—that justify punishment might be attained regardless of the guilt or innocence of the offender. Thus, consequentialism’s strength in furnishing a clear justification for punishment is also its weakness in seemingly allowing punishment of the innocent. For retributivism, it is just the opposite. Whereas consequentialism looks forward to the good results to be attained by punishing, retributivism looks backward to the wrongdoing already committed. The retributivist justification for punishment lies in “its relation to a past offence;” retributivism “is the notion that there is a mystic bond between wrong and punishment . . . .” Georg W.F. Hegel, who along with Kant is the leading influence on retributivism, arguably supplies the richest explication of that bond or relationship. Hegel seeks to forge a necessary connection between crime and punishment, and retribution is that connection. Central to Hegel’s account is that punishment

---

86 Morris, supra note 6, at 38; Bagoric & Amanasekara, supra note 12, at 129–131; Cragg, supra note 52, at 707; Fletcher, supra note 14, at 516; Oldenquist, supra note 17, at 749. 87 Duff, supra note 1, at 4 (finding the quoted explanation to be the common denominator of all retributivist accounts). For a similar view, see Bernard Bosanquet, Some Suggestions in Ethics 188 (1919) (“Punishment is prima facie retrospective; it deals with the past.”). 88 Oliver W. Holmes, Jr., The Common Law 42, (reprint ed. 1991) (1881). 89 For general discussions of Hegel’s theory of punishment see Ewing, supra note 57, at 21–26; Ted Honoreich, Punishment: The Supposed Justifications 45–48 (1969); Alan W. Norrie, Law, Ideology and Punishment: Retrieval and Critique of the Liberal Ideal of Criminal Justice 65–88 (1991); Primoratz, supra note 1, at 67–82. See generally Mark Tunick, Hegel’s Political Philosophy: Interpreting the Practice of Legal Punishment (1992); Markus Dubber, Rediscovering Hegel’s Theory of Crime and Punishment, 92 Mich. L. Rev. 1577 (1994) (book review); Peter J. Steinberger, Hegel on Crime and Punishment, 77 Am. Pol. Sci. Rev. 858 (1983). 90 See Primoratz, supra note 1, at 13 (“The most important and influential among classical retributivists are Kant and Hegel.”). 91 For the view that Hegel has developed a “richer” and more “systematic” account of retributivism than Kant, see id. See also Dubber, supra note 89, at 1581 (in contrast to Kant’s scant writings on punishment, “Hegel developed the deontological foundations of Kant’s theory into an all-encompassing theory of the logical connection between crime and punishment”). But see Honoreich, supra note 89, at 45 (dismissing Hegel’s account as “obscure” and “of very secondary interest”). 92 Hegel, supra note 22, at 129 (“punishment is merely a manifestation of the crime, i.e. it is one half which is necessarily presupposed by the other”); id. (committing the criminal “deed brings its own retribution with it”); Dubber, supra note 89, at 1608 (footnote omitted) (interpreting Hegel’s theory as “[t]he offender, by committing the criminal act, herself posits the law that compels her punishment”). 93 Hegel, supra note 22, at 129 (“Retribution is the inner connection and the identity of two determinations [the value and character of an offender’s crime and punishment] which are different in appearance and also have a different external existence . . . in relation to one another.”).
cancels, negates, or annuls crime. Crime, the infringement of the "Right," is conceived of as the initial use of force or coercion. In the abstract, force or coercion (and thus crime) is "contrary to right." Punishment is also a form of force or coercion, and thus (in the abstract) also wrong. But because "coercion is cancelled... by coercion; it [punishment] is therefore not only conditionally right but necessary—namely as a second coercion which cancels an initial coercion." Since crime is an infringement or negation of the right, and punishment cancels the crime, "punishment is merely the negation of the negation." In other words, it is not so much that two wrongs make a right but that the two wrongs (in the abstract) restore the right. But this retrospective outlook has led consequentialists to claim that retributivism is merely a philosophical rationalization for revenge, retaliation, and vengeance.

---

94 See PRIMORATZ, supra note 1, at 74. Punishment as the annulment of an offense is "one of the basic tenets of [Hegel's] theory of punishment." Id.
92 See HEGEL, supra note 22, at 122 ("Right, whose infringement is crime... ").
96 Id. at 121 ("The initial use of coercion, as force employed by a free agent in such a way as to infringe the existence [Dasein] of freedom in its concrete sense—i.e. to infringe right as right—is crime.").
97 Id. at 120.
98 PRIMORATZ, supra note 1, at 69 (interpreting Hegel's view to be that "retribution is a kind of coercion").
99 HEGEL, supra note 22, at 120 (§ 93).
100 Id.
101 Id. at 123.
102 Id. at 124 (referring to punishment of crime as "the restoration of right"); see id. at 252 (the law "restores and thereby actualizes itself as valid through the cancellation... of the crime").

Hegel's doctrine that punishment annuls the crime done by the offender has been subject to some ridicule. See, e.g., John L. Mackie, Morality and the Retributive Emotions, 1982 CRIM. JUST. ETHICS 3, 5 ("But [Hegel's annulment doctrine] really is incoherent. The punishment may trample on the criminal, but it does not do away with the crime."). Though predating Hegel, Cesare Beccaria also wonders how punishment can undo that which has been done. CESARE BECCARIA, ON CRIMES AND PUNISHMENTS 23 (Ch. XII) (David Young trans., 1986) (1764) ("The purpose of punishments is not to torment and afflict a sentient being or to undo a crime which has already been committed. ... Can the cries of a poor wretch turn back time and undo actions which have already been done?").

Others have wondered exactly how the second wrong (the suffering of the person punished) following the first wrong (the commission of a criminal offense) makes a right. See, e.g., HART, supra note 11, at 234–35 ("To some critics it appears to be a mysterious piece of moral alchemy in which the combination of the two evils of moral wickedness and suffering are transmuted into good ... "); Jean Hampton, The Moral Education Theory of Punishment, 13 PHIL. & PUB. AFF. 208, 236 ("How is the society that inflicts the second evil any different from the wrongdoer that has inflicted the first?").

103 The famous remark which perhaps has cemented the perception that retributivism is a rationale for vengeance, but which could equally be used to distinguish retributivism from vengeance, is as follows: "The criminal law stands to the passion of revenge in much the same relation as marriage to the sexual appetite." J.J. STEPHEN, GENERAL VIEW OF THE CRIMINAL LAW OF ENGLAND 99 (1st ed. 1863).

104 See PLATO, supra note 19, at 321 ("In punishing wrongdoers, no one concentrates on the fact that a man has done wrong in the past, or punishes him on that account, unless taking blind vengeance like a beast."); PRIMORATZ, supra note 1, at 83 ("One of the more popular objections to the retributive theory is the claim that it is in fact a philosophical rationalization of vengefulness."); Cragg, supra note 52, at 707 ("The most common and pressing concern with retributivism is its association in the minds of
For retributivism, punishment is not a means to other goods or consequences but is an end in itself. As such, punishment is an intrinsic good or has intrinsic value because it is valued for its own sake. "The state of affairs where a wrongdoer suffers punishment is morally better than the state of affairs where he does not; and it is better irrespective of any of the consequences of punishing him." In contrast, for a consequentialist theory, punishment is a qualified evil but is instrumentally good or of instrumental value because its value lies in what good consequences punishment might lead to. These contrasting views on punishment generate each theory's criticism of the other. Because for retributivism punishment is an end in itself, consequentialists ask why bother inflicting suffering "if no further good is achieved by doing so." But by consequentialism achieving a further good beyond punishment itself, retributivists claim that consequentialism immorally uses those punished as mere means or objects to attain that further good. In contrast, retributivism claims to respect and value the dignity and personhood of offenders by treating those punished as ends in themselves. Retributivism, however, honors offenders to such a degree that it conceives of punishment as the right of the offender. Recognizing many with the idea of vengeance.") Quinton, supra note 41, at 6 ("Retributivists, who seem to hold that there are circumstances in which the infliction of suffering is a good thing in itself, are charged by their opponents with vindictive barbarousness.").

For arguments attempting to distinguish retributivism from vengeance, revenge and retaliation, see, for example, Hegel, supra note 22, at 129, who notes: "What is at first sight objectionable about retribution is that it looks like something immoral, like revenge, and may thus be interpreted as a personal matter. Yet it is not the personal element but the concept itself which carries out retribution." Id.; see also Moore, supra note 73, at 89; Robert Nozick, Philosophical Explanations 366-68 (1981); Primoratz, supra note 1, at 83-85; Oldenquist, supra note 17, at 749.

105 Duff, supra note 6, at 6-7.
106 See infra notes 609-612 and accompanying text.
107 Rawls, supra note 1, at 5.
108 See infra notes 609-612 and accompanying text.
109 Duff, supra note 6, at 5-6.
110 Moore, supra note 73, at 93 n.19 (acknowledging the criticism as "[t]he main problem with the pure retributivist theory of punishment"); accord Cragg, supra note 52, at 708 (describing it as "[p]erhaps the most telling criticism" of retributivism).
111 See infra notes 397-399 and accompanying text.
112 See infra note 400 and accompanying text.
113 Hegel is perhaps the first retributivist to declare that the offender has a right to be punished. Hegel, supra note 22, at 126 (§ 100) ("The injury which is inflicted on the criminal is not only just in itself . . . it is also a right for the criminal himself . . . ."). Hegel explains that a criminal's right to be punished is located in the fact that a rational being has committed the act constituting the criminal offense:
For it is implicit in his action, as that of a rational being, that it is universal in character, and that, by performing it, he has set up a law which he has recognized for himself in his action, and under which he may therefore be subsumed as under his right.
Id.

For others holding the view that an offender has a right to be punished, see, for example, Morris,
the view's seeming absurdity, consequentialists have commented, "It is an odd sort of right whose holders would strenuously resist its recognition."

Who or what is the imposition of punishment supposed to serve? Retributivists and consequentialists supply differing answers. For retributivists, punishment serves abstract justice or right. Punishment is simply doing justice. If, for retributivism, punishment of deserving wrongdoers is a morally better state of affairs, then, for consequentialism, punishment generates a better state of affairs, regardless of one's conception of the good. (Typically, however, a consequentialist's conception of a better state of affairs will be less an abstract, but rather, a materially better state of affairs—societal welfare, and in particular, the diminution of crime). These are the responses to the "what" question. In response to the "who" question—who is the imposition of punishment supposed to serve—it might be argued that each theory serves different constituencies. If retributivism may be said to do justice to the criminal offender (by only imposing fair and deserved punishment), consequentialism might be said to do justice to future innocent crime victims (by seeking to deter or prevent future crime). Retributivists complain that consequentialism does justice to the latter by exploiting the former; consequentialists criticize retributivism for doing justice to the former by ignoring the latter. Which theory may lay claim to doing justice to past crime victims has recently become a contentious debate and one we will visit in Part V.

B. Varieties of Retributivism

There are a considerable number of versions of retributivism, and in order to see why retributivists living in glass houses should not throw stones at consequentialism, it may be helpful to have some rudimentary understanding of the variety of glass houses that retributivists occupy. Forms of retributivism differ as to how much a wrongdoer should be punished. Weak, or negative, retributivism merely requires that a wrongdoer must not be punished more than she deserves; strong, or positive, retributivism

\[\text{supra note 6, at 41 ("a person [i.e., a wrongdoer or offender] has a right to be punished, meaning by this that a person had a right to all those institutions and practices linked to punishment"); JEFFRIE MURPHY, RETRIBUTION, JUSTICE, AND THERAPY 134 (1979) ("The right to be punished and regarded as a responsible agent, though sometimes painful when honored, at least leaves one's status as a moral person intact") (emphasis added).}\]

\[114\] For discussion of this criticism and the retributivist responses, see PRIMORATZ, supra note 1, at 99–108.

\[115\] Quinton, supra note 41, at 8–9.

\[116\] John Cottingham, Varieties of Retribution, 29 PHIL. Q. 238 (1979) (finding nine distinct versions of retributivism); C.L. Ten, supra note 29, at n.366.

\[117\] The term "weak retributivism" may stem from HART, supra note 11, at 233. The term "negative retributivism" may derive from Mackie, supra note 102, at 4.

\[118\] Weak, or negative, retributivism would be satisfied, therefore, even if the wrongdoer received no
requires that a wrongdoer be punished to the fullest extent of his just de-
serts. Retributivist theories also vary as to whether justified punishment is 
obligatory or merely permissible. Versions of retributivism may differ as to what punishment is imposed for. Moralistic retributivism justifies the 
punishment of an actor's morally culpable wrongdoing even if it does not amount to a violation of law. Legalistic retributivism justifies the pun-
ishment of an actor's violation of law even if it does not amount to a moral 
wrong. And the combined version joins the moralistic and legalistic ver-
punishment at all. Because weak or negative retributivism fails to provide an affirmative reason to pun-
ish a wrongdoer at all, it is infrequently employed as a complete justification for punishment. David 
Dolinko, Some Thoughts About Retributivism, 101 ETHICS 537, 539–544 (1991); Duff, supra note 6, at 7 
(negative retributivism "clearly provides no complete justification for it tells us that we may punish the 
guilty (their punishment is not unjust), but not that or why we should punish them" (citation omitted)). 
Because a wrongdoer not only deserves to be punished, but also has a right to punishment, re-
tributivism generally conceives the punishment of wrongdoers to be obligatory. Kaplow & Shavell, su-
pra note 28, at 1229 n.660 (referring to obligatory retributivism as the "standard retributive view"); 
PRIMORATZ, supra note 1, at 110 (the “dominant” view in retributivism is that the duty to punish, 
though obligatory, is only of “paramount, but not absolute importance”). 
For obligatory retributivism, even if the consequences of punishment would be disastrous, the 
demands of justice necessitate punishment of culpable wrongdoing irrespective of the consequences. C.L. 
Ten, supra note 35, at 75 (criticizing the obligatory form of retributivism as “implausible” because punish-
ment would be obligatory even if “the skies will fall”). In contrast, under a consequentialist theory 
wrongdoers need not (and perhaps must not) be punished unless the good consequences generated by the 
punishment outweigh the various costs and harms of punishment. 
For examples of obligatory retributivism, see HEGEL, supra note 22, at 251 (§ 218) (“It would be 
impedible for society to leave a crime unpunished.”); id. at 127 (§ 100) (“Both the nature of crime and 
the criminal’s own will require that the infringement for which he is responsible should be cancelled 
punished.”) (emphasis added)); KANT, supra note 22, at 141: 
The principle of punishment is a categorical imperative, and woe to him who crawls through the 
windings of eudaemonism in order to discover something that releases the criminal from punish-
ment or even reduces its amount by the advantages it promises, in accordance with the Pharisaical 
saying, “It is better for one man to die than for an entire people to perish.” For if justice goes there 
is no longer any value in men’s living on the earth. 
Id.; Michael Moore, The Moral Worth of Retribution, in RESPONSIBILITY, CHARACTER, AND THE 
EMOTIONS, 179, 182 (Ferdinand Schoeman ed., 1987) (“For a retributivist, the moral culpability 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.”). 
While obligatory retributivism requires that culpable wrongdoing be punished, permissive re-
tributivism does not mandate punishment but merely justifies its permissibility. See, e.g., K.G. Arm-
strong, The Retributivist Hits Back, in THE PHILOSOPHY OF PUNISHMENT: A COLLECTION OF PAPERS, 
supra note 11, at 138, 155–57. For general discussions of permissive retributivism, see generally 
RICHARD B. BRANDT, ETHICAL THEORY: THE PROBLEMS OF NORMATIVE AND CRITICAL ETHICS 501– 
03 (1959); H.J. McCloskey, Utilitarian and Retributive Punishment, 64 J. PHIL. 91 (1967). For an ex-
ample of permissive retributivism, see H.J. McCloskey, A Non-Utilitarian Approach to Punishment, 8 
INQUIRY 249 (1965), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 
119, 131–34 [hereinafter McCloskey, Non-Utilitarian]. 
A critique of theories of punishment which couch their justification in moral terms, see Joel 
Feinberg, On Justifying Legal Punishment, in NOMOS III: RESPONSIBILITY 152, 156–57 (Carl J. Frie-
The terms “moralistic retributivism” and “legalistic retributivism” come from Feinberg, supra 
note 55, at 646–47.
sions: both morally culpable wrongdoing and the violation of a criminal law are independently necessary and jointly sufficient to justify punishment.\textsuperscript{123}

Though observers generally contrast retributivism with consequentialism on the basis that only the latter justifies punishment upon the good consequences attained by punishment,\textsuperscript{124} a class of retributive theories (termed conceptually consequential retributivism\textsuperscript{125}) justify punishment by resort to what is claimed to be a special type of consequences—conceptual or logical consequences.\textsuperscript{126} Conceptually consequential retributivism may be distinguished from pure or simple retributivism, which justifies punishment not on the basis of consequences but solely because the offender deserves it.\textsuperscript{127}

C. The Mixed Theory of Punishment

In response to the seemingly intractable debate between retributivist and consequentialist theories of punishment, there have been attempts to combine the two. These “mixed” theories are the intellectual heirs to Cesare Beccaria’s effort to fuse utility and justice.\textsuperscript{128} Among these mixed

\textsuperscript{123} See infra notes 236–238 and accompanying text.

\textsuperscript{124} Duff, supra note 6, at 5–6 (“The common feature of all “consequentialist” accounts . . . is that they justify punishment in terms of its contingent or instrumental contribution to an independently identifiable good.”); id. at 6 (“A nonconsequentialist insists that actions and practices may be right or wrong in virtue of their intrinsic character, independently of their consequences.”).

\textsuperscript{125} I take the term from Fletcher, supra note 14, at 516.

\textsuperscript{126} Conceptually consequential retributivism may be distinguished from consequentialism. Consequentialist theories justify punishment based on the attainment of actual or contingent consequences that may or may not occur. And whether or not they occur is subject to empirical verification. For example, the consequences of the offender being rehabilitated, the offender not committing crime while incapacitated, the offender being deterred not to commit future crimes, and members of the general public being deterred from committing future crimes, may or may not occur and are susceptible to empirical assessment. Conceptual or logical consequences, on the other hand, are those which are abstract or are claimed to necessarily follow as a result of punishment. Examples of the conceptual or logical consequences used to justify punishment by theories commonly perceived to be retributive include negation of the crime, avoidance of society’s complicity with the crime, vindication of the victim of the crime and, reallocation of the wrongdoer’s balance of society’s benefits and burdens. See infra notes 579–593 and accompanying text. All of these consequences are said to be abstract and necessarily follow from punishment.

\textsuperscript{127} Justifying punishment without resort to any consequences of any type is “[t]he most thoroughgoing” form of retributivism, Benn, supra note 45, at 30, and, “represents the paradigm of retributive thinking,” Fletcher, supra note 14, at 516.

\textsuperscript{128} Perhaps the first comprehensive mixed theory was Cesare Beccaria’s 1764 landmark On Crimes and Punishments. By combining elements of justice and utility (or retributivism and consequentialism), Beccaria sought to overturn the cruelty, torture and barbarism endemic to punishment practices of his age. For Beccaria, the best way to minimize the evil and cruelty of punishment was to reduce its occurrence. Taking the premise that “[i]t is better to prevent crimes than to punish them,” BECCARIA, supra note 102, at 74 (Ch. XLI), Beccaria concludes that “a punishment for a crime cannot be deemed truly just . . . unless the laws have adopted the best possible means . . . to prevent that crime,” id. at 60 (Ch. XXXI). Quoting Montesquieu, Beccaria asserted that “punishment which does not derive from absolute necessity . . . is tyrannical.” Id. at 8 (Ch. II).
theories of punishment, the most influential is H.L.A. Hart’s. Hart distinguished the competing justifications of punishment into the separate issues of what justified the general practice and institutions of punishment (the General Justifying Aim) as opposed to who should be punished (Distribution). For Hart, the General Justifying Aim of our institutional practice of punishment is the beneficial consequences that punishment generates. 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 deter-

Rather than separating out concerns of justice and crime prevention, as we tend to do today, for Beccaria the concerns were inextricably intertwined. Justice lays precisely in punishing no more than is necessary to preserve and promote public welfare. By justice Beccaria means nothing more than “the bond necessary to hold private interests together.” Id. at 9. Beccaria explains that “[a]ll punishments that exceed what is necessary to preserve this bond are unjust by their very nature.” Id. Contrary to our view today that the goals of retributivism and deterrence are incompatible, for Beccaria, the more just the punishment, the greater the crime prevention. As Beccaria puts it: “[T]he foundation of the sovereign’s right to punish crimes [is]: the necessity of defending the depository of the public welfare against the usurpations of private individuals. Further, the more just punishments are, the more sacred and inviolable is personal security . . . .” Id. at 8.

Though generally viewed as a utilitarian, Beccaria seems to be neither utilitarian nor retributivist. As distinct from a thoroughgoing utilitarian, Beccaria prohibits punishment even where its consequences are good or useful if the punishment is unjust. Id. at 43 (Ch. XXV) (“[E]ven if punishments produce a happy result, they are not therefore just, for, in order to be just, they must be necessary. A useful injustice cannot be tolerated . . . .”). Like retributivism’s lex talionis, Beccaria argues that “to tighten even further the connection between the misdeed and its punishment . . . the latter should conform as closely as possible to the nature of the crime.” Id. at 37 (Ch. XIX). For example, “[w]hoever seeks to enrich himself with the property of others ought to be deprived of his own.” Id. at 39 (Ch. XXII). And Beccaria shares the retributivist concern for respecting the offender’s rights and autonomy. See infra notes 404–406 and accompanying text. But Beccaria roundly rejects retributivism. Rather, the purpose of punishment is deterrence, both specific and general. BECCARIA, supra note 102, at 23 (Ch. XII) (“The purpose of punishment, then, is nothing other than to dissuade the criminal from doing fresh harm to his compatriots and to keep other people from doing the same.”).

If more consequentialist than retributivist, Beccaria might be considered a compassionate consequentialist, rather than a utility-calculating consequentialist as was Jeremy Bentham. H.L.A. HART, ESSAYS ON BENTHAM: STUDIES IN JURISPRUDENCE AND POLITICAL THEORY 51 (1982) (“there is in Beccaria a respect for the dignity and value of the individual person which is absent in Bentham”). For other notable efforts to provide a mixed theory of punishment, combining elements of consequentialism and retributivism see generally Ewing, supra note 57, at 300 (unlike retributivism, punishment justified by the good consequence of preventing crime by expressing moral condemnation but like retributivism, offenders should be punished only if guilty and only as much as their just deserts reflect); Quinton, supra note 41, at 12–15 (consequentialist concerns justify punishment while retributivism provides the definition of punishment); Rawls, supra note 1, at 5–6 (rule utilitarian approach in which the justification of punishment institutions is deterrence but the rules employed by the punishment institutions are retributive); C.L. Ten, supra note 35, at 79–81.

129 Id. at 8–13.
130 Id. at 11.
131 Id. at 8.
132 Id. at 11.
ence could be promoted.

The difficulty for most mixed theories is that either substantive issues of justification are assumed away by definition or the instability of the components leads to a collapse back into consequentialism. While Hart's mixed theory may avoid those problems, its own difficulty is in setting the amount of punishment. The negative retributivism component sets the upper limit, and deterrence concerns set the lower limit on how much punishment to impose. By ignoring positive retributivism, the mixed theory has been criticized by retributivists as being disproportionately lenient in (i) not justifying any punishment at all of guilty wrongdoers where deterrence could not be promoted, and (ii) not justifying the punishment of guilty wrongdoers to the fullest extent of their just deserts where less punishment satisfies the aim of deterrence. And by not fully embodying the aim of deterrence, the mixed theory is disproportionately lenient where promotion of deterrence requires a sentence exceeding in severity the limit set by negative retributivism.

The focus of this Article will be demonstrating that the four principal criticisms retributivists have leveled against consequentialist theories of punishment apply equally to retributivism. As a result, it will be concluded that retributivism's place on the high moral ground is misplaced. Retributivism is no better than consequentialist theories of punishment and retributivism is perhaps incoherent by failing to satisfy its own principles.

II. RETRIBUTIVISM'S INTENTIONAL PUNISHMENT OF THE INNOCENT

Two versions of retributivism—moralistic and legalistic—will be shown to justify the intentional punishment of particular, identifiable innocents. Moralistic retributivism justifies the intentional punishment of morally culpable wrongdoers who are legally innocent—by legally innocent, I mean those individuals who, despite their moral wrongdoing, could not be convicted of committing a criminal offence. Legalistic retributivism justifies the intentional punishment of the legally, or criminally, guilty who are morally innocent—by morally innocent I mean those individuals who, despite their violation of a criminal statute, have not committed moral wrongdoing. Before examining these two specific theories, let us first consider the state of the debate between retributivists and consequentialists over this

133 Primoratz, supra note 43, at 560.
134 PRIMORATZ, supra note 1, at 145.
135 MOORE, supra note 73, at 97–102.
137 For some additional criticisms of retributivism, see, for example, Cragg, supra note 52, at 708 (retributivism insufficiently values "compassion, forgiveness and mercy"); Kaplow & Shavell, supra note 28, at 1236 (retributivism fails to "include a definition of what constitutes wrongful behavior deserving punishment") (footnote omitted); Jeffrie G. Murphy, supra note 80, 78–79, 83–84 (retributivism is inconsistent with liberalism's harm principle and liberalism in general).
most crucial issue.

A. The Debate over Intentional Punishment of the Innocent

This subpart presents the retributivist charge that consequentialism justifies the intentional punishment of the innocent, and the various consequentialist responses to that charge.

1. The Retributivist Charge: Consequentialism Justifies the Intentional Punishment of the Innocent.—Retributivists' principal, and most devastating, criticism of consequentialist theories of punishment is that they justify punishment of the innocent.138 The force of this criticism has been so powerful that it alone is considered a sufficient basis to reject consequentialism.139 Because retributivism's justification for punishment is based on the desert of the punished, and an innocent presumably does not deserve to be punished, it would seem that retributivism cannot justify punishment of the innocent.140 Consequentialist theories are susceptible to the criticism precisely because they justify punishment by the good consequences to be attained by punishment.141 If one of the good consequences sufficient to justify punishment, for example, general deterrence, may be attained by

138 DUFF, supra note 1, at 154 ("A system of [consequentialist] punishment is most obviously defective if it involves the deliberate punishment of the innocent"); HART, supra note 11, at 5–6 (terming consequentialist theories justifying punishment of the innocent as the "stock 'retributive' argument"); MOORE, supra note 73, at 93 n.19 ("The main problem with the pure utilitarian theory of punishment is that it potentially sacrifices the innocent in order to achieve a collective good."); PRIMORATZ, supra note 1, at 113 (punishment of the innocent is the "main objection" to consequentialist theories of punishment); Benn, supra note 45, at 31 (justifying punishment of the innocent as the "dilemma" of consequentialist theories); Cragg, supra note 52, at 708 (consequentialist theories justifying punishment of the innocent is a "telling criticism"); Greenawalt, supra note 11, at 1341 ("the most damaging aspect of the [retributivist] attack is that utilitarianism admits the possibility of justified punishment of the innocent"); Mabbot, supra note 66, at 23 (terming punishment of the innocent as the "central difficulty" for consequentialist theories of punishment); Quinton, supra note 41, at 9 (Retributivists "crucial charge is that utilitarians permit the punishment of the innocent."); C.L. Ten, supra note 29, at 366–67 (punishment of the innocent is the "main problem" for consequentialist theories); see Murphy, supra note 6, at 4 ("Philosophers have written at great length about the moral problems involved in punishing the innocent—particularly as these problems have raised obstacles to an acceptance of the moral theory of Utilitarianism.").

139 Consider the following expression of the general view that justifying punishment of the innocent dooms consequentialism: "In short, if one maintains that punishment is justified by deterrence alone one seems committed to the immorality of punishing the innocent. It seems that one is committed to denying the rule that only the guilty may be punished. Surely, then, the deterrence theory is unacceptable." HONDERICH, supra note 89, at 62. For similar expressions of this prevailing view, see Bagoric & Amarasekara, supra note 12, at 133 (that consequentialism may justify punishment of the innocent "has been so persuasive that it alone has led many to reject utilitarianism as a general theory of morality"); Michael Philips, The Inevitability of Punishing the Innocent, 48 PHIL. STUD. 389, 389 (1985) (that consequentialism may justify punishment of the innocent "is often taken to be a sufficient refutation of Utilitarian theories of punishment and of utilitarianism in general").

140 Bagoric & Amarasekara, supra note 12, at 133.

141 For a good discussion of the range of various consequentialist responses to the criticism, see PRIMORATZ, supra note 1, at 43–65; C.L. TEN, supra note 35, at 13–37.
punishment of the innocent, then punishment of the innocent is justified under consequentialism.

Let us consider an example, based on H.J. McCloskey’s classic hypothetical, of consequentialism perhaps justifying punishment of the innocent. Suppose a particularly heinous crime is committed (whose perpetrator is unknown) and the community erupts in violence and riots in which many innocents are killed. If it is the case that many more innocents will be killed in the violence until a plausible suspect is arrested, convicted, and punished, a consequentialist “must conclude that he has a duty to bear false witness in order to bring about the punishment of an innocent person.” 143 Thus, if many innocent lives may be saved, consequentialism justifies the punishment of an innocent person. 143

2. Consequentialist Responses.—Consequentialist responses might be classified so as to fall into three categories: (i) denial, (ii) acceptance with a defense, and (iii) a countercharge that retributivism also intentionally punishes the innocent.

a. Denial.—Consequentialists deny the charge in a variety of ways. First, contrary to the general view, Benthamite classical utilitarianism does not require, or even allow for, the punishment of the innocent. 144

142 McCloskey, Non-Utilitarian, supra note 120, at 127.

143 For a defense of consequentialism in reply to McCloskey’s example, see T.L.S. Sprigge, A Utilitarian Reply to Dr. McCloskey, 8 INQUIRY 272 (1965), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 66, 70–75 (expressing skepticism that punishment of the innocent scapegoat, in McCloskey’s example, would actually produce, on balance, the most preferable consequences).

144 Frederick Rosen notes that “[i]f one examines Bentham’s writings on punishment and the discussions surrounding them, it soon becomes obvious that punishing the innocent was not originally envisioned as a problem in his theory.” Frederick Rosen, Utilitarianism and the Punishment of the Innocent: The Origins of a False Doctrine, 9 UTILITAS 23, 25 (1997) (footnote omitted). “Nor was it an issue for Kant, who feared that the guilty would go unpunished or be insufficiently punished with the adoption of utilitarianism rather than that the innocent would be punished.” Id. at 25 n.7.

On the general view that utilitarianism justifies punishment of the innocent, Rosen argues that this mistaken view is a result of a misunderstanding of Benthamite classical utilitarianism:

The mistake is to suppose that utilitarianism justifies punishment by deterrence alone, while deterrence in fact serves as one of several elements in the utilitarian theory of punishment. The mistake comes from ignoring the pleasure-pain dimension of traditional Benthamite utilitarianism and from thinking that utilitarianism begins with the general principle that utility (the maximization of happiness) is the sole criterion of right action, which is then applied to topics like punishment in a “top-down” fashion. In fact, classical Benthamite utilitarianism works in a “bottom-up” fashion from secondary principles to the principle of utility.

Id. at 26.

One of these “secondary principles,” justified by the principle of utility, is that the proportion between crimes and punishments should reflect the harm caused to the life and security of members of society. Id. at 32. Thus, “if there was no harm, there would not be a crime, and without a crime, punishment would be excluded.” Id. at 33. Since an innocent person would not commit a harm, there would be no crime and “punishment of the innocent would be generally excluded.” Id.

That punishment of the innocent would only be “generally excluded,” but not completely excluded,
It is more an institutional theory of law, not a theory of ethics governing the actions of individuals.\textsuperscript{145} As such, Benthamite utilitarianism rests on notions of transparency and openness of government actions which would preclude deception and manipulation of the general public.\textsuperscript{146} This, in turn, is claimed to preclude the secret framing of innocents.\textsuperscript{147} It is not clear, however, that the charge may be evaded this easily in light of Bentham's express statements that seem to allow for the possibility of justifying intentional punishment of the innocent.\textsuperscript{148} Second, punishment should be defined in such a way as to make the guilt of the person punished part of the definition of punishment.\textsuperscript{149} In this way, punishment of the innocent is a logical impossibility.\textsuperscript{150} As discussed above, this "definitional stop"\textsuperscript{151} simply assumes away the difficult issue of justification.\textsuperscript{152}

Third, punishment of the innocent would be counterproductive.\textsuperscript{153} Once it became known that innocents as well as offenders were punished, punishment would lose (some of) its deterrent value.\textsuperscript{154} There would be less reason to avoid committing criminal acts if that avoidance rendered one

allows the retributivist criticism to gain traction. Though perhaps retributivists have exaggerated the frequency with which innocents would be punished under consequentialism, their argument gains a foothold merely by the possibility that consequentialism might justify punishment of the innocent. And Bentham does seem to explicitly state that justifying punishment of the innocent by the principle of utility is possible. See infra note 161 and accompanying text.

\textsuperscript{145} The mistaken assumption of retributivists criticizing utilitarianism for justifying punishment of the innocent is that "if it is good for governments to seek to establish utility-maximizing institutions and policies it must be good for individuals to seek to perform utility-maximizing acts." Binder & Smith, supra note 27, at 120. As Binder and Smith further explain:

[T]he charge of framing the innocent rests on a misunderstanding of utilitarian penology as an application of an "act-utilitarian" ethic governing individual behavior. We contend that utilitarianism began as a normative theory of law and legal process aiming at not just happiness in general, but also at security in particular . . . .

\textit{Id.} at 118–19.

\textsuperscript{146} "[Benthamite utilitarianism is] methodologically committed to publicity, regularity and representativeness of legal decisionmaking. . . . Indeed, utilitarian penology cannot endorse any program of official manipulation of the public that restricts information to a putatively utilitarian elite." \textit{Id.} at 119 (footnote omitted).

\textsuperscript{147} In response to the retributivist charge that utilitarianism justifies the punishment of the innocent, Binder and Smith conclude "[t]hat charge is a frame-up." \textit{Id.} at 123. As Binder and Smith maintain:

"[U]tilitarian penology cannot endorse punishment of the innocent, which violates either the security aim, or the publicity condition, or both . . . . It is simply not true that eschewing unnecessary punishment of the guilty logically entails willingness to punish the innocent." \textit{Id.} at 119.

\textsuperscript{148} See infra note 161 and accompanying text.

\textsuperscript{149} See, e.g., Quinton, \textit{supra} note 41, at 10 ("The infliction of suffering on a person is only properly described as punishment if that person is guilty.").

\textsuperscript{150} \textit{Id.} at 12 ("[G]uilt is a logically necessary condition of punishment.").

\textsuperscript{151} See \textit{supra} note 44 and accompanying text.

\textsuperscript{152} See \textit{supra} notes 40–49 and accompanying text.

\textsuperscript{153} HOBBS, \textit{supra} note 21, at 249 ("[I]f there can arrive no good to the commonwealth by punishing the innocent."); C.L. Ten, \textit{supra} note 35, at 17–18; Benn, \textit{supra} note 45, at 31; Rawls, \textit{supra} note 1, at 11–13; C.L. Ten, \textit{supra} note 29, at 367.

\textsuperscript{154} Benn, \textit{supra} note 45, at 31; Rawls, \textit{supra} note 1, at 12.
subject to punishment anyway. Therefore, consequentialists argue, punishment of the innocent is not justified by their theories. But retributivists respond that punishment of the innocent would still be justified as long as the general public never found out that innocents were being punished. If innocents were punished, and the general public never found out, the punishment of the innocent would not undermine deterrence but promote deterrence. Consequentialists claim that the systematic punishment of the innocent would inevitably be found out. Even so, retributivists rejoin, a consequentialist would always be tempted to do it just once, figuring that one instance would escape detection. Even if consequentialist theories of punishment would not ever punish the innocent, they fail to provide a satisfactory account of why the innocent should not be punished. The proper rationale should address the injustice, and not merely the inefficacy, of punishing the innocent. That deterrence might be undermined by punishing the innocent fails to adequately explain our intuitive abhorrence to the moral injustice of punishing the innocent.

b. Acceptance with a Defense.—Rather than deny the charge that consequentialism justifies punishment of the innocent, some consequentialists might accept the charge but defend it by invoking a broader perspective of what constitutes moral injustice. While punishment of the innocent may be unjust, what is also unjust is retributivism’s insistence on inflicting punishment on an offender even if no good was to come of it and despite the needless suffering of future crime victims which might have been avoided by a deterrence-based approach to punishment. For example, although Bentham explicitly maintains that the innocent ought to be punished in some special circumstances where utility could thereby be

155 Benn, supra note 45, at 31; Rawls, supra note 1, at 12.
156 One prominent retributivist, however, suggests an additional reason why punishment of the innocent might be counterproductive for consequentialist theories of punishment. Hegel claimed that an unjust threat of punishment might not deter crime, but rather, induce it. The unjust threat of punishment “may ultimately provoke someone into demonstrating his freedom in defiance of it . . . .” Hegel, supra note 22, at 126 (§ 99). In other words, punishment of the innocent may be counterproductive by promoting criminal violations rather than deterring them.
157 Murphy & Coleman, supra note 54, at 122; C.L. Ten, supra note 35, at 18.
158 Duff, supra note 1, at 160–161; Benn, supra note 45, at 31; C.L. Ten, supra note 29, at 367.

One possible consequentialist reply is to counter that retributivism cannot explain how much society should spend both to increase the number of the guilty that are punished and to decrease the number of innocents that are punished. Kaplow & Shavell, supra note 28, at 1240–1242. Presumably a retributivist “would be willing to spend, say, one penny to ensure that an innocent individual be acquitted rather than executed by mistake. But what is the most he would spend—the entire GDP?” Id. at 1241. While consequentialism could easily reach a resolution, id. at 1242 n.686, the claim is that retributivism has no satisfactory answer, nor could it even supply an answer to such a question in principle. Id. at 1242–43. And as a result, retributivism “is seriously incomplete.” Id. at 1242 (footnote omitted).
159 Duff, supra note 1, at 160–161; Benn, supra note 45, at 31.
160 One example might be collective criminality where it is known that some members of a group are responsible for a crime and some are innocent, but it is not known which are which. If the punish-
maximized, retributivist punishment is unjust both to the general public and to the offender. In other words, though deterrence theories may be unjust in justifying the intentional punishment of innocents, retributivism is unjust in exposing the innocent general public (innocent future crime victims) to a greater risk of victimization.

The second controversial consequentialist response, and perhaps the "more promising," is to embrace the charge and defend it on grounds of utility. If punishing an innocent truly yields a better state of affairs than not punishing, then of course we would want to punish that innocent. Why would retributivism prefer a worse state of affairs over a better state of affairs? Though this response has been subject to some ridicule—consequentialists have been said to "outsmart" themselves—nonetheless it does have an undeniable common-sense appeal. It certainly might be dis-

1\footnote{Bentham, supra note 63, at 476:}

ment of all the members of the group would maximize utility, then the intentional punishment of the innocent might be justified for Bentham. \textsc{Primoratz}, supra note 1, at 26.

\footnote{To inflict punishment when . . . the infliction of such punishment is avoidable, is in the case of the innocent, contrary to the principle of utility. . . . To punish, where . . . such punishment is unavoidable, is not in either case contrary to the principle of utility;—not in the case of the guilty: no, nor yet in the case of the innocent.}


\footnote{Bagoric & Amarasekara, supra note 12, at 141–42.}

\footnote{For discussion of this claim, see \textsc{Honderich}, supra note 89, at 66–75.}

\footnote{See C.L. Ten, supra note 29, at 367 ("[R]etributivists have difficulty in explaining why the guilty should be punished at all if punishment fails to produce any good consequences . . . .").}

\footnote{\textsc{Primoratz}, supra note 1, at 54.}

\footnote{In response to the retributivist criticism that consequentialism justifies punishment of the innocent, a consequentialist, J.J.C. Smart, embraced the criticism. Smart defends the consequentialist implication that to avert great misery, punishment of the innocent might be justified:}

\footnote{In response to Smart’s argument, the philosopher Daniel C. Dennett included the term "outsmart," as a pun on Smart’s name and as a comment on the wisdom of Smart’s argument, in his satirical philosophical dictionary. \textsc{Daniel C. Dennett}, \textsc{The Philosophical Lexicon} (8th ed. 1987). The entry is as follows: "[O]utsmart, v. —To embrace the conclusion of one’s opponent’s \textit{reductio ad absurdum} argument. ‘They thought they had me, but I outsmarted them. I agreed that it was sometimes just to hang an innocent man.’" \textit{Id}. Dennett’s dictionary is available online at http://www.blackwellpublishers.co.uk/LEXICON/default.htm.}

874
puted whether punishment of an innocent would lead to a better state of affairs. But if it did not then consequentialism would not justify punishment of the innocent. But the consequentialist argument is that if punishment of an innocent did lead to a better state of affairs, then that punishment is justified.

The argument may be put even more persuasively. If a retributivist system of punishment dramatically increased the incidence of crime and the number of innocent victims who were victimized by crime, as opposed to a system of no punishment or a consequentialist system of punishment, retributivism would still be committed to punishing.\(^\text{168}\) And despite the possibility "that more innocents would be punished in a society governed by retributivist, rather than utilitarian, principles,"\(^\text{169}\) retributivists would be committed to retributivism and the punishment of a greater amount of innocents.\(^\text{170}\) Even if retributivism led to (i) greater victimization of innocent victims,\(^\text{171}\) (ii) a greater number of guilty who were acquitted,\(^\text{172}\) and (iii) a greater number of innocents punished,\(^\text{173}\) as opposed to a consequentialist

---

\(^{168}\) C.L. Ten, supra note 29, at 369.

\(^{169}\) Ezorsky, supra note 72, at xviii.

\(^{170}\) Ezorsky explains why it would be possible:

[F]or utilitarians, punishment is justified only if, by comparison with other alternatives, e.g., treatment, punishment would maximize utility. Suppose a painless but expensive pill were devised which cured any propensity to commit crime. The pill's utility surpassed that of other alternatives to punishment and slightly exceeded that of punishment itself. In that case utilitarians should endorse adoption of the crime cure pill and abolition of punishment. But retributivists believe that punishment of the guilty is either necessarily just or has some intrinsic value. In that case retributivists might refuse to substitute the crime cure pill for punishment. But juries would most likely remain fallible. Thus some innocents would be punished along with the guilty. Hence if retributivist principles held sway, some innocents would, as a consequence, be punished. While if utilitarians had their way, not a single innocent would suffer punishment. Id. at xviii–xix.

\(^{171}\) Kaplow & Shavell, supra note 28, at 1243–44. Louis Kaplow and Steven Shavell question retributivism's indifference to higher crime rates in determining just punishments:

[S]uppose that if the fair [retributivist] punishment for murder is applied, ten murders will be committed, whereas if a greater and unfair [consequentialist] punishment is threatened, none will be. Given that murder is considered by the retributivist to be a wrongful act—that the world in which an act of murder occurs is less just than one in which it does not—it is hard to understand how the retributivist can say that it does not matter in determining whether the supposedly fair punishment is indeed fair that, as a consequence of insisting on fair punishment, ten murders will occur.

Id. at 1244–45. Kaplow and Shavell argue:

In the example just mentioned [see supra note 171], when the fair punishment is employed and ten murders occur, some of the murderes may escape . . . let us say that five do so. Thus, for the retributivist to recommend the fair punishment for murder means that he favors a regime under which five people suffer no sanction for their wrong, which itself is unfair under retributive theory. But if the unfairly high sanction is announced, it will by hypothesis discourage all murders, so that no individuals will in fact experience an unfair sanction (either too high a sanction or getting away scot-free).

Id. (footnote omitted).

\(^{172}\) By extending Kaplow & Shavell's reasoning, see supra notes 171–172, it could be shown that
system of punishment, retributivists would still be committed to retributivist punishment.\textsuperscript{174}

Or instead suppose that not punishing an innocent would lead to catastrophic consequences, such as the destruction of an entire society or nation.\textsuperscript{175} In such cases, our intuition might be that even though the one should not be sacrificed to save the many, nonetheless the one should be sacrificed to prevent all from being lost.\textsuperscript{176} Throughout history,\textsuperscript{177} as well as today,\textsuperscript{178} we have often sacrificed the few to avoid catastrophic consequences and often much less severe results—consider pretrial detention, quarantine, and involuntary enlistment, to name only a few.\textsuperscript{179}

Retributivists claim that the demands of justice are absolute and no innocent may be punished even if it were to prevent the extinction of mankind. This is the view of Kant, who declared that “if justice goes [perishes] there is no longer any value in men’s living on the earth.”\textsuperscript{180} But this may be too impracticable and implausible.

more innocents will be punished under the fair or retributivist system than the unfair or consequentialist system. Suppose that in their example, under the retributivist system, some innocents will be brought to trial and convicted for some of the ten murders, let us say two innocents. But under the consequentialist system, which by hypothesis discourages all murder, there will be no crimes to be prosecuted and thus no innocent will be convicted and punished. As a result, a retributivist is forced to favor a system in which more innocents will be punished than under a consequentialist system.

For Kaplow & Shavell’s own argument that more innocents might be mistakenly punished under a retributivist system than a consequentialist system, see Kaplow & Shavell, \textit{supra} note 28, at 1266–69.

\textsuperscript{174} A retributivist, however, could simply reply that the ends do not justify the means. Even when the use of unfair or consequentialist means may produce a better state of affairs, we are not morally permitted to do wrong (by punishing unfairly) to attain the good.

\textsuperscript{175} \textit{See, e.g., John} 11:47–53, describing an innocent punished to prevent catastrophic consequences.

\textsuperscript{176} \textit{W.D. Ross, The Right and the Good} 61 (1967) (“[T]he interests of the society may be so deeply involved as to make it right to punish an innocent man ‘that the whole nation perish not.’”).

\textsuperscript{177} Consider the following instance of a lesser number of innocents being sacrificed to save the many:

A pointed example is the decision by the British Prime Minister of the day, Winston Churchill, to sacrifice the lives of the residents of Coventry in order not to alert the Germans that the English had deciphered German radio messages. On 14 November 1940 the English decoded plans that the Germans were about to air bomb Coventry. If Coventry were evacuated or its inhabitants advised to take special precautions against the raid, the Germans would have known that their code had been cracked and the English would have been unable to obtain future information about the intentions of its enemy. Churchill elected not to warn the citizens of Coventry, and many hundreds were killed in the raid that followed. Their lives were sacrificed in order not to reveal a secret that would hopefully save many more lives in the future.

\textit{Bagoric & Amarasekara, \textit{supra} note 12, at 143.}

\textsuperscript{178} On September 11, 2001, President Bush ordered that hijacked civilian passenger planes heading toward significant targets with hostile intent be shot down by American fighter jets. Apparently, the basis for the order is that it is preferable to knowingly kill hundreds in order to save the lives of thousands. \textit{See Dan Balz & Bob Woodward, America's Chaotic Road to War}, \textit{WASH. POST.}, Jan. 27, 2002, at A1; Bradley Graham, \textit{Military Alerted Before Attacks}, \textit{WASH. POST.}, Sept. 15, 2001, at A18.

\textsuperscript{179} The retributivist objection to consequentialist theories countenancing intentional punishment of the innocent “loses its force when it is shown that punishing the innocent is in fact no worse than other activities we condone.” \textit{Bagoric & Amarasekara, \textit{supra} note 12, at 143.}

\textsuperscript{180} \textit{KANT, supra} note 22, at 141.
c. Retributivism Does It Too.—As a final response to the charge that their theory justifies punishing innocents, consequentialists might claim that regardless of the truth of that claim, retributivism also justifies intentional punishment of the innocent. Unlike Kant, some retributivists might concede the permissibility of punishing the innocent in order, for example, "that the whole nation perish not."\textsuperscript{181} This seems to be the view of Michael Moore, perhaps the most committed advocate of retributivism among legal scholars,\textsuperscript{182} who would justify punishment of the innocent if the adverse consequences stemming from not punishing were severe enough.\textsuperscript{183} But does not this concession\textsuperscript{184} defuse retributivism's critique of consequentialism's punishment of the innocent?\textsuperscript{185} After all, that even retributivists suggest that we turn to consequentialism in the very moments of peril and crisis—when the heat is on—suggests consequentialism's plausibility.\textsuperscript{186} A retributivist could still argue that retributivism might permit the punishment of the innocent when it is the only way to prevent a catastrophic state of affairs, whereas consequentialism might justify punishment of the innocent whenever it produced even a slightly preferable state of affairs.\textsuperscript{187} Though a

\textsuperscript{181} ROSS, supra note 176, at 61.

\textsuperscript{182} Douglas Husak, Retribution in Criminal Theory, 37 SAN DIEGO L. REV. 959, 960 (2000) ("No contemporary criminal theorist rivals Moore in his unqualified enthusiasm for retribution.").

\textsuperscript{183} Moore rejects the absolutism of deontological duties which is "often attributed to Kant, who held that though the heavens may fall, justice must be done." MOORE, supra note 73, at 719. Moore explains that at a certain threshold of grievous harm, consequential principles outweigh deontological ones:

Despite my nonconsequentialist views on morality, I cannot accept the Kantian line. It just is not true that one should allow a nuclear war rather than killing or torturing an innocent person. It is not even true that one should allow the destruction of a sizable city by a terrorist nuclear device rather than kill or torture an innocent person. To prevent such extraordinary harms extreme actions seem to me to be justified.

\textsuperscript{184} For other nonconsequentialists also seeming to make this concession see NOZICK, supra note 36, at 29 n.*; Thomas Nagel, War and Massacre, in THOMAS NAGEL, MORTAL QUESTIONS 53, 56 (1979).\textsuperscript{185} For Moore's argument that it does not collapse into consequentialism see MOORE, supra note 73, at 721–22. As Moore explains:

[T]here is no collapse of agent-relative views [deontology] into consequentialism just because morality's norms can be overridden by horrendous consequences. A consequentialist is committed by her moral theory to saying that torture of one person is justified whenever it is necessary to prevent the torture of two or more. The agent-relative view, even as here modified, is not committed to this proposition. To justify torturing one innocent person requires that there be horrendous consequences attached to not torturing that person—the destruction of an entire city, perhaps, or of a lifeboat of building full of people.

\textsuperscript{186} Bagoric & Amarasekara, supra note 12, at 143 ("The decisions we make in a real life crisis are the best evidence of the way we actually prioritise [sic] important competing principles and interests.").

\textsuperscript{187} PRIMORATZ, supra note 1, at 60.
Retributivists might make the concession in extraordinary circumstances, retributivism per se, as a theory, would have difficulty in justifying punishment of the innocent to avoid catastrophic consequences, without collapsing into consequentialism.\textsuperscript{188}

Though not directly addressing the claim that retributivism would collapse into consequentialism, H.L.A. Hart offers a partial way out for retributivism:

In extreme cases [of impending social catastrophe] many might still think it right to resort to these expedients [punishing the innocent to avert the catastrophe] but we should do so with the sense of sacrificing an important principle. We should be conscious of choosing the lesser of two evils, and this would be inexplicable if the principle sacrificed to utility were itself only a requirement of utility.\textsuperscript{189}

Retributivists have seized this argument to claim that even if retributivism were to justify punishment of the innocent on consequentialist grounds, retributivists "would also realize that this imposed a significant moral cost, a significant wrong done to the innocent person, which the utilitarian cannot recognize."\textsuperscript{190} Such a version of retributivism might be thought of as consequentialism with the guilt. Thus even where retributivism resorts to consequentialist reasoning, rather than conceding that in particular extreme cases consequentialism gets it right, retributivists proclaim that retributivism does consequentialism better. That is, retributivists make a virtue out of theoretical inconsistency (its collapse into consequentialism) by trumpeting that at least they suffer the appropriate pang of conscience. Regardless of whether retributivists embrace the use of consequences with open arms or with hands wringing, the resort to consequences to justify punishment still may smack of consequentialist reasoning.

In addition to claiming that retributivism would collapse into consequentialism, consequentialists might also deny Hart's claim. That is, consequentialists are aware of the sacrifice being made and consciously choose the lesser of two evils when an innocent is punished. Bentham states that all punishment is evil; punishment may be justified, if at all, only to the extent that it may prevent some greater evil.\textsuperscript{191} So contrary to Hart, for conse-

\textsuperscript{188} The view that a retributivist or deontologist might justify intentional wrongdoing in order to avoid catastrophic consequences, without collapsing into consequentialism, has been termed "threshold deontology." See Moore, supra note 73, at 719–24. That is, the consequences of our actions are not sufficient to justify them unless and until a sufficiently high threshold has been met whereby the consequences of an action, or the failure to take action, are sufficiently severe. Id. For skeptical discussions of whether threshold deontology is tenable, see Nancy (Ann) Davis, Contemporary Deontology, in A Companion to Ethics 205, 214-16 (Peter Singer ed., 1993); Larry Alexander, Deontology at the Threshold, 37 San Diego L. Rev. 893, 895 (2000).

\textsuperscript{189} Hart, supra note 11, at 12. For Hart's explanation of the same point in a slightly different way see id. at 81.

\textsuperscript{190} Duff, supra note 2, at 423.

\textsuperscript{191} See supra note 68–71 and accompanying text.
Consequentialists all punishment (not just punishment of the innocent) constitutes
a conscious choosing of what is the lesser evil.\textsuperscript{192} And if consequentialism
does justify punishment of the innocent, it would do so with the sense that
important interests are being sacrificed.\textsuperscript{193}

Consequentialists might also turn back on retributivism Hart’s argument\textsuperscript{194}
that when we intentionally impose pain, hardship, deprivation, and
suffering on someone—that is, when we punish—we should do so with the
sense of sacrificing important principles and interests. We should be con-
scious of choosing the lesser of two evils, and this would be inexplicable if
the principles or interests sacrificed were not a sacrifice at all but were con-
sidered inherently good. Since, for retributivism, punishment is inherently
or intrinsically good or right, retributivism has no sense that punishment in-
volves any choosing of the lesser evil or involves any sacrifice at all.

Another way that retributivism is claimed to justify intentional pun-
ishment of the innocent is that infliction of punishment on a guilty offender
will most likely inflict pain and suffering on his or her friends and family
who are innocent of the offense.\textsuperscript{195} Thus, “[a] retributively just punishment
for one person . . . would thus involve a retributively unjust punishment for
several.”\textsuperscript{196} Of course, a retributivist might reply that such suffering by
friends and family members fails to fall within the notion of what is meant
by punishment. Punishment is only inflicted on a person for wrongdoing or
an offense committed by that person. Though such a definition would pre-
clude the pain suffered by friends and family members of an offender from
falling within the ambit of punishment, such a definition would constitute a
“definitional stop”\textsuperscript{197} making punishment of the innocent a logical impos-
sibility.\textsuperscript{198} Thus, if a retributivist is to answer the criticism by relying on a
particular definition of punishment, the retributivist charge that conseque-

\textsuperscript{192} While it is true that consequentialism cannot punish an innocent without the awareness that an
important principle has been sacrificed, no theory with but a single principle can sacrifice it to another
principle without being inconsistent. Hart seems to be criticizing consequentialism for nothing more
than theoretical consistency.

\textsuperscript{193} See, e.g., Binder & Smith, supra note 27, at 136 (“[E]ven if a utilitarian were to accept punish-
ment of the innocent in some exceptional circumstance, she would do so with considerable regret.”);
Shavell & Kaplow, supra note 28, at 1269 (“A [A]n analyst employing welfare economics [consequential-
ism] acts with a heavy heart when proposing a system of punishment that may involve any punishment
of the innocent.”).

\textsuperscript{194} See supra text accompanying note 189.

\textsuperscript{195} Binder & Smith, supra note 27, at 130 (“Punishing an offender usually does inflict suffering on
an offender’s family, an undeserved consequence of retribution which retributivists should not evade by
means of a definitional stop.”). For further explanation of the concept of a “definitional stop,” see supra
note 44 and accompanying text.

\textsuperscript{196} This is because if punishment can only be imposed on a person for wrongdoing or an offense
committed by that same person, an innocent person, who by definition has not committed the offense,
cannot be subject to punishment so conceived.
tialism justifies the innocent will be rendered logically impossible.\textsuperscript{199}

To conclude, if retributivism is willing to justify punishment of the innocent in special circumstances, the difference between consequentialism and retributivism may be small. In most instances, punishment of the innocent, under consequentialism, would be groundless, unprofitable, and needless.\textsuperscript{200} Because punishing an innocent causes more suffering than punishing a guilty person, Bentham states that “[t]o inflict punishment when . . . the infliction of such punishment is avoidable, is in the case of the innocent, contrary to the principle of utility.”\textsuperscript{201} It is only in special circumstances, then, that consequentialism would justify punishing the innocent.\textsuperscript{202} But if it is only in special circumstances that both retributivism and consequentialism justify the intentional punishment of the innocent, how is retributivism morally superior to consequentialism? While admittedly the special circumstances warranting punishment for consequentialism might be broader than for retributivism, the difference is only a matter of degree. A consequentialist would justify punishment of the innocent, if at all, only in those special circumstances where greater good would result; a retributivist might restrict such special circumstances to the prevention of catastrophic consequences. But if so, there is not a significant moral or qualitative difference between the two theories as to punishment of the innocent, but only a matter of degree.

An additional facet of the debate over the punishment of the innocent\textsuperscript{203} is the significance of retributivism mistakenly punishing the innocent, which we will address in Part IV. The next subpart, II.B, will present the novel argument that retributivism justifies intentional punishment of the innocent, not merely to prevent catastrophic consequences, but under substantially broader circumstances.

\section*{B. The Moralistic Version of Retributivism}

Although most (if not all) retributivist theories justify the imposition of punishment based on the offender deserving the punishment, retributivists differ over what makes a wrongdoer or offender deserving of punishment.

\textsuperscript{199} A retributivist might more profitably answer the criticism by invoking some variant of the doctrine of double effect, see infra notes 306–308 and accompanying text, which we will focus on in Part IV. In short, the pain and suffering inflicted on the offender’s innocent friends and family by the offender being punished might be punishment after all. However, the infliction of punishment on the friends and family members is merely foreseen but not intended and thus, according to the doctrine of double effect, permissible.

\textsuperscript{200} See supra notes 78, 161.

\textsuperscript{201} Bentham, supra note 63, at 476.

\textsuperscript{202} Bentham suggests that ordinarily punishment of the innocent would be unprofitable. “Punishment then, as applied to delinquency, may be unprofitable . . . . By the danger there may be of its involving the innocent in the fate designed only for the guilty.” BENTHAM, THE PRINCIPLES, supra note 57, at 316 (emphasis added). For a similar expression of this view, see id. at 318.

\textsuperscript{203} For other consequentialist responses, see C.L. TEN, supra note 35, at 13–14.
Perhaps the first to supply labels to the differing forms of retributivism as a function of their differing conceptions of what it is that makes someone deserve punishment, Joel Feinberg distinguished between “moralistic” and “legalistic” retributivism. Under moralistic retributivism, moral guilt or (moral) desert is a necessary and sufficient condition for the imposition of justified punishment, even if the moral wrongdoing does not constitute the commission of a criminal offense. In contrast, legalistic retributivism justifies the imposition of punishment for the commission of a criminal offense, even if the violation does not constitute moral wrongdoing. To continue Feinberg’s typology of forms of retributivism as a function of what it is that makes a wrongdoer or offender deserve punishment, we might also speak of two additional versions of retributivism. A conjunctive moralistic and legalistic retributivism, or combined retributivism, would justify punishment when the wrongdoer or offender committed an act that constitutes or entails both (i) moral wrongdoing or moral guilt, and (ii) the commission of a criminal offense. Another conceivable form of retributivism, which so far as I know has not previously been identified and has no adherents, is a disjunctive moralistic or legalistic retributivism, in which


205 Id. Feinberg supplies two accounts of moralistic retributivism. The more general account is as follows:

Punishment is justified only on the ground that wrongdoing merits punishment. It is morally fitting that a person who does wrong should suffer in proportion to his wrongdoing. That a criminal should be punished follows from his guilt, and the severity of the appropriate punishment depends on the depravity of the act. The state of affairs where a wrongdoer suffers punishment is morally better than one where he does not, and is so irrespective of consequences.

Justification, according to these accounts, must look backward in time to guilt rather than forward to “advantages”; the formulations are rich in moral terminology (“merits,” “morally fitting,” “wrongdoing,” “morally better”); there is a great emphasis on desert. For these reasons we might well refer to this as a “moralistic” version of the retributive theory.

Id. at 515 (footnote omitted). The comparatively more specific account, termed “pure moralistic retributivism,” Feinberg describes “as consisting (at least) of the following propositions”:

1. Moral guilt is a necessary condition for justified punishment.
2. Moral guilt is a sufficient condition (“irrespective of consequences”) for justified punishment.
3. The proper amount of punishment to be inflicted upon the morally guilty offender is that amount which fits, matches, or is proportionate to the moral gravity of the offense.

Id. at 516. By “pure” moralistic retributivism, Feinberg means only that it does not add consequentialist considerations. Id. at 515 (“Both moralistic and legalistic retributivism have ‘pure’ and ‘impure’ variants. In their pure formulations, they are totally free of utilitarian admixture.”).

206 As Feinberg explains:

[Legalistic retributivism justifies punishment for] lawbreaking, not (necessarily) for wrongdoing. Legalistic retributivism holds that the justification of punishment is always to be found in the fact that a rule has been broken for the violation of which a certain penalty is specified, whether or not the offender incurs any moral guilt. The offender, properly apprised in advance of the penalty, voluntarily assumes the risk of punishment, and when he or she receives comeuppance, he or she can have no complaint.

Id.

207 This form of retributivism will be discussed infra, Part III.
punishment is justified if an actor’s conduct constitutes either (i) moral wrongdoing or moral guilt, or (ii) the commission of a criminal offense.\textsuperscript{208}

It is not entirely clear which particular retributivists subscribe to which forms of retributivism, or which form is more influential. Not only do most retributivists not identify their accounts as either moralistic or legalistic, they also fail to carefully identify the sufficient conditions for justified punishment, from which we might identify which version they endorse. Kathleen Moore explains a possible source of the difficulty: “[T]he should not be inferred that any given retributivist holds one to the exclusion of the other. Most retributivist positions can be found along a continuum connecting the two views . . . .”\textsuperscript{209} Although Feinberg supplies only one example each of moralistic retributivism\textsuperscript{210} and legalistic retributivisn,\textsuperscript{211} he does suggest that moralistic retributivism is the more prevalent form.\textsuperscript{212} Martin Gardner suggests the opposite—that legalistic retributivism is more common.\textsuperscript{213} Moore asserts that retributivists are “split” between the two versions.\textsuperscript{214}

The moralistic version of retributivism justifies the punishment of morally culpable wrongdoers, even if their moral wrongdoing is not sufficient to warrant conviction for a criminal offense.\textsuperscript{215} For example, Michael Moore maintains that morally culpable wrongdoers must be punished in accordance with their just deserts.\textsuperscript{216} Another commentator declares that punishment must “be justified by reference to a moral wrong which the person punished has allegedly committed.”\textsuperscript{217} This is because a defendant may be found guilty in an invalid legal proceeding, or even in a valid adjudication, but be nonetheless innocent as a matter of fact. Or, a defendant may be fac-

\textsuperscript{208} Because this form has no known adherents and does not avoid any of the problems that will be raised for the other versions of retributivism, it will not be discussed further.


\textsuperscript{210} Feinberg, supra note 204, at 515 (after the explanation of the account of moralistic retributivism, Feinberg cites Rawls, supra note 1, at 4, 5).

\textsuperscript{211} Feinberg, supra note 204, at 515 (citing, as an example of an account adhering to legalistic retributivism, Mabbott, supra note 66).

\textsuperscript{212} Feinberg, supra note 204, at 516 (referring to moralistic retributivism as “that popular variant of the theory”).

\textsuperscript{213} Gardner, supra note 4, at 781 (noting that a particular theory is “somewhat uncommon in its emphasis on moral, as opposed to merely legal, desert”).

\textsuperscript{214} MOORE, supra note 209, at 94. Referring to the moralistic and legalistic versions, Moore notes that “[t]hrough all the different permutations of retributivist theory, there can be seen two different, identifiable traditions or strains.” Id.

\textsuperscript{215} See, e.g., C.L. TEN, supra note 35, at 46 (“We can then formulate retributive theories of punishment as those theories which maintain that punishment is justified because the offender has voluntarily committed a morally wrong act.”).

\textsuperscript{216} Moore, supra note 119, at 179 (“[P]unishment is justified by the moral culpability of those who received it.”); MOORE, supra note 73, at 173 (“[T]he moral desert of an offender is a sufficient reason to punish [the offender].”).

\textsuperscript{217} DUFF, supra note 1, at 153. Duff, however, would not subscribe to moralistic retributivism. Id. at 153–54.
Deterring Retributivism

Finally and legally guilty of committing an act that may have been unjustly defined to be a crime. (For example, a status crime such as being a drug addict\textsuperscript{218} or a crime of engaging in consensual, homosexual intercourse by adults in the privacy of one’s home.\textsuperscript{219}) To avoid justifying the punishment of moral innocents, moralistic retributivism\textsuperscript{220} limits the application of its justification of punishment to those who are morally culpable, but not necessarily legally convicted, wrongdoers.\textsuperscript{221} This limitation triggers two problems. Applying the moralistic version of retributivism to actual, legal punishment is both underinclusive and overinclusive. First, it is underinclusive because some of those validly adjudicated to be guilty may not justifiably be punished. This underinclusiveness, however may not be a serious problem. As discussed above, retributivism may have a compelling reason for not justifying punishment of those who are legally guilty but not morally culpable: they are morally innocent. Thus, moralistic retributivism’s underinclusiveness may be less a defect than a virtue. Its overinclusiveness, however, is more serious in that it justifies punishment of those who are morally culpable yet legally innocent.

To truly see the problem of moralistic retributivism justifying punishment of the morally guilty but legally innocent, it is helpful to consider those morally culpable wrongdoers who have a valid legal defense that does not negate their moral culpability. These so-called extrinsic\textsuperscript{222} or nonexculpatory defenses\textsuperscript{223} include diplomatic immunity, immunity pursuant to a plea agreement, double jeopardy, the statute of limitations, and the principle of legality.\textsuperscript{224} It would seem that moralistic retributivism would justify the

\textsuperscript{218} See, e.g., Robinson v. California, 370 U.S. 660 (1962) (overturning the conviction of a defendant in violation of a statute making it a crime to be addicted to the use of narcotics).

\textsuperscript{219} See, e.g., Bowers v. Hardwick, 478 U.S. 186 (1986) (upholding the conviction of a defendant for violating a state statute prohibiting consensual homosexual intercourse, even in one’s home).

\textsuperscript{220} For other examples of moralistic retributivism, see Gardner, supra note 4, at 799 n.123 (citing, as examples, KANT, supra note 22; HONDERICH, supra note 89, at 13; JEROME HALL, GENERAL PRINCIPLES OF CRIMINAL LAW (2d ed. 1960); JOHN KLEINIG, PUNISHMENT AND DESERT 32-33 (1973)).

\textsuperscript{221} Moralistic retributivists typically do not specify precisely what constitutes moral wrongdoing sufficient to justify the imposition of punishment. Obviously, most moralistic retributivists would agree that murder, rape, and torture are moral wrongs. But they might disagree, for example, as to whether consensual homosexual intercourse, or abortion, is or is not a moral wrong. But what unites moralistic retributivists is that moral wrongdoing, however that may variously be understood, and even if it does not amount to a criminal offense, is sufficient to justify the imposition of punishment.

\textsuperscript{222} JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 184 (2d ed. 1995) (so-called extrinsic defenses “raise public-policy factors extrinsic to substantive criminal law doctrine” and which are unrelated to the culpability or wrongdoing of the defender).

\textsuperscript{223} PAUL ROBINSON, CRIMINAL LAW 569 (1997). Though in a sense any defense that successfully prevents a finding of guilt, as do the defenses above, is exculpatory, these defenses have come to be called nonexculpatory because they do not affect our substantive assessment of the defendant’s conduct. They are procedural defenses representing public policy interests.

\textsuperscript{224} Arguably, a defense based on the principle of legality might be thought to be an exculpatory defense that relates to a person’s culpability or blameworthiness. But for malum in se crimes (where the act, even if unprohibited, is obviously wrong), a defense based on legality may be considered to be an
punishment of a morally culpable wrongdoer despite the fact that she is legally innocent by virtue of satisfying one of the above defenses.

Retributivists acknowledge the problem. To avoid it, Michael Moore, for example, concedes that retributivists are not "monomaniacal" in the pursuit of the goals of retributivism. Punishing deserving wrongdoers is "subject to being overridden by other goods that outweigh the good of achieving retributive justice." One such good that, for Moore, outweighs the good of achieving retributive justice is the principle of legality. Therefore, despite moralistic retributivism justifying the punishment of a morally culpable wrongdoer that could not legally be punished because of legality concerns, a consideration extrinsic to retributivism, in this case legality, overrides the dictates of moralistic retributivism and precludes the imposition of punishment. In essence, Moore's move is that a valid legal adjudication of guilt is a precondition for punishment under moralistic retributivism. In this way, a moralistic retributivism system of punishment would be precluded from punishing a legal innocent.

This solution, however, is not entirely satisfactory. That a consideration extrinsic to moralistic retributivism precludes the imposition of punishment of legal innocents does not alter the fact that moralistic retributivism per se does justify the punishment of particular, identifiable legal innocents. Moralistic retributivism is saved from that unfortunate result only by something outside itself. Therefore, moralistic retributivism per se is still subject to the criticism that it justifies the intentional punishment of particular, identifiable legal innocents.

As a result, retributivism's principal argument against a consequentialist theory of punishment also applies to retributivism. Moralistic retributivism, no less than consequentialist theories, justifies the intentional punishment of particular, identifiable legal innocents. If such an abhorrent result renders consequentialist theories of punishment untenable, then moralistic retributivism should be considered equally untenable.

On the other hand, if employing a valid legal adjudication of guilt as a precondition for justifying punishment was considered a satisfactory solution, it might also serve as a solution for consequentialism. If we allow the valid legal adjudication of guilt to be a precondition of the application of retributive punishment—in order to avoid the intentional punishment of particular, identifiable legal innocents—then consequentialist theories could equally adopt that precondition. Though a consequentialist theory per se might justify the punishment of the innocent, it could equally avoid doing so by justifying punishment of only those who had been validly adjudicated to be legally guilty. As a result, with the precondition in place, consequential...
tional theories of punishment would no longer justify intentional punishment of the innocent. Retributivism’s principal argument against consequential theories of punishment would then no longer apply.

Either way, moralistic retributivism fares no better than consequentialism.

C. The Legalistic Version of Retributivism

To avoid the criticism of justifying the intentional punishment of particular, identifiable legal innocents, retributivists could simply shift to another version of retributivism—legalistic retributivism. Under this theory, punishment is justified not for those who with moral culpability commit wrongdoing regardless of legal culpability, but instead for those who are legally convicted of committing a criminal offense regardless of their moral culpability. Legalistic retributivism thereby avoids the intentional punishment of particular, identifiable legal innocents.

In avoiding the problem of moralistic retributivism, however, legalistic retributivism is subject to the claim that it is violative of the very principles underlying retributivism. Legalistic retributivism “would imply that punishment is sometimes morally legitimate precisely because it is retribution for the violation of a morally illegitimate law.” In finding legalistic retributivism to be incompatible with the essence of retributivism, C.L. Ten declared that “to restrict the justification of punishment to illegal acts, irrespective of the moral quality of these acts, would also sever the link between punishment and moral wrongdoing which lies at the heart of retributive theories.” And from the perspective of moralistic retributivism, legalistic retributivism is also underinclusive in failing to justify the punishment of morally culpable wrongdoing that is not defined to be a criminal offense.

Even though it avoids the problem incurred by moralistic retributivism, legalistic retributivism incurs one problem moralistic retributivism avoids. A defendant who is factually and legally guilty of committing an act which has been unjustly defined to be a crime—such as the status crime of being a drug addict or the crime of engaging in consensual, homosexual intercourse between adults in the privacy of one’s home—would not justifiably be punished under moralistic retributivism but would be justifiably punished under legalistic retributivism. As a result, legalistic retributivism is overinclusive in justifying the intentional punishment of particular, identifiable moral innocents.

Both versions of retributivism—moralistic and legalistic—justify the

---

228 See, e.g., Mabbott, supra note 66, at 28 ("No punishment is morally retributive or reformative or deterrent. Any criminal punished for any of these reasons is certainly unjustly punished. The only justification for punishing any man is that he has broken a law.").

229 PRIMORATZ, supra note 1, at 13.

230 C.L. TEN, supra note 35, at 72.
intentional punishment of particular, identifiable innocents. Moralistic retributivism justifies the intentional punishment of particular, identifiable legal innocents (those who have not even committed an act satisfying the definition of a criminal offense or those who have done so but have a valid, extrinsic defense). Legalistic retributivism justifies the intentional punishment of particular, identifiable moral innocents (those who are guilty of committing an act which is unjustly defined to be a criminal offense). Under either version, retributivism justifies the intentional punishment of particular, identifiable innocents.

D. Objection

It might be objected that moralistic and legalistic retributivism’s justification of the intentional punishment of particular, identifiable legal and moral innocents, respectively, is not problematic. Neither theory is punishing innocents relative to its own definition of innocent. Moralistic retributivism, justifying the punishment of actors committing morally culpable wrongdoing, is not designed to preclude the punishment of legal innocents and is properly indifferent to the punishment of legal innocents. Similarly, legalistic retributivism, justifying the punishment of actors committing criminal offenses, is not designed to preclude the punishment of moral innocents and is properly indifferent to the punishment of moral innocents. It would only be a failing of moralistic retributivism if it justified the punishment of moral innocents; and, it would only be a failing of legalistic retributivism if it justified the punishment of legal innocents. Neither theory is justifying the punishment of what each theory would consider to be innocents.

The objection is unpersuasive for a number of reasons. First, retributivists have criticized consequentialist theories for not only justifying punishment of the innocent but, more importantly, for not being able to explain the wrong of punishing the innocent. Moralistic and legalistic retributivism succumb to the same criticism, for they are equally unable to explain the wrong of justifying the punishment of legal and moral innocents, respectively. The above objection seems to deny the importance for a theory of being able to explain why punishment of the innocent is wrong. Second, it is not merely from a consequentialist perspective but also a retributivist perspective that justifying the intentional punishment of particular, identifiable innocents is problematic. From the viewpoint of moralistic retributivism, legalistic retributivism’s punishment of moral innocents is a serious failing. And from the perspective of legalistic retributivism, moralistic retributivism’s punishment of legal innocents is a serious failing.

But let us assume arguendo that the objection is persuasive. If so, then

---

231 I am grateful to Ken Ehrenberg for posing this possible objection.
232 See supra notes 158–159 and accompanying text.
consequentialism may equally avail itself of the argument contained in the objection, for if it is good for moralistic and legalistic retributivism, then it also is good for consequentialist theories. As a result, against the retributivist criticism that consequentialist theories justify punishment of the innocent and are unable to explain the wrong of punishing the innocent, consequentialists now have a persuasive reply. Consequentialist theories, justifying punishment on the basis of the good consequences it generates, are not designed to prevent punishment of the innocent and may be properly indifferent to punishment of the innocent. Consequentialism's justification of punishment of the innocent and its inability to explain why doing so is wrong is not a failing from its own perspective. It is only problematic from a retributivist perspective. And as such, it should not be counted as a serious failing for consequentialism.

Moreover, if for determinations of whether a theory justifies punishment of the innocent, innocence is to be defined from each theory's own perspective, then consequentialism could also define innocence in such a way that it never justifies punishment of the innocent. A consequentialist might define innocence functionally: a defendant is "innocent" if his punishment would fail to generate good consequences sufficient to outweigh its costs. Under this conception, consequentialism would never justify punishment of the innocent. Of course, this is an odd and unfortunate conception of innocence. But if we allow moralistic and legalistic retributivism to define innocence from their own respective perspectives, then we must allow consequentialism to do so as well. As a result, if innocence may be defined from the perspective of each theory, neither moralistic retributivism nor legalistic retributivism nor consequentialism justifies intentional punishment of the innocent.

Regardless of whether the objection is persuasive or not, moralistic and legalistic retributivism fare no better than consequentialist theories. If the objection is unpersuasive, then moralistic and legalistic retributivism justify the intentional punishment of particular, identifiable innocents and are unable to explain why doing so is wrong. If the objection is persuasive, then neither these retributivist theories nor consequentialism justify the intentional punishment of the innocent. Either way, as to justifying the intentional punishment of the innocent, the moralistic and legalistic versions of retributivism are no better than consequentialist theories.

E. Conclusion

The most important criticism retributivists have leveled against consequentialist theories of punishment is that they justify the intentional punishment of particular, identifiable innocents. But this criticism may be
turned back against some forms of retributivism. Moralistic retributivism justifies the intentional punishment of particular, identifiable legal innocents; legalistic retributivism justifies the intentional punishment of particular, identifiable moral innocents.

One possible solution for moralistic retributivism is to require the precondition of a valid legal adjudication of guilt before punishment could be imposed. Although this would prevent legal innocents from being punished, it is a requirement extrinsic to moral retributivism itself. As such, moralistic retributivism per se is still subject to the criticism that it intentionally justifies the punishment of particular, identifiable legal innocents. If the solution was deemed to be satisfactory, however, consequentialism could equally avail itself of it. By incorporating a valid legal adjudication of guilt precondition into their theories, consequentialists also would no longer intentionally punish particular, identifiable innocents. Without the precondition, moralistic retributivism is subject to the same criticism retributivists have leveled at consequentialist theories. With the precondition, consequentialist theories are no longer subject to their most significant criticism. Either way, as to justifying the intentional punishment of particular, identifiable innocents, moralistic retributivism fares no better than consequentialist theories.

One possible objection is that neither moralistic nor legalistic retributivism justifies the punishment of innocents as they define innocent. But this critique fails to recognize that neither moralistic nor legalistic retributivism can explain why the punishment of legal and moral innocents, respectively, is wrong. If consequentialism’s failure to explain the wrong of punishing the innocent renders consequentialism untenable, then the same failure of these two versions of retributivism renders them equally untenable. On the other hand, if the objection was considered persuasive—that is if each theory is entitled to construe what constitutes punishment of the innocent from its own perspective—then consequentialism can likewise define innocent in such a way that consequentialism also does not punish the innocent. Either way, as to the intentional punishment of particular, identifiable innocents, neither moralistic nor legalistic retributivism fares better than consequentialism.

III. RETRIBUTIVISM’S INABILITY TO PROVIDE A PRINCIPLED ACCOUNT OF THE DEGREE OF DESERT

The previous Part demonstrated that two versions of retributivism—moralistic and legalistic—justify the intentional punishment of particular, identifiable legal and moral innocents, respectively. This Part will consider another version of retributivism—combined moralistic and legalistic retributivism or, simply, combined retributivism—which, in part, avoids the problem of justifying the intentional punishment of particular, identifiable
innocents. But the combined version does this, it will be shown, only by losing the capacity to provide a principled account of the degree of an offender’s desert and punishment. This is a significant shortcoming of the combined version of retributivism because it is retributivists who have criticized consequentialist and mixed theories for their failure to supply a principled determination of the degree of an offender’s desert and punishment.

This Part presents retributivists with the following dilemma: either retributivism justifies the intentional punishment of particular, identifiable innocents (under moralistic and legalistic retributivism) or it is unable to provide a principled, coherent account of the degree of an offender’s desert and punishment (under combined retributivism). The concern of this Part is not the familiar objection that retributivism cannot determine the precise degree of desert and amount of punishment to be imposed. Instead, a novel problem will be identified: the combined version of retributivism supplies contradictory determinations of the degree of desert and punishment to be imposed. This will arise where the degree or gravity of the moral wrong of an action or course of conduct significantly diverges from the degree or gravity of the legal wrong.

A. The Moralistic and Legalistic Combined Version of Retributivism

To avoid one of the shortcomings of both the moralistic and legalistic versions of retributivism, C.W.K. Mundle was perhaps the first to suggest a form of retributivism combining both the moralistic and legalistic versions. Under the combined version, morally culpable wrongdoing and legal guilt are independently necessary and jointly sufficient to justify the imposition of punishment. Thus, the combined version only justifies the punishment of morally culpable wrongdoers who are legally guilty, and would not punish moral or legal innocents. However, it is not entirely without difficulties.

Though it avoids the problem of overinclusiveness (in punishing the

234 The combined moralistic and legalistic version of retributivism does not completely avoid the problem because it would still justify strict liability, see infra section IV.B.1, which a prominent retributivist has declared amounts to justifying intentional punishment of the innocent, see infra text accompanying note 335.

235 See infra notes 250–270 and accompanying text.

236 C.W.K. Mundle, Punishment and Desert, in THE PHILOSOPHY OF PUNISHMENT, supra note 11, at 65, 75–79. “Punishment of a person by the State is morally justifiable, if and only if he has done something which is both a legal and moral offence . . . .” Id. at 79.

237 For other accounts of retributivism as the combined moralistic and legalistic version, see, e.g., DUFF, supra note 1, at 153–154 (suggesting that punishment is only applicable for the commission of a criminal offense the violation of which constitutes moral wrongdoing); PRIMORATZ, supra note 1, at 13 (“When retributivists claim that the moral justification of punishment is in the offense committed, by ‘offense’ they mean only a violation of a morally legitimate criminal law.”).

238 For an interpretation of retributivism that assumes the combined version, see PRIMORATZ, supra note 1, at 95–96 (discussing, and defending retributivism from, criticism that it justifies punishment of violators of any positive law, no matter how unjust).
innocent) that both the moralistic and legalistic versions possess, the combined form shares the problem of underinclusiveness with moralistic and legalistic retributivism. From the perspective of moralistic retributivism, the combined version is underinclusive in failing to justify the punishment of morally culpable wrongdoing that does not satisfy the definition of a criminal offense. From the perspective of legalistic retributivism, the combined version is underinclusive in failing to justify the punishment of those who are legally guilty of a criminal offense that does not involve morally culpable wrongdoing. Thus, the combined version of retributivism avoids the problems of overinclusiveness in both moralistic and legalistic versions, but compounds the problems of underinclusiveness. By avoiding the justifications of the intentional punishment of moral and legal innocents, however, the combined version would seem preferable to either of the other two versions. But the combined version is beset by a more serious problem than underinclusiveness—the inability to supply a principled determination of an offender’s degree of desert and punishment. Before demonstrating that combined retributivism generates contradictory desert determinations, let us briefly canvass the debate as to the proper role of the degree of desert in the theory of punishment.

B. The Retributivist Claim That Only Retributivism Can Provide a Principled Determination of the Degree of Desert

Retributivists criticize consequential and mixed theories for justifying the degree or extent of punishment based not on the gravity of the offender’s wrongdoing but, rather, on the amount of deterrence that is needed or by what is necessary for an effective rehabilitation.\(^{239}\) Bentham, for example, propounds an elaborate set of rules for determining the degree of punishment that includes, but fails to treat as dispositive, the offender’s desert or culpability. The lower limit of punishment for a particular offense is based on “what is sufficient to outweigh that of the profit of the offence,”\(^{240}\) the relative degree of temptation to commit the offense,\(^{241}\) and the relative degree of “mischief” or harm the offense entails.\(^{242}\) The upper limit of the degree of punishment, because punishment itself is an evil, is no more than what is “necessary” to deter crime.\(^{243}\)

Retributivists claim their approach avoids this failing because they

\(^{239}\) Id. at 37–38; Benn, supra note 45, at 32; Leo Katz, Criminal Law, in A COMPANION TO THE PHILOSOPHY OF LAW AND LEGAL THEORY, supra note 14, at 80, 80; C.L. Ten, supra note 29, at 367–68.

\(^{240}\) BENTHAM, THE PRINCIPLES, supra note 57, at 179 (Ch. XIV) (emphasis omitted) (footnotes omitted).

\(^{241}\) Id. at 180 (the greater the degree of temptation to commit the offense, the greater the degree of punishment is required).

\(^{242}\) Id. at 181 (emphasis omitted).

\(^{243}\) Id. at 182 (emphasis omitted). In addition to these three rules, Bentham formulates ten other rules regarding the proper proportion between punishments and offenses. See id. at 181–88.
Deterring Retributivism

base the degree of punishment solely on the desert of the wrongdoer. Minor wrongdoing warrants only minor punishment, grave wrongdoing warrants grave punishment. But for consequentialist theories, the degree of punishment depends not on the offender’s desert but instead on how best to effectuate the good consequences to be promoted. For example, consider the offense of jaywalking, which is ordinarily a minor offense. Suppose there is an epidemic of jaywalking that the threat of minor punishment has not sufficiently detered. In order to successfully reduce the incidence of jaywalking, a deterrence-based theory of punishment might justify a disproportionally harsh penalty. Now consider the grave offense of murder, which ordinarily carries a substantial punishment. But suppose that in a society there has not been a single murder in fifty years. Because in such a society there is little need to deter its citizens from committing murder, a deterrence theory of punishment might only justify a comparatively minor punishment. Thus, under extreme circumstances, retributivists argue, a deterrence theory of punishment might justify the same or greater punishment of a jaywalker than of a murderer.

Consequentialists have offered a number of responses. First, consequentialism does include a limited proportionality requirement. As Justice Scalia notes, however, proportionality is essentially a retributive concept. Second, the retributivist argument for the proportionality principle is more assertion than argument. Though retributivism’s rationale

244 Hegel, supra note 22, at 127–29 (§ 101); Kant, supra note 22, at 141; Benn, supra note 45, at 32 (retributivism “insists that the punishment must fit the crime”); Greenawalt, supra note 73, at 347–48 (under retributivism, “the severity of punishment should be proportional to the degree of wrongdoing”).

245 For the view that the crime prevention rationale properly justifies disproportionate punishments, see generally Alexander, supra note 36; Quinn, supra note 36.

246 See Primoratz, supra note 1, at 37–38.

247 The consequentialist sense of a proportionality requirement, however, allows significantly greater latitude. The minimum amount of punishment is determined by a number of factors. First, it must be sufficient to outweigh the offender’s incentive to commit the offense. See supra notes 240–241 and accompanying text. Second, it must not “depreciate the seriousness” of the offense. American Law Institute, Model Penal Code and Commentaries § 7.01 1(c) (1985) [hereinafter MPC] (providing for incarceration when “a lesser sentence will depreciate the seriousness of the defendant’s crime”). Third, it must not “imply a license to commit” the offense. Herbert Wechsler, Sentencing, Corrections, and The Model Penal Code, 109 U. Pa. L. Rev. 465, 468 (1962) (the minimum amount of punishment for an offense must be such “that it does not depreciate the gravity of the offense, whatever that may be, and thus imply a license to commit it”). The upper limit is set by the consideration that the degree of punishment be no more than what is necessary to deter crime. See supra note 243 and accompanying text.

248 Harmelin v. Michigan, 501 U.S. 957, 989 (1991) (Scalia, J.) (“[I]t becomes difficult even to speak intelligently of ‘proportionality,’ once deterrence and rehabilitation are given significant weight. Proportionality is inherently a retributive concept . . . .”).

249 Kaplow & Shavell, supra note 28, at 1235:

[The] retributive theory’s rationale for the degree of punishment . . . is difficult to identify. It seems that most retributivists believe that the proportionality principle is self-evident or follows directly from the understanding that the purpose of punishment is to restore moral balance in the universe; consequently, the principle is not seen to require explicit justification.
for the proportionality principle is undoubtedly thin (and perhaps nonexistent), that punishment should be in some way proportional to the crime is an intuition (like the wrong of punishing the innocent) that is so widely shared as to make its attack unpersuasive.

The third, and most successful, consequentialist response is that retributivism has no precise calculus for determining what a wrongdoer's desert is, and thus how much punishment is deserved. However flawed, consequentialist theories at least have a comparatively more precise method for determining how much punishment is justified. For retributivism, the idea is that the punishment should fit the crime, it should be proportional to the crime. But this gives insufficient guidance as to how much punishment a specific crime warrants. For example, is twenty years imprisonment the deserved punishment for armed robbery, or is twenty-five years imprisonment? Retributivism has no basis to answer this question.

The lex talionis is inadequate. As expressed in the Bible, "You shall give your life for life, eye for eye, tooth for tooth, hand for hand, foot for foot, burn for burn, wound for wound, stripe for stripe." Blackstone rejected the lex talionis as insufficient to guide the amount of punishment for all crimes: "[T]here are very many crimes, that will in no shape admit of these penalties, without manifest absurdity and wickedness. Theft cannot be punished by theft, defamation by defamation, forgery by forgery, adul-

Id. (citing Jeffrie Murphy, Does Kant Have a Theory of Punishment?, 87 COLUM. L. REV. 509, 532 (1987) ("And why should we aim at proportionality so characterized? Kant does not attempt to argue for this view, but seems to think that the claim is self-evident or intuitively obvious.").

Hart, supra note 11, at 233 ("To many the most perplexing feature of the [retributive] model is its requirement that the punishment should in some way 'match' the crime."); Primoratz, supra note 1, at 85 ("Of all arguments against the retributive theory, probably the most popular is the one directed at the demand that punishment should fit the offense committed."); Bedau, supra note 83, at 611-14; Benn, supra note 45, at 32; Cragg, supra note 52, at 707-08; Dolinko, supra note 2, at 1636 ("[I]t has long been a stock objection to retributivism that there is simply no workable way to determine just what punishment a criminal deserves.") (emphasis omitted).

See supra notes 239-243, 247 and accompanying text. For example, take the society with the jaywalking epidemic. The degree of punishment can be increased until the optimal level of deterrence is reached. This optimum level would be determined by maximizing the margin by which the benefit of the punishment exceeds its costs. Of course, whatever precision may result from such a determination depends on accurate empirical assessments of the benefits and costs of imposing the punishment.

Cragg, supra note 52, at 702; Dolinko, supra note 2, at 1636.

Katz, supra note 239, at 80-81.

The most important challenge to retributivism has been its alleged vagueness: everyone may agree that five years in prison is unjustly harsh desert for shoplifting, or that a five dollar fine is unjustly lenient desert for rape, but beyond such clear cases our intuitions seem to fail us. Is two years, five years, or ten years the proper sanction for a rape? ... Our sense of just deserts here seems to desert us.

Id.

Contra id. at 81 (citing empirical evidence of psychophysics studies which suggest "that underlying our punishment practices is a fairly refined sense of just deserts").

Cragg, supra note 52, at 707-08; C.L. Ten, supra note 29, at 370.

tery by adultery, and the like. The problem is even more acute when translating the gravity of offenses to numerical terms of imprisonment.

Perhaps, at best, retributivism can determine the roughly appropriate punishment by comparatively ranking offenses in such a way as murder warrants greater punishment than rape, which warrants greater punishment than armed robbery, and so on. But it cannot determine whether rape warrants twenty, thirty, forty years imprisonment. Though retributivism cannot set cardinal or absolute levels of punishment, its advocates insist that they can set ordinal, or relative, levels of punishment (for example, murder warrants greater punishment than larceny). But retributivism cannot even satisfactorily determine degrees of punishment ordinally. For example, even if we assume that, all other things being equal, murder warrants greater punishment than armed robbery, does negligent homicide warrant greater punishment than intentional rape or intentional armed robbery? Retributivism has no answer. This is the serious flaw in Kant's lex talionis and even G.W.F. Hegel's more sophisticated version. Both fail to take into account differing culpability levels stemming from the various levels of mens rea or mental states that accompany the commission of wrongdoing. Retributivism has no answer to the issue of whether greater wrongdoing done with lower culpability (for example, negligence or recklessness) warrants more or less punishment than comparatively minor wrongdoing with a greater level of culpability (such as intention or purpose). Thus, retributivism can determine neither the ordinal nor the cardinal ranking of crimes and their concomitant degrees of punishment.

Punishment theorists contest the significance of retributivism's inability to impose punishments precisely reflecting offenders' just deserts. One commentator argues that retributivism's inability to impose punishments that are precisely what wrongdoers deserve constitutes an immoral use of the wrongdoers as mere means. Even Hegel himself, along with Kant the most influential retributivist, concedes retributivism's inability to determine the just punishment for any particular offense:

It is impossible to determine by reason, or to decide by applying a determination derived from the concept, whether the just penalty for an offence is corpo-
nal punishment of forty lashes or thirty-nine, a fine of five dollars as distinct from four dollars and twenty-three groschen or less, or imprisonment for a year or for 364 days or less, or for a year and one, two, or three days.\textsuperscript{265}

And quite significantly, Hegel goes on to explain that even a slightly incorrect punishment constitutes an injustice: "And yet an injustice is done if there is even one lash too many, or one dollar or groschen, one week or one day in prison too many or too few."\textsuperscript{266} If it is virtually impossible to give an offender his precise just deserts through punishment,\textsuperscript{267} and if even a slightly incorrect punishment constitutes an injustice, then virtually every instance of retributive punishment will be unjust.\textsuperscript{268} As a result, by retributivism’s very principles, retributive punishment should not be imposed.\textsuperscript{269} Seemingly contradicting Hegel, modern retributivists respond that this imprecision in determinations of desert and the degree of punishment is not problematic.\textsuperscript{270}

\textbf{C. Combined Retributivism Supplies Contradictory Determinations of the Degree of Desert}

The difficulty of the combined version of retributivism, in some situations, is not merely imprecision, but rather internal inconsistency in determining an offender’s appropriate level of desert and punishment. The moralistic version determines the level of desert and degree of punishment based on the gravity of the morally culpable wrongdoing.\textsuperscript{271} The legalistic version determines the amount of desert and punishment based on the severity of the legal violation.\textsuperscript{272} But under the combined version, what determines the degree of punishment and desert? Perhaps in most instances the severity of the moral wrong and the legal violation will coincide. But in some instances they will surely diverge. That which is a grave moral wrong

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{265} \textcite{Hegel, supra note 22, at 245} (§ 214). The editor explains that in Hegel’s time there were 24 groschen in one dollar. \textit{Id.} at 446.
\item \textsuperscript{266} \textit{Id.} at 245.
\item \textsuperscript{267} "There is no estimate which can determine degrees of moral guilt in actual individual cases. Such a thing is wholly inconceivable. It would demand an insight into motive and temptation which it is impossible to possess for others, and all but impossible for oneself." \textcite{Bosanquet, supra note 87, at 203}.
\item \textsuperscript{268} \textcite{Ewing, supra note 57, at 40}:
\item [O]wing to the impossibility of estimating the moral guilt of the offender and the degree of punishment proportionate to it … almost all punishments will be unjust. But to do injustice seems worse than to do nothing at all. . . . Ought the State to aim at retributive justice, if the overwhelming probability is that each time it tries to inflict it it will do serious retributive injustice?
\item \textsuperscript{269} \textit{Id.} (footnote omitted).
\item \textsuperscript{270} \textcite{Ewing, supra note 57, at 39} (explaining that because of the near impossibility of punishing according to an offender’s actual desert under retributivism, “it is wrong even to attempt to apply it”).
\item \textsuperscript{271} \textcite{Kleinig, supra note 220, at 114; Primoratz, supra note 1, at 94; Wojciech Sadurski, Giving Desert Its Due 239 (1985)}.
\item \textsuperscript{272} See supra notes 215–217 and accompanying text.
\end{itemize}
\end{footnotesize}
might be a minor legal violation and that which is a grave legal violation might be a minor legal wrong. In such cases, how would the combined version determine the amount or degree of an offender's desert and punishment?

Let us first consider examples of serious legal violations that are comparatively less serious moral wrongs. While violations of federal drug laws carry quite stiff penalties, their moral wrongness is arguably less serious than the legal wrong. To take a more specific example, the penalties for offenses involving crack cocaine are substantially higher than penalties for cocaine in powder form. But the moral wrongness of possessing or selling cocaine is arguably the same regardless of the particular form of the cocaine. Or consider the "three strikes and you're out" legislation. The commission of a third (even minor) felony may carry a life sentence, but the moral wrongness of the third felony may be considerably less than a felony that would alone merit a life sentence. Under the combined version of retributivism, should the punishment of these offenses reflect their substantial legal wrongness or their comparatively less serious moral wrongness? Proponents of combined retributivism have no answer.

Let us next consider examples of minor legal violations that might constitute grave moral wrongs. While lying is a serious moral wrong, at least under Kantian ethics, perjury is a comparatively minor criminal offense. Similarly, blackmail is perceived to be a great moral wrong but

---

273 Controlled Substances Act, 21 U.S.C. § 841(b)(1)(A)-(B) (1994) (imposing a mandatory minimum sentence of ten years for sale of either 5000 grams of powder cocaine or 50 grams of crack or rock cocaine, and a five year mandatory minimum sentence for either 500 grams of powder or 5 grams of crack cocaine). This amounts to a 100:1 ratio in the severity of punishment for crack cocaine as compared to the same amount of powder cocaine.

274 See, e.g., William Spade, Jr., Beyond the 100:1 Ratio: Towards a Rational Cocaine Sentencing Policy, 38 Ariz. L. Rev. 1233, 1279–83 (1996) (describing the unfairness of the disparate sentences, as expressed by prosecutors, juries and Justice Kennedy).

275 In perhaps the most well-known example, in Rummel v. Estelle, the defendant, Rummel, defrauded a bar owner in the amount of $129.75. 445 U.S. 263 (1980). But this was Rummel's third fraud conviction and thus he qualified, under the statute, for life imprisonment without parole. Id. at 284–85.

276 See George Fletcher, The Recidivist Premium, Crim. Just. Ethics, Summer-Fall 1982, at 54, 57–59 (arguing that increased penalties for habitual offenders is not morally justified within a retributivist framework). But see Von Hirsch, supra note 6, at 84–88 (defending increased penalties for recidivists as morally justified within retributivism).

277 For Kant, lying violated a person's duty to himself:

The greatest violation of man's duty to himself regarded merely as a moral being (the humanity in his own person) is the contrary of truthfulness, lying .... By a lie a man throws away and, as it were, annihilates his dignity as a man. A man who does not himself believe what he tells another ... has even less worth than if he were a mere thing .... [Lying is] a crime of a man against his own person and a worthlessness that must make him contemptible in his own eyes.

KANT, supra note 22, at 225–26 (emphasis omitted).

278 Perjury is a third-degree (minor) felony. MPC, supra note 247, § 241.1(1).

279 Consider the following account:

The criminal offence of blackmail is universally feared. By many in this country it is "considered the foulest of crimes—far crueler than most murders because of its cold-blooded premeditation
in the hierarchy of criminal offenses it is comparatively low. Under the combined version of retributivism, should the punishment of these offenses reflect the seriousness of the moral wrong or their comparatively less serious legal wrong?

Let us consider a more abstract example. Suppose that conduct $X$ committed in jurisdiction $A$ renders an offender subject to little, or even no, punishment, and the identical conduct committed in jurisdiction $B$ qualifies the offender for severe punishment. Though the degree of punishment specified by the legislatures for the identical conduct varies from jurisdiction $A$ to jurisdiction $B$, presumably the degree of moral wrongdoing is the same in both jurisdictions. For the purposes of our argument it will not matter what that degree of moral wrongdoing may be. That different people will have differing views on the degree of moral wrongness does not mean that a single act can be both a minor moral wrong and a grave moral wrong. The important point is that though the offender’s legal desert will vary from jurisdiction to jurisdiction her moral desert will not. As a result, in (at least) one of the jurisdictions there will necessarily be a substantial disparity between the offender’s moral and legal desert. In such a jurisdiction, how would combined retributivism determine the degree of the offender’s desert and just punishment? The legalistic retributivism component would give one answer and the moralistic retributivism component would supply a sub-

and repeated torture of the victim; incomparably more offensive to the public conscience than the vast majority of other offences which the law seeks to punish.”


280 Blackmail, under the rubric of criminal coercion, is either a misdemeanor or a third-degree felony. MPC, supra note 247, § 212.5(2).

281 This is not merely of hypothetical interest. George Fletcher explains that common law jurisdictions (as well as France) have a dramatically different view of complicity than some civil law jurisdictions, notably Germany. Common law jurisdictions hold to the “equivalence” theory of complicity in which all participants or accessories should be punished to the same degree as the principal perpetrators. See, e.g., MPC, supra note 247, § 2.06. In contrast, in Germany an accessory is only punished to the extent of her participation or contribution. An accessory who only makes a minor contribution to the commission of the offense would receive comparatively light punishment and an accessory that made a major contribution would receive comparatively greater punishment. For a discussion of these differing approaches to complicity see FLETCHER, supra note 46, at 188–200; FLETCHER, supra note 1, at 649–57.

So let us suppose that our offender is a minor accessory in a murder. In America, the offender’s conduct would constitute a grave legal wrong and she might receive a severe punishment, perhaps life imprisonment. In Germany, the very same conduct by our offender would be a substantially lesser legal wrong and she might receive an appreciably lighter sentence. Though there may be disagreement as to the extent of the moral wrong, regardless of the degree of moral wrong, presumably it is the same degree of moral wrong whether it is committed in Germany or in the United States. Thus, in at least in one of the jurisdictions, the two components of combined retributivism will supply substantially different answers as to the degree of the offender’s desert and just punishment. Because of the incompatibility of the two components, combined retributivism could not arrive at a principled determination of our offender’s desert. The two components of combined retributivism would supply contradictory answers.
Deterring Retributivism

stantially different answer.

In all of these examples where there is a disparity between the gravity of the moral and the legal wrong, the combined version of retributivism could come to no principled resolution of the amount of desert of the offender and the degree of just punishment. I take it that this criticism of the combined version of retributivism—that in certain instances it cannot arrive at a principled determination of the amount of desert and punishment of the offender—is fair. After all, this is the same criticism retributivists have leveled against consequentialism\(^{282}\) and the mixed theory of punishment (deterrence justifying punishment with negative retributivism as a side-constraint).\(^{283}\)

If consequentialist and mixed theories are considered untenable by retributivists, then the combined legalistic and moralistic version of retributivism is equally untenable. Combined retributivism, consequentialism, and the mixed theory all fail to determine the degree of desert and punishment in a principled manner. Since the same failing besets all three, and retributivists deem that failing to be sufficient to render consequential and mixed theories unsatisfactory, then retributivists must agree that the combined version of retributivism is also unsatisfactory for the same reason. Retributivists can, of course, avoid this criticism by shifting to either the legalistic or moralistic versions of retributivism. Either version is able to give a noncontradictory account of the degree of an offender’s desert and punishment. But shifting to either of these versions of retributivism retriggers the problem of retributivism justifying the intentional punishment of particular, identifiable innocents.

The combined version of retributivism is the only version that affords partial escape from the criticism that retributivism justifies the intentional punishment of particular, identifiable innocents. If combined retributivism is untenable, the problem of retributivism justifying the intentional punishment of particular, identifiable innocents is reintroduced.

D. Conclusion

Retributivists have criticized consequentialist and mixed theories for being unable to give a principled account of an offender’s degree of desert and punishment. But this criticism may be turned back on retributivism. While the combined (moralistic and legalistic) version of retributivism partially avoids justifying the intentional punishment of particular, identifiable moral and legal innocents, it does so only at the expense of being unable to give a principled account of an offender’s desert. In some instances, the moralistic and legalistic components of the combined version diverge and

\(^{282}\) Retributivists have criticized consequentialism for failing to adhere to any proportionality between the gravity of the offense or wrongdoing and the degree of punishment. See supra notes 239, 245-246 and accompanying text.

\(^{283}\) See supra notes 135-136 and accompanying text.
give contradictory assessments of an offender’s degree of desert and punishment. Under either the combined version, which partially avoids justifying the intentional punishment of particular, identifiable innocents but cannot give a principled account of the degree of desert and punishment, or either of the moralistic and legalistic versions, which do give a principled account of the degree of desert and punishment but cannot avoid intentional punishment of the innocent, retributivism is subject to the very criticisms retributivists have directed against consequentialist theories.

The resulting dilemma for retributivism is that either (i) it intentionally justifies punishment of particular, identifiable innocents (under legalistic or moralistic retributivism) or, (ii) it avoids the above criticism but supplies a contradictory account of the degree of an offender’s desert and punishment (under the combined version of retributivism). Retributivism, therefore, can avoid either one of the criticisms, but not both. Furthermore, retributivism’s susceptibility to either one of the criticisms is sufficient to render it untenable. After all, it is retributivists themselves who leveled these criticisms at consequential and mixed theories and, as a result, found both defective. Since the same criticisms apply to retributivism, then retributivism is equally unsatisfactory.

IV. RETRIBUTIVISM’S INDIRECT PUNISHMENT OF THE INNOCENT

In Part II, the principal retributivist criticism against consequential theories of punishment was turned back on to retributivism. Moralistic and legalistic retributivism justify the intentional punishment of particular, identifiable legal and moral innocents, respectively. Combined retributivism does not justify the punishment of innocents, but as seen in Part III, is subject to other difficulties. This Part will demonstrate that all versions of retributivism justify indirect punishment of the innocent.

Punishment of the innocent may take on a variety of forms. Perhaps the most serious form, at least from a retributivist perspective, is the intentional punishment of particular, identifiable innocents. But other varieties include punishing offenders who lack fault, culpability, blameworthiness, or mens rea, and punishing offenders with valid excuses. Punishing such offenders is often expressed as punishment under strict liability. Another form of punishing the innocent may be the erosion or elimination of procedural safeguards designed to prevent conviction and punishment of the innocent. Retributivists maintain that consequentialism does, and retributivism does not, justify punishment under strict liability.

---

284 See supra notes 138–139.
285 Combined retributivism, in some instances cannot arrive at a principled determination of the degree of offenders’ desert and punishment because its two components are in contradiction. See supra subpart III.C.
286 See infra notes 303–304 and accompanying text.
287 See infra note 325 and accompanying text.
under eroded procedural safeguards\textsuperscript{288} and without due regard to offenders' fault,\textsuperscript{289} culpability,\textsuperscript{290} mens rea\textsuperscript{291} and excuses.\textsuperscript{292} One leading retributivist argues that consequentialism's justification of punishment under strict liability is not merely tantamount to punishing the innocent but constitutes \textit{intentionally} punishing the innocent.\textsuperscript{293}

These retributivist criticisms of consequentialism will be turned back on to retributivism by showing that retributivism is compatible with justifying punishment under (i) strict liability, (ii) a presumption of guilt standard of proof, (iii) guilty verdicts by a non-majority of jurors, and (iv) an extreme form of absolute liability. While punishment under these conditions, practices, or standards may or may not constitute intentional punishment of the innocent, it does dilute or weaken the concepts of guilt and innocence to such a degree as to undermine our sense that offenders being punished truly "deserve" their punishment.\textsuperscript{294} In addition, the number of innocent defendants punished greatly increases. By punishing under the first three practices above, many, if not virtually all, innocent defendants would be convicted and punished. And under the fourth practice—an extreme form of absolute liability—\textit{all} innocent defendants would be convicted and punished.

\textbf{A. The Debate over Mistaken Punishment of the Innocent and the Doctrine of Double Effect}

To see how retributivism justifies adoption of, and punishment under, standards and practices in which most or all innocent defendants would be convicted and punished, we first need to briefly examine the debate between retributivists and consequentialists over the significance of mistaken punishment of the innocent. It is retributivists' reliance on the doctrine of double effect,\textsuperscript{295} in this debate, that renders retributivism subject to the charge of justifying the punishment of most or all innocent defendants.

\textit{1. The Consequentialist Charge: Retributivism Is Unjust by Mistak-
only Punishing the Innocent.—In response to retributivists' criticisms that consequentialism justifies punishment of the innocent, consequentialists argue that any actual punishment institution, including retributivist systems of punishment, will inevitably result in the punishment of the innocent.296 Retributivism justifies the knowing or intentional implementation and utilization of necessarily fallible legal institutions of punishment in which it is foreseeable that some unknown innocents will inevitably be mistakenly punished.297 By justifying such fallible institutions, retributivism is knowingly punishing the innocent.298 In addition, "retributivists who advocate punishment are relevantly like utilitarians who will sacrifice the welfare of innocents for the greater good, since retributivists are willing to trade the welfare of the innocents who are punished by mistake for the greater good of the punishment of the guilty."299 Under this critique, retributivism faces the dilemma of either giving up all legal institutions of punishment300 or conceding that a retributivism system of punishment is unjust by its own criteria.301

2. The Retributivist Response: Invocation of the Doctrine of Double Effect.—Retributivists readily acknowledge that innocents will mistakenly be punished, but dismiss it as an insignificant problem.302 Retributivists claim that though they know that some unknown innocents will be mistakenly punished, no particular, identifiable innocent is intentionally punished under retributivism.303 In contrast, consequentialist theories of punishment

296 Jeffrey Reiman & Ernest Van Den Haag, On the Common Saying That It Is Better That Ten Guilty Persons Escape Than That One Innocent Suffer: Pro and Con, 7 SOC. PHIL. & POL. 226, 240–41 (1990) (“The conviction of some innocents is an unintended, yet unavoidable, byproduct of bringing the guilty to justice. We are prepared to risk punishing some innocents for the sake of punishing the guilty.”); George Schedler, Can Retributivists Support Legal Punishment?, 63 MONIST 185, 187 (1980) (“We know all too well that a system of legal punishment will on occasion convict and punish by mistake individuals who are innocent . . . .”).


298 Schedler, supra note 296, at 187 (“[R]etributivists who would support systems of legal punishment are subject to the same moral criticism that they lodge against utilitarians [who would intentionally punish the innocent] . . . such people support systems of punishment that they know will condemn innocent people.”).

299 Id. at 189.

300 Id. (“We could simply refuse to lend support to any system which we know will cause innocent people to suffer.”); id. at 189–90 (“[W]e have open to us the alternative of refusing to establish systems that cause innocent people to suffer.”).

301 Id. at 196 (“[A] consistent and coherent retributivist must condemn all systems of legal punishment, and . . . retributive justice is not a standard that can be met by legal systems of punishment.”).

302 See, e.g., Moore, supra note 73, at 158.

303 For example, Michael Moore makes this point:

The probable punishment of the innocent by any real-world punishment scheme is not much of a worry. . . . We rightly set up many social institutions where we know that some percentage of in-
do not merely lead to the mistaken punishment of the innocent, but also justify the intentional punishment of particular, identifiable persons known or believed to be innocent. As R.A. Duff declares, retributivists, or nonconsequentialists:

[A]re not yet committed to the kind of perversion of punishment which a consequentialist must allegedly sanction, so long as we do not aim to procure the punishment of the innocent. There is, for a non-consequentialist, a crucial difference between the demand that we should take every possible precaution against the risk of punishing those who are in fact innocent, and the demand that we must never aim to procure the punishment of those whom we believe to be innocent. . . . [W]e thus consent to the expected punishment of some who are in fact innocent; but their punishment is, in terms of our and the system's aims, mistaken—it occurs despite our intention and our efforts to avoid it. But a system of punishment can, and properly should, satisfy the latter demand; and it is that demand which a consequentialist is allegedly willing to flout, by intending to procure the punishment of those whom she believes to be innocent.  

Implicit in the retributivist argument is that the unintended but foreseen punishment of some unknown innocents is distinguishable, in a morally significant sense, from the intentional punishment of particular, identifiable innocents. The former, attributable to retributivism, is permissible, but the latter, attributable only to consequentialism, is impermissible. This is essentially an appeal to some form of St. Thomas Aquinas's doctrine of double effect, as Duff acknowledges. Simply put, the doctrine or principle is that an act which has two effects or consequences, one bad and one good, is impermissible if the bad consequence is intended; but the act is permissible if the good consequence is intended and the bad consequence is neither intended as an end nor as a means to one's end, even if the bad con-

---

901
sequence is known or foreseen. Thus, applying the doctrine to retributivism’s punishment of the innocent, justifying institutions of punishment known to be inevitably fallible will have (at least) two effects—one bad and one good. The good effect, which is intended, is punishment of the guilty; the bad effect, which is known or foreseen but unintended, is mistaken punishment of the innocent. Since to retributivists the bad effect is neither intended as the end of punishing the guilty nor the means to that end, justifying institutions of punishment known or foreseen to punish some unidentifiable innocents may be permissible.

Let us term retributivism’s application of some variant or version of the doctrine of double effect to the issue of mistaken punishment of the innocent “DDE.” That is, let DDE be understood as the claim that a retributivist system of punishment knowingly justifying the punishment of some unknown innocents is permissible as long as it does not justify the intentional punishment of particular, identifiable innocents. This is the most prevalent and significant of the retributivist replies to the consequentialist claim that retributivism’s knowing punishment of some unknown innocents forces retributivists to either abandon actual punishment or concede that a retributivist system of punishment is unjust. As Duff admits, a retributivist must invoke some variant of the doctrine of double effect “so that the [retributivist] system can therefore be justified as being just.”

It also forms the basis of the most significant criticism of consequentialist theories of punishment—that consequentialism justifies the intentional punishment of

---

308 For a more complete exposition of the doctrine, see BLACKBURN, supra note 23, at 109 (“[A]n action is permissible if (i) the action is not wrong in itself, (ii) the bad consequence is not that which is intended, (iii) the good is not itself a result of the bad consequence, and (iv) the two consequences are commensurate.”); Philip E. Devine, Principle of Double Effect, in THE CAMBRIDGE DICTIONARY OF PHILOSOPHY, supra note 56, at 737, 738 (“One may produce a forbidden effect, provided (1) one’s action also had a good effect, (2) one did not seek the bad effect as an end or as a means, (3) one did not produce the good effect through the bad effect, and (4) the good effect was important to outweigh the bad one.”).

The fourth clause or condition, in each of the above formulations appears to involve consequentialist reasoning. Some have argued that the doctrine of double effect collapses into consequentialism:

Underlying the doctrine of double effect, and the only coherent basis for distinction adverted to by the doctrine, is nothing more than the consequentialist view that it is permissible to do that which is “merely foreseen” if the adverse consequences of the act are outweighed by the “intended” good consequences.

Bagaric & Amarasekara, supra note 12, at 146–47 (footnote omitted). While Duff acknowledges that the doctrine utilizes some aspect of consequentialist reasoning, he denies that it constitutes a collapse into consequentialism per se. Duff, supra note 2, at 424 (footnote omitted) (“Whilst the doctrine does serve to allow deontologists to attend to consequences in their practical reasoning, this does not make their views consequentialist, since consequentialists hold that only consequences ever matter.”).

Warren Quinn, the doctrine’s “most sophisticated and resourceful recent exponent,” Dolinko, supra note 2, at 1634, argues that the “most important and plausible” version of the doctrine is found in its first three conditions and ignores the more controversial fourth condition. Warren Quinn, Actions, Intentions, and Consequences: The Doctrine of Double Effect, 18 PHIL. & PUB. AFF. 334, 334 n.3 (1989). For a critique of Quinn’s version of the doctrine, see Dolinko, supra note 2, at 1634 n.42.

309 Duff, supra note 2, at 423.
Deterring Retributivism

particular, identifiable innocents.

Consequentialists and critics of retributivism have sought to resist DDE by criticizing the significance of the distinction between foreseen and intended consequences in general, and in particular, the significance of the distinction between the foreseen punishment of some unknown innocents and the intentional punishment of particular, identifiable innocents. But the debate as to whether a retributivist system is proper in justifying the foreseen punishment of some unknown innocents has stalled; it has degenerated into an argument over the validity of the doctrine of double effect. Though the doctrine of double effect is "controversial," and faces "fierce attacks," rather than fight it let us accept the DDE, for the sake of argument, and demonstrate that retributivism's reliance on the doctrine entails its acceptance of practices that retributivists have criticized consequentialists for justifying.

With respect to two versions of retributivism (moralistic and legalistic retributivism), DDE is inapplicable. For DDE to apply, the bad effect—punishment of the innocent—must be unintentional. But moralistic and legalistic retributivism do justify the intentional punishment of particular, identifiable innocents, as we have already demonstrated in Part II. Thus for moralistic and legalistic retributivism, the conditions for DDE's application are not satisfied. As a result, without recourse to DDE, moralistic and legalistic retributivism, for their mistaken punishment of the innocent, are subject to the charge that they are unjust by retributivism's own criteria.

Nonetheless, let us assume, arguendo, that DDE is applicable in order to demonstrate that all versions of retributivism punish the innocent in a variety of ways. That is, if it is permissible for retributivism to justify the knowing punishment of some foreseen, but unidentifiable, innocents, then it may be shown that retributivism justifies punishment under standards, practices and doctrines which fail to give due regard to offenders' fault, culpability, mens rea, excuses, and innocence. As a result, retributivism will be shown to justify punishment of virtually all innocent defendants.


For a more sophisticated exposition of the doctrine designed to remedy the doctrine's persistently pointed out shortcomings, see Quinn, supra note 308.

311 For criticisms of the doctrine of double effect in general, as well as its application to retributivism's mistaken punishment of the innocent, see Bagaric & Amrasekara, supra note 12, at 144–47; Dolinko, supra note 2, at 1632–35; Lempert, supra note 297, at 1184; Schedler, supra note 296, at 187–90. For a particularly detailed and comprehensive critique, see Kaplow & Shavell, supra note 28, at 1273–77 & n.757.

For a nonconsequentialist's defense of the doctrine of double effect and its application to retributivism's mistaken punishment of the innocent, see Duff, supra note 2, at 423–24.

312 BLACKBURN, supra note 23, at 109.

313 Dolinko, supra note 2, at 1634.
B. Retributivism’s Punishment of the Innocent by Substantive Means

In this subpart, retributivism, via its reliance on DDE, will be shown to justify punishment of the innocent under substantive standards of liability in which the absence of offenders' fault and culpability is irrelevant. The two standards of liability on which we will focus are strict liability and an extreme form of absolute liability.

1. Retributivism’s Justification of Punishment Under Strict Liability.—While subject to various conceptions,\textsuperscript{314} criminal strict liability is generally understood as a type of liability pertaining to offenses for which a defendant may be convicted despite the absence of mens rea, culpability, or fault, and despite the presence of an excuse.\textsuperscript{315} Finding offenders guilty of serious offenses such as murder, rape, or robbery under a strict liability standard has generally been considered anathema in the criminal law.\textsuperscript{316} As

\textsuperscript{314} There is some confusion over the precise parameters of what constitutes strict liability. For the view that no single conception of strict liability enjoys a consensus, see Douglas N. Husak, \textit{Varieties of Strict Liability}, 8 CAN. J.L. & JURISPRUDENCE 189 (1995).

For the distinction between strict liability doctrines and strict liability offenses, see DRESSLER, supra note 222, at 125; for the distinction between strict liability and absolute liability, see Michael Davis, \textit{Strict Liability, Criminal, in THE PHILOSOPHY OF LAW: AN ENCYCLOPEDIA}, supra note 17, at 841, 844; for the distinction between strict liability and negligence, see Richard Wasserstrom, \textit{Strict Liability in the Criminal Law}, 12 STAN. L. REV. 731, 741-45 (1960); for the distinction between strict liability and vicarious liability, see DRESSLER, supra note 222, at 125 n.3 and Davis, supra, at 841.

\textsuperscript{315} DRESSLER, supra note 222, at 125 (defining strict liability simply as “conviction in the absence of mens rea”); FLETCHER, supra note 1, at 716 (“We define strict liability to mean liability imposed for an act or omission in violation of the law, without considering at trial whether the defendant may exculpate himself by proving a mistake or accident bearing on the wrongfulness of the violation.”); ROBINSON, supra note 223, at 251 (“An offense may be made one of strict liability either by a determination that no culpability is required or by a determination that a reasonable mistake is no defense.”); Davis, supra note 314, at 841 (“If conduct is faulty only insofar as it does wrong intentionally, knowingly, recklessly, or negligently, then strict liability is liability without fault.”); Wasserstrom, supra note 314.

[S]trict liability offenses might be tentatively described (although not defined) as those in which the sole question put to the jury is whether the jury believes the defendant to have committed the act proscribed by the statute. If it finds that he did the act, then it is obliged to bring in a verdict of guilty. . . . [I]t is perhaps sufficient to observe that whatever it is that the concept of mens rea is thought to designate, it is this which needs not be shown to be predicable of the defendant.

\textit{Id.} at 733–34 (footnotes omitted).

\textsuperscript{316} See MPC, supra note 247, § 205 cmt., at 282 (footnote omitted) (the Code “makes a frontal attack on . . . strict liability in the penal law”); HART, supra note 11, at 17–24, 195–209. Wasserstrom provides an account of the degree of disapproval strict liability has engendered:

The proliferation of so-called “strict liability” offenses in the criminal law has occasioned the vociferous, continued, and almost unanimous criticism of analysts and philosophers of the law. The imposition of severe criminal sanctions in the absence of any requisite mental element has been held by many to be incompatible with the basic requirements of our Anglo-American, and, indeed, any civilized jurisprudence.

Wasserstrom, supra note 314, at 731. \textit{But see} LADY BARBARA WOOTTON, CRIME AND THE CRIMINAL LAW: REFLECTIONS OF A MAGISTRATE AND SOCIAL SCIENTIST 46–57 (1963) (forcefully advocating for strict liability throughout the criminal law); Wasserstrom, supra note 314, at 734–45 (mounting a limited defense of strict liability).
far back as the thirteenth century, St. Thomas Aquinas pronounced that "a man should never be condemned without fault of his own to an inflictive punishment."317 And today, "[m]ost modern criminal law scholars do not look kindly upon the abandonment of the mens rea doctrine."318 H.L.A. Hart commented that "strict liability" is generally viewed with great odium and admitted as an exception to the general rule, with the sense that an important principle has been sacrificed."319 The Commentary to the Model Penal Code trumpets "the Code's insistence that an element of culpability is requisite for any valid criminal conviction."320 The Supreme Court has held that as to the necessity for mens rea (thereby precluding the strict liability standard), "The contention that an injury can amount to a crime only when inflicted [with mens rea] is no provincial or transient notion. It is . . . universal and persistent in mature systems of law."321

Suppose Lady Wootton's infamous suggestion322 of punishing serious crimes, such as murder, rape, or robbery, as strict liability offenses323 was incorporated into a retributivist system of punishment. Presumably, at least some of those convicted under a strict liability standard would, in fact, be morally culpable. Many offenders would also be convicted who lacked mens rea, had valid excuses, and in general were not morally culpable. But since such criteria of culpability would be irrelevant as to determining their guilt under this system, evidence of the lack of moral culpability would be inadmissible. As a result, the retributivist system would never know whether any particular, identifiable defendant was morally culpable, though

For a review of the arguments both for and against strict liability see, e.g., FLETCHER, supra note 1, at 713–36; Larry Alexander, Reconsidering the Relationship Among Voluntary Acts, Strict Liability, and Negligence in Criminal Law, 7 SOC. PHIL. & POL'Y. 84, 87–88 (1990).

For an overview of the debate over strict liability as it relates to issues of punishment see C.L. TEN, supra note 35, at 86–122.

317 AQUINAS, supra note 20, II, part 2-2, quest. 108, 4th art.
318 DRESSLER, supra note 222, at 127.
319 HART, supra note 11, at 20. "We condemn legal systems where they disregard this principle" of foregoing punishment of actors who have valid excuses. Id. at 21.
320 MPC, supra note 247, § 2.02, explanatory note at 23.
322 WOOTTON, supra note 316, at 46–57. For H.L.A. Hart's famous rejoinder, see HART, supra note 11, at 195–209.
323 The basis for Lady Wootton's promotion of strict liability throughout the criminal law was to increase the preventive effect of criminal prohibitions:

If the law says that certain things are not to be done, it is illogical to confine this prohibition to occasions on which they are done from malice aforethought; for at least the material consequences of an action, and the reasons for prohibiting it, are the same whether it is the result of malicious plotting, of negligence or of sheer accident. . . .

. . . [A]n action does not become innocuous merely because whoever performed it meant no harm. If the object of the criminal law is to prevent the occurrence of socially damaging actions, it would be absurd to turn a blind eye to those which were due to carelessness, negligence or even accident.

WOOTTON, supra note 316, at 51–52.
it would know that many unidentifiable defendants would be convicted who were not morally culpable. No particular, identifiable innocent (that is, a defendant without moral culpability) would be intentionally convicted and punished.

Would this system be compatible with retributivism? Yes, by recourse to DDE. Just as retributivism claims it does not intentionally punish the innocent despite intentionally erecting and maintaining fallible institutions of punishment in which, inevitably, innocents are mistakenly punished, so too a retributivist system punishing offenders under a strict liability standard in which the nonculpable are mistakenly punished could claim that no particular, identifiable nonculpable defendants are intentionally punished. The adoption of, and punishment under, a strict liability approach to serious criminal offenses would have two effects. The intended good effect would be the punishment of the guilty; the unintended bad effect would be the punishment of innocent or faultless defendants. Consequently, via DDE, retributivism can permit and justify the adoption of, and punishment under, a strict liability approach.

That a retributive system would justify, via DDE, punishment under strict liability is particularly striking because commentators often claim that strict liability is incompatible with retributivism. Moreover, compatibility with strict liability is usually a criticism retributivists direct at consequentialist or deterrence-based theories. Part of that criticism is that

324 Fletcher, supra note 1, at 713 (“Liability despite a good excuse is ‘strict’ in the sense that the criteria of liability deviate from the principles of just punishment.”); Morris, supra note 6, at 44 (noting that punishment under a strict liability approach is a “glaring example” of that which would not satisfy a “just system” of punishment); Alexander, supra note 316, at 88 (“Strict liability crimes . . . are surely inconsistent with any version of retributivism that absolutely condemns punishment of the nonculpable.”); see Dressler, supra note 222, at 105 (“[T]he principle of mens rea has its roots far deeper in retributive than in utilitarian soil.”); id. at 127 (“[T]he requirement of mens rea is consistent with the retributive principle that one who does not choose to cause social harm, and who is not otherwise morally to blame for its commission, does not deserve to be punished.”).

[The avoidance of strict liability is] necessitated by the need to preserve the essentially punitive [retributivist] function of the criminal law . . . If that function is conceived less in terms of [retributivist] punishment than as a mechanism of prevention these fears [of strict liability’s problems] become irrelevant . . .

If, however, the primary function of the courts is conceived as the prevention of forbidden acts, there is little cause to be disturbed . . .

Wootton, supra note 316, at 50–51. Another account of strict liability’s incompatibility with retributivism is based on strict liability imposing a duty of “super care”:

Retributivists object . . . to the justice of demanding super care. For (most) retributivists, a law imposing criminal liability for negligence is justified, if it is, only because there is a natural duty of reasonable care. “Evil mind”—a mind not up to the standard of the ordinary reasonable mind—is still part of justifying punishment for negligent crime. The retributive objection to strict criminal liability is that no such natural duty exists for super care. Since failing to exercise super care is not failing in a natural duty, crimes of strict liability must (it is said) punish for failure to take unreasonable precautions, bad luck, or some other nonfault. No one can deserve punishment for that.

Davis, supra note 314, at 843.

325 See, e.g., Benn, supra note 45, at 31 (arguing that consequentialist justification for legal excuses is “unsatisfactory”); Ezorsky, supra note 72, at xvi (criticizing consequentialist theories for their com-
Deterring Retributivism

consequentialist theories cannot convincingly explain why excused conduct should be exempt from punishment.\textsuperscript{326} Bentham attempted to account for excuses by claiming that since excused conduct could not be deterred, punishment would be "inefficacious" or "unprofitable," that is, useless.\textsuperscript{327} H.L.A. Hart famously replied that "Bentham's argument is in fact a spectacular non sequitur."\textsuperscript{328} All Bentham's argument demonstrates is that "the threat of punishment will be ineffective so far as the class of persons who suffer from these [excusatory] conditions is concerned."\textsuperscript{329} But "the actual infliction of punishment on those persons, may secure a higher measure of conformity to law on the part of normal persons than is secured by the admission of excusing conditions."\textsuperscript{330} In other words, though Bentham may be right that punishing excused actors will not generate specific deterrence,\textsuperscript{331} he overlooks the possibility that punishing excused actors will still yield general deterrence.\textsuperscript{332} Thus, Bentham's rationale for not punishing actors who have excuses under a consequentialist theory fails.\textsuperscript{333} Hart concludes that modern consequentialist theories fare no better.\textsuperscript{334}

One prominent critic of consequentialist theories, R.A. Duff, makes the case against a deterrence-based theory that might employ a strict liability standard to promote deterrence:

If for the sake of general deterrence we make liability for an offence so strict

patibility with justifying punishment under strict liability); Huigens, supra note 3, at 956 (footnote omitted) ("Consequentialism authorizes not only nonintentional criminal liability, such as negligence, but also nonfault, or strict, criminal liability."); id. at 957 (footnotes omitted) ("The elimination of fault from the criminal law has been a goal of the reformist impulse in deterrence theory."); id. at 958 ("The central failure of... deterrence theory involves the inability to develop a conception of nonintentional fault."); McCloskey, Non-Utilitarian, supra note 120, at 124 (arguing that consequentialism justifies punishment "of those not responsible for their acts"); id. at 125 ("Utilitarianism involves the conclusion that if it is useful to punish lunatics, mentally deranged people... it is obligatory to do so.").

\textsuperscript{326} The consequentialist rationale for not punishing excused conduct "is nothing but a fragile construct of logical fallacies." FLETCHER, supra note 1, at 813; cf. Alexander, supra note 316, at 88:

[Strict liability crimes do serve some retributive and consequentialist values at the same time that they disserve other retributive values. Strict liability crimes serve the retributive value of seeing that the culpable do not go unpunished in that by omitting the necessity to prove culpability beyond a reasonable doubt, they make it easier to convict the culpable.]

\textit{Id.} (footnote omitted).

\textsuperscript{327} BENTHAM, THE PRINCIPLES, supra note 57, at 173–75 (Ch. XIII).

\textsuperscript{328} HART, supra note 11, at 19.

\textsuperscript{329} \textit{Id.}

\textsuperscript{330} \textit{Id.}

\textsuperscript{331} See supra note 61 and accompanying text.

\textsuperscript{332} See supra notes 61–64 and accompanying text.

\textsuperscript{333} For another critique of Bentham's explanation of why punishing excused actors is pointless, see FLETCHER, supra note 1, at 813–17.

\textsuperscript{334} HART, supra note 11, at 20. Hart anticipates a wide variety of arguments consequentialists might make in attempting to account for excuses. \textit{Id.} at 19–21. But even if they were to be successful, Hart maintains, consequentialism still cannot explain why it would be wrong to punish actors with valid excuses. \textit{Id.} at 21.
that one who conducts a legitimate activity with all reasonable care may still be guilty in law of an offense ... we do not simply consent to the mistaken punishment of some who are actually innocent: we intend to convict and punish some whom we believe to be properly innocent, in order to encourage others or to ensure the punishment of the guilty.335

Though speaking about the possibility of a deterrence approach utilizing a strict liability standard, Duff's comments are also applicable to retributivism using the same standard. Duff concludes that the adoption of a strict liability standard amounts to intentionally punishing the innocent! Since we have shown that retributivism, via DDE, would justify punishment under strict liability, then, according to Duff's comment, retributivism's justifying punishment under strict liability might also amount to intentionally punishing the innocent.

That retributivism is compatible with a standard of culpability for serious offenses that fails to take into account any criteria of mens rea, excuse, culpability, or fault may demonstrate that either retributivism or DDE is problematic. Either way, retributivism faces a serious dilemma. Retributivism is either (i) with DDE, subject to the criticism that it justifies punishment under strict liability for serious offenses or, (ii) without DDE, unjust by retributivism's own criteria by intentionally maintaining punishment institutions that inevitably mistakenly punish the innocent.

2. Retributivism's Justification of Punishment Under an Extreme Form of Absolute Liability.—In this section it will be shown that retributivism justifies punishment under a standard of liability—an extreme form of absolute liability—in which not merely many, but every, innocent defendant would be convicted and punished. Absolute liability may perhaps be best understood as distinguished from strict liability. Whereas strict liability eliminates the necessity of mens rea (and some excuses) as a condition for liability and punishment, absolute liability eliminates both the mens rea and the actus reus, as well as the applicability of all excuses, as conditions for liability and punishment.336 That is, under absolute liability, a defendant might be convicted and punished despite the presence of an excuse, and despite the absence of mens rea, culpability, fault, or even an act.337 But not all defendants would be convicted under absolute liability. Certain defenses could still apply. The so-called policy, extrinsic, or nonexculpatory defenses might still exonerate a defendant under absolute liability.

Although retributivism, via DDE, could justify punishment under absolute liability, retributivism is also compatible with justifying punishment

---

335 DUFF, supra note 1, at 159–60 (emphasis added). For a similar view, see MORRIS, supra note 6, at 35 (punishment of persons who have valid excuses constitutes punishment of the innocent).
336 Davis, supra note 314, at 841.
337 Id. (the basis for liability may be nothing more than "a mere event, reflex, or external cause").
338 See supra notes 222–223. For examples of such defenses, see text accompanying note 224.
under an even more extreme form of liability—what I shall term “absolute liability plus.” Under this fictional absolute liability plus, as I conceive it, every defendant is automatically found guilty and punished. Suppose that even serious offenses such as homicide, rape, and robbery were punished under this form of liability. Under this form of liability, all the “actually guilty” would be convicted and punished, but so would all of the “actually innocent.”

How could retributivism possibly be compatible with justifying punishment under such a standard? By recourse to DDE. The adoption and use of “absolute liability plus” would have two effects—one good and one bad. The intended good effect would be punishment of the guilty; the bad side effect of punishment of every innocent defendant would be unintended. Though it would be known or foreseen that every innocent defendant would be convicted and punished, no particular, identifiable innocent would intentionally be convicted and punished. This is because it would not be known which defendants are actually guilty and which are innocent. Via DDE, retributivism permits and justifies the conviction and punishment of defendants under “absolute liability plus.”

The dilemma for retributivism is that with DDE, retributivism justifies the adoption of, and punishment under, an approach to liability in which every innocent defendant would be convicted and punished. But without DDE, retributivism is conceded to be unjust for justifying punishment in an inevitably fallible system in which it is known that at least a few unknown innocents will be mistakenly punished.

C. Retributivism’s Punishment of the Innocent by Procedural Means

Retributivists contend that consequentialism, if it is in the best interests of society, would justify the erosion or elimination of procedural safeguards in criminal trials designed to prevent punishment of the innocent and insure the accuracy of trial outcomes. Herbert Morris notes that a consequential, rehabilitative system of punishment would disregard a number of procedural safeguards due to a lack of concern with punishment of the innocent. Ronald Dworkin argues that “a society that submits questions

339 See supra text accompanying note 309.
340 See, e.g., McCloskey, Non-Utilitarian, supra note 120, at 124. McCloskey details the variety of ways in which the elimination of safeguards against punishment of the innocent would be justified under a deterrence-based approach to punishment:

We may sometimes best deter others by punishing, by framing, an innocent man who is generally believed to be guilty, or by adopting rough and ready trial procedures, as is done by army courts martial in the heat of battle in respect of deserters, etc.; or we may severely punish a person not responsible for his actions, as so often happens with military punishments for cowardice, and in civil cases involving sex crimes where the legal definition of insanity may fail to cover the relevant cases of insanity.

Id. (emphasis added).
341 Morris compares the level of procedural safeguards in a punishment, or retributive, system with those in a therapy, or consequential, rehabilitative system:
of criminal procedure to an ordinary utilitarian calculus does not recognize the independence or importance of moral harm, or, if it does, does not recognize that even an accidental conviction of an innocent person is an occasion of moral harm.\(^\text{342}\) As a result, Dworkin assumes that a consequentialist or "cost-efficient" society is defective.\(^\text{343}\)

But these criticisms of consequentialism may be turned back on retributivism. In this subpart retributivism, via DDE, will be shown to justify punishment under procedures of determining guilt that dilute the concept of guilt to such an extent that virtually all innocent defendants would be convicted and punished. These procedures include a presumption of guilt standard of proof and guilty verdicts by a non-majority of jurors.

1. Retributivism's Justification of Punishment Under a Presumption of Guilt Standard of Proof.—One way to punish the innocent (albeit not necessarily intentionally) is to lower the standard of proof required for conviction. The lower the standard of proof, the greater the number of defendants, both guilty and innocent, who will be convicted and punished.\(^\text{344}\) Our standard of proof in criminal trials\(^\text{345}\)—the presumption of innocence unless guilt is established beyond a reasonable doubt—serves to insure that many of the guilty, but few of the innocent, will be convicted and punished.

\[^{\text{342}}\] Dworkin sets out to determine the degree of procedural protections to which a defendant is entitled. \textit{Id.} at 72. In trying to find a middle ground, he rejects both extremes of a defendant being entitled to the most accurate trial possible (which would be very expensive) as well as a defendant not being entitled to any procedural protections. \textit{Id.} at 72–73. The latter extreme he associates with consequentialism:

\begin{quote}
If people are not entitled to the most accurate trials possible, hang the cost, then to what level of accuracy are they entitled? Must we flee to the other extreme, and hold that people accused of crime are entitled to no particular level of accuracy at all? That would be our assumption if we chose trial procedures and rules of evidence entirely on the basis of cost-benefit calculations about the best interests of society as a whole, balancing the interests of the accused against the interests of those who would gain from public savings in a greatest-good-of-the-greatest-number way. Would that cool utilitarian approach be consistent with our fervent declaration that the innocent have a right to go free?
\end{quote}

\textit{Id.} at 84.

\[^{\text{344}}\] Binder & Smith, \textit{supra} note 27, at 131 ("The lower the proof standard, the more guilty offenders will be punished, but also the more innocent suspects."); Reiman & van den Haag, \textit{supra} note 296, at 240 ("[H]igh standards of proof make sure that fewer innocents are convicted, but also that fewer criminals are, and... low standards of proof make sure that more criminals are convicted, but also more innocents."); \textit{id.} at 241 ([S]tandards of proof codetermine conviction rates... ).

\[^{\text{345}}\] The Fifth and Fourteenth Amendments' due process clauses guarantee that a defendant should not be found guilty unless the prosecutor persuades the judge or jury of her guilt "beyond a reasonable doubt of every fact necessary to constitute the crime" for which she is charged. \textit{In re Winship}, 397 U.S. 358, 364 (1970).
Given that mistakes are inevitable, it reflects a preference for false acquit-tals over false convictions, and it finds expression in William Blackstone's adage that it is better that ten guilty escape punishment than one innocent be punished.\textsuperscript{346} Given the importance retributivists place on not punishing the innocent, it is not surprising that retributivism has been said to have profoundly influenced the presumption-of-innocence standard.\textsuperscript{347} In contrast, retributivists point out, consequentialist theories justify intentional punishment of the innocent and are unable to explain the wrong of doing so. As Dworkin's critique suggests, if it would benefit society (for example, by increasing deterrence) consequentialism might well justify lowering the standard of proof required for conviction and punishment,\textsuperscript{348} thereby increasing the number of innocents mistakenly punished. And Morris maintains that a consequential, rehabilitative system would have little need for a presumption-of-innocence standard of proof.\textsuperscript{349}

Before turning back these criticisms on to retributivism by showing how retributivism justifies dramatically lowered standards of proof, let us briefly visit the debate as to whether retributivism is indeed compatible with the presumption of innocence. Ernest van den Haag argues that retributivism cannot justify the presumption-of-innocence standard\textsuperscript{350} because it is incompatible with retributivism's\textsuperscript{351} twin absolute dictates to punish the guilty and to acquit the innocent.\textsuperscript{352} Van den Haag explains the difficulty in

\textsuperscript{346}The maxim is generally, and most famously, associated with the great English jurist William Blackstone. IV WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 358 (Ch. 27) (1844 ed.) (1765) ("[T]he law holds that it is better that ten guilty persons escape, than that one innocent suffer."). For a brief history of the maxim, see Jeffrey Reiman & Ernest van den Haag, supra note 296, at 226.

\textsuperscript{347}Cragg, supra note 52, at 707.

\textsuperscript{348}For example, one consequentialist argues in favor of lower standards of proof—either a clear-and-convincing-evidence or preponderance-of-the-evidence standard—depending on the circumstances. Reiman & van den Haag, supra note 296, at 241, 248. Van den Haag (writing separately from his co-author) sets the lower limit on the standard of proof by "when the number of innocents being convicted so nearly approaches the number of criminals that punishment would no longer be perceived to be directed at [criminals]." Id. at 241.

\textsuperscript{349}MORRIS, supra note 6, at 40.

\textsuperscript{350}More precisely, van den Haag argues that retributivism cannot justify the "punishment constraint," by which he means our criminal law standard of presumed innocence unless guilt is proved beyond a reasonable doubt as expressed in the saying that it is better that ten guilty persons escape than that one innocent is punished. Reiman & van den Haag, supra note 296, at 226, 240–44.

\textsuperscript{351}Van den Haag is presumably referring to obligatory or positive retributivism. See supra notes 118–119 and accompanying text. In contrast, negative retributivism merely demands that no innocents be punished. But even the reasonable doubt standard ensures that some, even if only a few, innocents will be convicted and punished. As a result, negative retributivism would seem to be incompatible with our presumption of innocence and reasonable doubt standard. To satisfy negative retributivism, the reasonable doubt standard of guilt would have to be changed to a more stringent presumption of innocence standard—"innocence unless there is absolute certainty as to guilt." Only in this way could the dictate of negative retributivism, that no innocents be punished, be satisfied. While such a standard would satisfy negative retributivism, virtually no guilty would be punished under that standard.

\textsuperscript{352}See supra notes 73, 75, 119 and accompanying text.
determining the standard of proof justified by retributivism:

How, then, do retributivists determine which has priority: the obligation to punish the guilty, or the obligation not to punish the innocent? They don't. . . . [R]etributivism tells us that we have an obligation to punish the guilty and not to punish the innocents. We are not told which duty has priority. Should we try to convict fewer innocents and risk letting more of the guilty escape, or try to convict more of the guilty, and, unavoidably, more of the innocent? Retributivism (although not necessarily retributivists) is mute on how high standards of proof ought to be; it can neither impose nor oppose the [presumption-of-innocence standard].

That is, since retributivism does not prioritize between the two duties, retributivism cannot justify the presumption-of-innocence standard, which reflects the comparatively greater value placed on not punishing the innocent. It might be added that the presumption-of-innocence standard, which inevitably will result in some mistaken punishment of the innocent, not only does not satisfy the categorical duty to punish the guilty but also fails even to satisfy the categorical duty not to punish the innocent.

While some retributivists insist that punishment of the innocent is a greater evil than letting the guilty go unpunished, that our standard of proof should reflect that, and that such a presumption-of-innocence standard is compatible with retributivism, others maintain that both the pre-

---

353 Reiman & van den Haag, supra note 296, at 242-43.

354 For the view of a retributivist that punishing the innocent is worse than foregoing punishment of the guilty, see, for example, MORRIS, supra note 6, at 35:

The resolution arrived at in the system I am describing consists in weighing as the greater evil the punishment of the innocent. . . . [I]t is determined to be a greater evil for society to interfere unjustifiably with an individual by depriving him of good than for the society to fail to punish those that have unjustifiably interfered.

Id.

355 For the view that standards of proof should reflect the greater evil of punishing the innocent than foregoing punishment of the guilty, see, for example, id. at 36:

[The deprivation, in this just system of punishment, is linked to . . . procedures that strike some balance between not punishing the guilty and punishing the innocent . . . in which it is evident that the evil of punishing the innocent is regarded as greater than the nonpunishment of the guilty.

Id.

356 Reiman & van den Haag, supra note 296, at 230–34. Terming the saying that it is better to let ten guilty escape than punish one innocent the "punishment constraint," Reiman (writing separately from his co-author) argues that for retributivism to be plausible, punishment of the innocent must be worse than nonpunishment of the guilty:

I shall make use of the assumption that (for crimes and their punishments of comparable harm a la lex talionis) punishment of the innocent and crime against innocent are equally bad. If this is so, then, it follows that to deny the punishment constraint, retributivism must insist that nonpunishment of the guilty is as bad as crime against the innocent. But, once we consider the losses that the various failures of duty impose on those to whom they are owed, we shall see that the claim that nonpunishment of the guilty is as bad as crime against the innocent is implausible. To be plausible, then, retributivism must accept that crime against the innocent is substantially worse than nonpunishment of the guilty. But, since punishment of the innocent is as bad as crime against the innocent, it follows that, to be plausible, retributivism must grant that punishment of the innocent is substantially worse than nonpunishment of the guilty.
sumption-of-innocence standard and the preponderance-of-the-evidence standard (roughly equalizing the number of false acquittals and false convictions) are compatible with retributivism.\textsuperscript{357} Though there is no standard that satisfies both duties,\textsuperscript{358} perhaps the civil law preponderance-of-the-evidence standard constitutes a compromise that equalizes the importance of both duties.\textsuperscript{359} The compromise,\textsuperscript{360} however, entails the elimination of the presumption of innocence, which fundamentally marks the difference between civil and criminal law. Moreover, the use of a preponderance-of-the-evidence standard would substantially increase the number of innocents punished. That retributivism, which severely criticizes other theories of punishment for justifying punishment of the innocent, is compatible with the civil law standard of liability is a startling admission.

Retributivism, however, may be shown to justify punishment under a dramatically lower standard—a fictitious standard of proof in which virtually all innocent defendants would be convicted and punished. This standard, let us term it the “presumption-of-guilt” standard of proof, presumes the guilt of a defendant unless absolute evidence of innocence is proven to the satisfaction of the judge or jury. The adoption of the standard would

\textit{Id.} at 232–33.

\textsuperscript{357} \textsc{Moore, supra note 73}, at 157:

The retributivist might adopt a principle of symmetry here—the guilty going unpunished is exactly the same magnitude of evil as the innocent being punished—and design his institutions accordingly. Or the retributivist might share the common view (that the second is a greater evil than the first) and design punishment institutions so that "ten guilty persons go unpunished in order that one innocent not be punished."

\textit{Id.} at 157 n.11.

\textsuperscript{358} There is no standard that is capable of satisfying the twin dictates of positive retributivism. Positive retributivism demands (i) that not a single innocent be punished, and (ii) that not a single deserving wrongdoer escape the punishment that she deserves. To satisfy the first demand that no innocents be punished, the standard would have to be “innocence unless there is absolute certainty as to guilt.” But such a standard would allow almost all guilty defendants to be acquitted. To satisfy the second duty that all guilty defendants be punished, we might choose what I will term the “presumption of guilt standard.” Under that standard, guilt would be presumed unless evidence of absolute innocence was established. But such a standard would ensure that virtually every innocent would be convicted and punished. Therefore, neither standard, nor any single conceivable standard, would satisfy both duties or dictates of positive retributivism.

\textsuperscript{359} If utilized in a criminal trial, a defendant would be convicted if her guilt was established by a preponderance of the evidence—that is, if the evidence established that her guilt was more likely than not. The preponderance standard would serve to roughly equalize the number of convictions and acquittals. It would reflect the thought that punishing the guilty and acquitting the innocent are equally important; and that acquitting the guilty was as wrong as punishing the innocent. The preponderance-of-the-evidence standard constitutes a compromise on the standard of guilt required by both of the absolute dictates of positive retributivism—to compromise between no false guilty verdicts and no false acquittals.

\textsuperscript{360} The compromise would be satisfactory to neither negative nor positive retributivism. Negative retributivism would not be satisfied because many innocents would be found guilty and punished. The compromise standard satisfies negative retributivism even less well than our reasonable doubt standard. The compromise standard also fails to satisfy either of the twin dictates of positive retributivism. Though equalizing the number of false acquittals and false guilty verdicts, the preponderance standard would convict many innocents and acquit many of the guilty.
serve to insure that no guilty person would escape conviction and punish-
ment. It would also have the effect that virtually every innocent defendant
would be convicted and punished. 361

How could retributivism, which stresses the importance of not punish-
ing the innocent, justify such an extreme standard? By recourse to DDE. Though it would be known or foreseen that virtually all innocent defendants
would be convicted and punished, it would also be known or foreseen that
at least some of the defendants found guilty would actually be guilty. The
intended good effect in adopting the presumption-of-guilt standard would
be to punish the guilty, the bad side effect of punishment of virtually all in-
ocent defendants would be unintended. Retributivism would not be inten-
tionally punishing any particular, identifiable innocent. As a result,
retributivism, via DDE, can permit and justify the adoption and utilization
of the presumption-of-guilt standard.

This presumption-of-guilt standard does satisfy one of the categorical
demands of retributivism—that the guilty be punished. That the standard
fails to satisfy the duty that no innocent be punished hardly makes it incompat-
ible with retributivism, because no single standard can satisfy both of the
duties, 362 and standards which satisfy neither mandate are nonetheless
claimed to be compatible with retributivism. 363 That retributivism would be
compatible with a standard that leads to the punishment of virtually all in-
ocent defendants while simultaneously (i) demanding that no innocent be
punished, and (ii) criticizing consequentialist theories for justifying the pun-
ishment of the innocent is, to say the least, implausible. But retributivism,
via DDE, justifies punishment under the presumption-of-guilt standard.

Retributivism, then, is subject to the same claim that retributivists have
levered against consequentialist theories. According to Dworkin, it is con-
sequentialism that justifies the erosion or elimination of procedural safe-
guards designed to prevent innocents being from punished and to insure the
accuracy of criminal trials. 364 But as demonstrated, retributivism also
would justify the elimination of a procedural safeguard—the presumption-
of-innocence standard of proof. While retributivists claim that consequen-
tialism cannot explain why punishment of the innocent is wrong, retributiv-
ism cannot explain why it would be wrong to punish defendants after
finding them guilty under a presumption of guilt standard.

That DDE can justify implementation, in a retributivist system, of a
standard in which almost all innocent defendants are found guilty and pun-
ished may demonstrate either that retributivism justifies punishment of the
innocent on a mass scale or that DDE is untenable or proves too much. The

361 After all, how many defendants, even truly innocent ones, could demonstrate absolute evidence
of innocence?

362 See supra notes 350–351, 358–360 and accompanying text.

363 See supra notes 356–357 and accompanying text.

364 See supra notes 342–343 and accompanying text.
resulting dilemma for retributivism is that either (i) with DDE, retributivism is subject to the criticism that it justifies punishment under a standard of proof that will find guilty virtually every innocent defendant, or (ii) without DDE, retributivism is subject to the criticism that, by intentionally maintaining fallible punishment institutions which inevitably, mistakenly punish the innocent, it is unjust by its own criteria.

2. Retributivism’s Justification of Punishment By Guilty Verdicts from a Non-Majority of Jurors.—Retributivists make a general critique of consequentialism as justifying the erosion or elimination of procedural safeguards designed to protect the innocent in, and insure the accuracy of, criminal trials. One particular example on which they focus is the jury system. Dworkin suggests that consequentialism would justify weakening the size, and unanimity requirement, of juries. Morris declares that a consequential, rehabilitative approach might eliminate juries altogether. But like other attacks on consequentialism, these criticisms may also be turned back on retributivism. For Dworkin, even a minor erosion of the jury system, as justified by consequentialism, would constitute an "injustice." This section will demonstrate that retributivism, via DDE, justifies punishment of offenders found guilty in a jury system in which both the size and unanimity requirements are not merely slightly, but dramatically, eroded.

The traditional arrangement of a jury rendering a guilty verdict is that all twelve jurors must vote unanimously in favor of conviction. While this still holds in federal courts and in most states, both the size and the unanimity requirement have been relaxed in some states. The Supreme Court has held that state laws allowing juries as small as six members are constitutional, and non-unanimous guilty verdicts are permissible as long

---

365 DWORKIN, supra note 303, at 72, 90–91.
366 Morris compares the role of the jury in a system of justice, or retributivism, with a consequential, or rehabilitative, system: “[A] jury system which, within a system of justice, serves to make accommodations to the individual situation and to introduce a human element, would play no role or a minor one in a world where expertise is required in making determinations of disease and treatment.” MORRIS, supra note 6, at 40–41.
367 As Dworkin explains it:
[The number of jurors is plainly so important a consideration in guarding an accused against injustice, when a unanimous verdict is required to convict him, that any substantial change in that number for capital cases or cases threatening severe punishments—say reducing the number to six—would count as a violation of the rights of the accused just because it would be a substantial diminution in the level of safety provided at the center of the criminal process for so long.]
DWORKIN, supra note 303, at 90–91.
368 FED. R. CRIM. P. 23(a), 31(a) (requiring, in order for a conviction, a jury of twelve persons unanimously voting for a guilty verdict).
369 DRESSLER, supra note 222, at 4.
370 Williams v. Florida, 399 U.S. 78 (1970) (holding that a six-person jury is constitutional under the Sixth Amendment right to a jury trial); see Ballew v. Georgia, 435 U.S. 223 (1978) (finding a five-person jury unconstitutional).
as they represent a "substantial majority."\textsuperscript{371} The requirement of unanimity, or at least a substantial majority, in order to convict and punish, roughly parallels our presumption of innocence under which a defendant may not be convicted unless the prosecution establishes proof of guilt beyond a reasonable doubt. Both the size and unanimity (or substantial majority) requirements reflect a bias in favor of acquitting the guilty rather than punishing the innocent.

But even with a requirement of a unanimous verdict of twelve jurors to convict, some innocents will be convicted and punished.\textsuperscript{372} "Perhaps if the concern were solely with guaranteeing that no innocent man be convicted, a twenty or thirty man jury in which unanimous consent was required for conviction would do a better job."\textsuperscript{373} And, of course, a fifty or one-hundred-person jury requiring unanimous consent for conviction would do an even better job still of avoiding punishment of the innocent. While doing so would come closer\textsuperscript{374} to satisfying one of the duties of retributivism—not punishing the innocent—it would do a poor job of satisfying the other duty of retributivism—punishing the guilty.\textsuperscript{375}

If, instead, equal importance were placed on not punishing the innocent and on punishing the guilty, then the traditional use of a unanimous vote of twelve jurors would be eliminated. Rather than a unanimous vote of twelve jurors, a non-unanimous vote of, for example, six (out of twelve) jurors would be implemented. This might very roughly equalize the number of false acquittals and false convictions. Though not satisfying either one of the categorical duties of retributivism, a vote of six jurors for conviction on a twelve-person jury would very roughly value the two duties equally. Of course, implementation of such a jury system would lead to a substantial increase (over our traditional unanimous twelve-person jury practice) in punishment of the innocent.

And if, instead, comparatively greater importance was placed on the retributivist absolute dictate of conviction and punishment of the guilty, then our traditional practice of conviction based on a unanimous vote of twelve jurors could be replaced with a system of guilt unless there was a unanimous vote for acquittal. That is, even, an 11-1 vote in favor of acquittal would constitute a guilty verdict. And, of course, a system of guilt unless a fifty or one-hundred person jury unanimously voted for acquittal


\textsuperscript{372} Dworkin, supra note 303, at 72 ("If we continue to use only twelve jurors in order to save the extra expense [of a 25-person jury], that will result in some people being convicted though innocent.").

\textsuperscript{373} Wasserstrom, supra note 314, at 739.

\textsuperscript{374} Of course, even with the unanimous consent of a one-hundred person jury, at least some innocents would be mistakenly convicted.

\textsuperscript{375} If a unanimous vote of, for example, a 100-person jury was required for conviction, substantially fewer guilty defendants would be convicted. The implementation of such a jury in a retributivist system would reflect the comparatively greater importance placed on avoiding punishment of the innocent over foregoing punishment of the guilty.
Deterring Retributivism would come even closer still to attaining the retributivist duty of punishing the guilty. While implementing such a system would allow markedly fewer guilty defendants to escape punishment, it would greatly increase the number of innocents who would be punished.

Given a standard of "guilt unless there is a unanimous vote for acquittal of an \( X \) person jury," where \( X \) is a sufficiently high number, virtually every innocent defendant would be convicted. Retributivism, via DDE, would justify such a jury standard for guilt in criminal trials. Though it would be known or foreseen that virtually all innocent defendants would be convicted and punished under such a jury standard, it would also be known or foreseen that at least some (if not virtually all) factually guilty defendants would be found guilty. It would not be known that any particular, identifiable innocent was found guilty; no particular, identifiable innocent would be intentionally convicted and punished. The intended good effect in adopting and maintaining such a jury standard would be conviction and punishment of the guilty; the bad side effect of conviction and punishment of virtually every innocent defendant would be unintended. Thus, via DDE, retributivism permits and justifies the adoption of, and punishment under, such a jury standard.

The retributivist criticism—that consequentialism justifies the erosion of procedural safeguards designed to protect the innocent—may be turned back on retributivism. It is retributivism, because of DDE, that is unable to explain the wrong of such a jury system in which virtually every innocent defendant would be found guilty and punished. Again, perhaps the problem lies not with retributivism but with its reliance on DDE. But without DDE, retributivism's intentional maintenance of fallible punishment institutions which are known to inevitably and mistakenly punish the innocent is unjust by retributivism's own criteria.

D. Objections

This subpart will consider three possible objections to the conclusion that DDE renders retributivism compatible with justifying indirect punishment of the innocent, that is, punishment without due regard to defendants' fault, culpability, or guilt. First, retributivism would not justify punishment under standards of presumed guilt, strict liability, an extreme form of absolute liability, or guilty verdicts by a non-majority of jurors because that would fail to constitute efforts to avoid punishment of the innocent. Second, retributivism does not justify punishment under such standards and practices because it would amount to failing to take reasonable efforts against punishing the innocent. And third, even if retributivism does justify punishing under such standards, this is not a shortcoming of retributivism as a theory but is due only to its application to the real world. All of these objections will be shown to be either unpersuasive or, even if persuasive, equally applicable to absolve consequentialism of retributivists' criticisms.
1. **Efforts to Avoid Punishment of the Innocent.**—The creation and maintenance of institutions that seek to punish defendants under standards of presumed guilt, strict liability, an extreme form of absolute liability, or non-majority jury verdicts might fail to constitute efforts to avoid punishment of the innocent, faultless, or non-culpable.\(^{376}\) According to Duff, though a retributivist need not “take every possible precaution against the risk of punishing those who are in fact innocent,” punishment of the innocent should only occur “despite our intention and our efforts to avoid it.”\(^{377}\) Thus, according to the objection, retributivism would not be compatible with, and would not justify, punishment under these standards.

The objection is unpersuasive. DDE does make retributivism compatible with such standards. As Duff noted, retributivism need not take every possible precaution against punishing the innocent. After all, the standards of proof beyond a reasonable doubt and preponderance of the evidence (both of which are claimed to be compatible with retributivism\(^{378}\)) do not take every possible precaution against punishing the innocent. Retributivists acknowledge that even under the reasonable-doubt standard it is expected and known that some percentage of defendants found guilty and punished will be innocent. If retributivism wished to take every possible precaution against punishing the innocent, it is free to adopt a standard of “presumption of innocence unless *absolute* evidence of guilt is established.” But retributivists’ concern with avoiding punishment of the innocent does not extend this far (perhaps because of the equal duty to punish the guilty). The standards of strict culpability, presumption-of-guilt standard of proof, and non-majority jury verdicts do constitute some effort to avoid punishment of the innocent. The standards allow some innocents to escape conviction and punishment.\(^{379}\)

Since the standards of (i) preponderance of the evidence and (ii) presumption of innocence are claimed to be compatible with retributivism, and thus apparently constitute some effort to avoid punishment of the innocent, then the standards of (iii) strict culpability, (iv) presumption of guilt and (v) non-majority jury verdicts are also compatible with retributivism and constitute efforts to avoid punishment of the innocent. None of these five standards constitutes taking every possible precaution against punishing the innocent. True, the former two perhaps constitute more of an effort and the latter three constitute less of an effort. Nonetheless, all five constitute some effort to avoid punishment of the innocent. The only difference between the former two standards and the latter three is a matter of degree.\(^{380}\)

---

\(^{376}\) I am grateful to George Fletcher for posing this possible objection.

\(^{377}\) **DUFF, supra** note 1, at 159.

\(^{378}\) See *supra* notes 356–357 and accompanying text.

\(^{379}\) The same is not true, however, for punishing under our extreme form of absolute liability.

\(^{380}\) The objection might still be made that it is not merely a matter of degree. Retributivism justifying punishment (i) of strict liability offenses, (ii) under a presumption of guilt standard and, (iii) non-
2. Reasonable Efforts to Avoid Punishment of the Innocent.—A retributivist might still object that something more than “some” effort to avoid punishment of the innocent might be required for a standard or practice to be justified by retributivism. So, between making some effort and every possible effort to avoid punishment of the innocent, what exactly is the quanta of effort required to avoid punishment of the innocent? Duff suggests that the proper measure is those efforts that are “reasonable.”381 Unfortunately, Duff does not elucidate what would constitute reasonable precautions.

Let us assume that from a common-sense perspective the standards and practices we have sought to make compatible with retributivism would plainly not constitute reasonable efforts to avoid punishment of the innocent. This is because under such standards and practices many, or virtually every, innocent defendant would be convicted and punished.382 Implementing standards and practices that lead to punishing virtually every innocent defendant hardly constitutes, from a common-sense perspective, reasonable efforts to avoid punishment of the innocent.

But the common-sense perspective may not be available to retributivism. To see why, we need only consider why Duff and other retributivists refuse to take all possible precautions against punishing the innocent. For majority jury verdicts is not merely quantitatively, but also qualitatively, different from punishing under the standards of reasonable doubt, preponderance of the evidence and jury verdicts meeting constitutional size and unanimity (or substantial majority) requirements. Though punishment under both the former three and latter two standards will inevitably lead to the punishment of the innocent, the former will lead to punishment of the innocent to such a greater extent as to constitute either aiming at punishment of the innocent or for the extent to be so great as to not be a matter of degree but qualitatively different. As such, retributivism does not justify punishment under the former three standards.

The objection is unpersuasive. Justifying punishment of defendants convicted under nonmajority jury verdicts is merely quantitatively different than justifying punishment under our current jury size and unanimity (or substantial majority) requirements. Requiring a unanimous vote of 12 jurors for conviction is only different by a matter of degree from a unanimous vote of 10, 6 or 2 jurors. But every decrease in the number of jurors will incrementally increase the odds of punishing the innocent. Similarly, replacing the present system of a unanimous vote of 12 jurors for conviction is only different by a matter of degree from a standard requiring (out of a 12 person jury) only 10, 6 or 2 votes for conviction. But each decrease in the number of juror votes for guilt required for a conviction will incrementally increase the number of innocent defendants found guilty. And requiring for conviction only 2 (out of 12) guilty votes is different only by a matter of degree from requiring, for an acquittal, a unanimous not guilty vote of 12 jurors, 10 jurors or 1000 jurors. But each increase in the number of unanimous juror votes for an acquittal will incrementally increase the number of innocent defendants mistakenly convicted.

381 DUFF, supra note 1, at 159 (“A practicable system of punishment cannot satisfy the former demand [of taking every possible precaution against punishing the innocent]; it can take all ‘reasonable’, but not all possible, precautions against the risk.”); Duff, supra note 2, at 424 (requiring “reasonable safeguards against mistaken convictions (this being what retributivists must accept as a feature of any human system of punishment”).

382 It might be quibbled that punishing under a strict liability standard will not significantly lead to more innocents punished than under a preponderance of the evidence standard. And since the latter is claimed to be compatible with retributivism, see supra note 357 and accompanying text, so should the former.
all of retributivists' rhetoric about the moral abhorrence of punishing the innocent, it might seem odd that retributivists do not advocate making all possible efforts to avoid punishing the innocent. Even short of making all possible efforts, retributivists do not even advocate making additional efforts (beyond those already in place in our current systems) to avoid punishment of the innocent. Surprisingly, retributivists even suggest that we could make less effort to avoid punishment of the innocent. Recall that retributivists have suggested the possibility of punishing under the standard of proof used in civil cases (guilt by a preponderance of the evidence). This is all because in addition to not punishing the innocent, retributivism has the affirmative duty to punish the guilty. The more efforts are made to avoid punishing the innocent, fewer of the guilty are punished.

From the retributivist perspective, then, what shall be deemed reasonable? Given the two duties of retributivism, which in a practical sense are countervailing, it does not seem to be intelligible for retributivism to speak of reasonable efforts to avoid punishment of the innocent. Any practicable effort made to avoid punishing the innocent is equally an effort to avoid punishing the guilty. Because of the affirmative duty to punish the guilty and the negative duty to not punish the innocent, to the extent that the two duties are practicably reconcilable, it would seem that any arrangement that achieved one or the other duty, or some combination thereof, would be considered reasonable. If so, then standards and practices that convict and punish many or all guilty defendants, despite convicting and punishing virtually all innocent defendants, are reasonable, not from a common-sense perspective, but from a retributivist perspective.

But let us suppose for the sake of argument that (i) the requirement of making reasonable efforts to avoid punishment of the innocent is intelligible under retributivism, and (ii) standards and practices in which all (or virtually all) of the guilty and the innocent are punished do not constitute reasonable efforts to avoid punishment of the innocent. DDE still makes such standards compatible with retributivism. These standards and practices do not entail the intentional punishment of any particular, identifiable innocent. Though it is known or foreseen that virtually every innocent defendant would be convicted and punished, no particular, identifiable innocent is intentionally punished. Since only the good effect of punishment of the guilty is intended and the bad side effect of punishment of the innocent or nonculpable is unintended, the conditions for DDE’s applicability are satisfied. This holds true even for the extreme form of absolute liability in which every innocent defendant would be convicted and punished. It is retributivists themselves who have informed us, via DDE, that known or fore-

383 See, e.g., DWORKIN, supra note 303, at 79–92 (defending the position that further procedural safeguards are not necessary in order for our system of criminal procedure to be just).
384 See supra note 357 and accompanying text.
385 See supra notes 344, 367, 373 and accompanying text.
seen punishment of some unknown innocents is permissible as long as no innocents are intentionally punished. If justifying the implementation of standards and practices in which all, or virtually all, innocents would be convicted and punished satisfies DDE, but still does not constitute reasonable efforts to avoid punishing the innocent, then the reasonable efforts requirement is inconsistent with DDE.\textsuperscript{386}

To avoid the inconsistency, retributivism would be forced to eliminate its reliance on one or the other. If DDE were eliminated, retributivism would have no defense to consequentialist criticisms that retributivism is unjust by its own criteria because it justifies the implementation of punishment institutions where it is foreseen that innocents will inevitably be convicted and punished. But if retributivism eliminates the requirement of making reasonable efforts to avoid punishment of the innocent, it is compatible with standards and practices in which virtually every innocent defendant will be convicted and punished. Either way, retributivism, as to punishment of the innocent, is no better than consequentialism, and perhaps worse.

3. Theory Versus Application.—The objection might be that, as a theory, retributivism has no problem with justifying punishment of the innocent, and that it is only retributivism’s application to the uncertainties of real world punishing that raises difficulties.\textsuperscript{387} Suppose the finder of guilt and innocence is God\textsuperscript{388} or is omniscient. Guilt, innocence, and culpability would be known with absolute certainty. There might then be no need for standards of guilt, innocence, and culpability.\textsuperscript{389} The guilty would simply be punished and the innocent acquitted under retributivism; no guilty would be mistakenly acquitted and no innocent would be mistakenly punished, because there would be no mistakes. Retributivism’s problems with justifying extreme standards of proof and culpability would thereby be eliminated. It is not retributivism as a theory that is problematic, one could argue, but merely its application to punishment under conditions of epistemic uncertainty in the real world.

The objection is unpersuasive for a number of reasons. First, retributivism does not merely aspire to theoretical validity. Retributivism is invoked as relevant for punishment in the real world where virtually every

\textsuperscript{386} Even if the reasonable efforts requirement is consistent with DDE, punishing under a strict liability standard might still be justified, via DDE, by retributivism. Presumably, punishing under a strict liability standard would not lead to punishing an appreciably greater number of innocents than punishing under a preponderance of the evidence standard, which is claimed to be compatible with retributivism and thus presumably constitutes reasonable efforts to avoid punishing the innocent. Therefore, if the preponderance of the evidence standard is compatible with retributivism and constitutes reasonable efforts to avoid punishment of the innocent, then the same holds true for strict liability.

\textsuperscript{387} I am grateful to Ken Levy for posing this possible objection.

\textsuperscript{388} See Schedler, supra note 296, at 193–94 (suggesting that retributivism might require God or an omniscient factfinder for retributivism to realize its ambitions).

\textsuperscript{389} Id.
instance of punishing is under conditions of epistemic uncertainty. Unless retributivism is willing to be purely an ethical theory with no application to legal punishment, its application to the real world is relevant and crucial.

Second, it is retributivism as a theory that is compatible with, or justifies (via DDE), punishing under the standards of strict liability, an extreme form of absolute liability, presumption of guilt, and non-majority guilty verdicts. The criticism is, in fact, directed at retributivism as a theory, a theory that justifies or is compatible with extreme standards of guilt, proof, and liability. The issue is not whether retributivism does use such standards, or would not under certain conditions, but whether retributivism as a theory justifies (via DDE) the use of such standards and whether retributivism can explain why the use of such standards is wrong. (Recall that retributivists have criticized consequential theories not so much for actually punishing innocents but for their inability to explain the wrong of punishing the innocent.) Even if retributivism’s factfinder was God or was omniscient (and thus had no need for standards in order to determine guilt and innocence), because of DDE, retributivism would be unable to explain why the adoption of, and punishment under, such standards is wrong.

Let us, however, assume arguendo that the objection is persuasive. Could consequentialism equally avail itself of the same argument? That is, under conditions of perfect knowledge of who committed each crime and which defendants are guilty and which are innocent, would a consequentialist theory nonetheless justify punishment of the innocent? No. Consequentialist theories hold that punishment of an innocent lowers societal welfare to a greater extent than punishment of a guilty person. This is because the innocent would suffer more than the guilty and because punishment of the innocent would probably be counterproductive. For Bentham, because punishment by itself is an evil, punishment should only be imposed where unavoidable. In situations where the identity of the guilty offender is known, punishment of an innocent would be avoidable. As Bentham states, “To inflict punishment when . . . the infliction of punishment is unavoidable, is in the case of the innocent, contrary to the principle of util-

---

390 As Schedler explains:

We know all too well that a system of legal punishment will on occasion convict and punish by mistake individuals who are innocent . . . . To admit this is merely to admit that well-intentioned human beings are fallible and imperfect and that, as a consequence, their legal systems will reflect these characteristics.

Id. at 187.

391 See supra notes 158–159 and accompanying text.

392 See supra note 201 and accompanying text.

393 See supra notes 153–155 and accompanying text.

394 See supra notes 161–162 and accompanying text.

395 See supra notes 160–161 and accompanying text.
ity. Thus, under conditions of perfect knowledge consequentialism would also no longer be subject to the criticism of punishing the innocent.

Moreover, if retributivism’s employment of constructs of idealized worlds (for example, where the factfinder is omniscient) persuasively eludes retributivism’s difficulties, then consequentialist theories could equally avoid criticism by constructing its own idealized worlds. Rather than a world of perfect knowledge, suppose a world with perfect deterrence. In such an idealized world of perfect deterrence, no crimes would be committed because all crime would be deterred. Thus, a consequentialist theory would have no need to justify intentional punishment of the innocent. It is only the messy real world where people actually commit crimes that causes difficulty for consequentialism.

If it is not retributivism as a theory that is problematic as to punishment of the innocent, but only retributivism’s application to the real world, then equally it is not consequentialism as a theory that is problematic but consequentialism’s application to the real world. If instead the real world is relevant in assessing a theory of punishment, then retributivism is compatible with justifying punishment under standards and practices in which every, or virtually every, innocent defendant would be convicted and punished. Either way, as to punishment of the innocent, retributivism fares no better than consequentialist theories of punishment.

E. Conclusion

Retributivists have found consequentialist theories to be hopelessly defective for justifying the intentional punishment of the innocent. In response, consequentialists argue that by intentionally creating and maintaining necessarily fallible punishment institutions in which it is known that some innocents will inevitably be punished, retributivism is unjust by its own criteria. While conceding that some innocents will inevitably be punished by mistake, retributivists deny that it is unjust because no particular, identifiable innocent will be knowingly or intentionally punished under retributivism. Invocation of some variant on the doctrine of double effect, or DDE, then, purports to whitewash over the problem of punishing the innocent for retributivism. So long as no particular, identifiable innocent is intentionally punished, retributivists claim that punishment of the innocent does not render their theory unjust or inconsistent with its own criteria.

Even if DDE is valid and if its conditions can be satisfied, however, it causes as many difficulties as it solves. Via DDE, retributivism justifies the punishment of defendants without due regard to fault or culpability (a criticism which retributivists have directed at consequentialist theories). By its reliance on DDE, retributivism justifies the adoption of, and punishment

396 Bentham, supra note 63, at 476.
under, standards of (i) strict liability, even for serious offenses, (ii) an extreme form of absolute liability, (iii) a presumption of guilt standard of proof, and (iv) non-majority jury verdicts. Retributivism is compatible with justifying punishment under such extreme standards because even though it would be foreseen that virtually every innocent defendant would be convicted and punished, no particular, identifiable innocent would be intentionally punished. While the good effect of punishing the guilty is intended, the bad side effect of punishing the innocent is unintended. Moreover, just as retributivists have claimed that consequentialist theories cannot explain why intentional punishment of the innocent is wrong, so too retributivism (because of DDE) cannot explain why justifying punishment under extreme standards of liability and procedure in which all or virtually all innocent defendants would be convicted and punished is wrong. By avoiding intentional punishment of particular, identifiable innocents, retributivism is subject to its own criticism of justifying punishment without due regard to fault or culpability.

Retributivism could avoid the above difficulties, of course, by not invoking DDE. But without it, retributivism is unjust by its own criteria by intentionally creating and implementing necessarily fallible institutions that are known to inevitably punish innocents. Either way, as to punishment of the innocent, retributivism fares no better than consequentialist theories.

V. RETRIBUTIVISM USES PERSONS NOT AS ENDS BUT AS MERE MEANS

Perhaps the “primary” or most “fundamental” criticism retributivists level against consequentialist theories of punishment is that they use those who are punished as mere means to attain the good consequences sought to be promoted, rather than as ends in themselves. Retributivism, in contrast, allegedly treats offenders as ends in themselves by giving off-

---

397 MORRIS, supra note 6, at 46 ("The primary reason for preferring the system of [retributive] punishment as against the [consequentialist] system of therapy might have been expressed in terms of the one system treating one as a person and the other not.").
398 DUFF, supra note 1, at 149 ("The fundamental moral objection to any consequentialist account . . . will be that it falls to accord to the citizen, and to the criminal, the respect which is their due."); Greenawalt, supra note 11, at 1341 ("The most fundamental objection [to consequentialist accounts] is to treating the criminal as a means to satisfy social purposes rather than as an end in himself.").
399 For example, a deterrence theory of punishment uses those who are punished as mere means in order to attain the end of the deterrence of crime. The offender is used as a mere means in order to benefit society. Though an offender might be used as a means to benefit others (general deterrence), the goals of special deterrence, incapacitation and rehabilitation are arguably to the benefit of the offender being punished. How then are consequentialist theories of punishment using offenders as mere means when the offender himself is benefitted? Retributivists argue that the prohibition against using offenders as mere means extends even to instances where the offenders are benefitted. Special deterrence and rehabilitative theories of punishment give offenders what they need. General deterrence theories give society what it needs (the reduction of crime). But according to retributivism, only by giving offenders what they deserve, and not what they need, are wrongdoers respected as fully autonomous agents and as ends in themselves.
Deterring Retributivism

fenders only what they deserve. But retributivism, it will be demonstrated, also uses persons as mere means.

After explaining the means/ends distinction, we will briefly consider the state of the debate as to the distinction’s relevance in the theory of punishment. The principal focus of this Part, however, will be to make the novel argument that retributivism, in violation of Kant’s maxim, 

versely uses crime victims as mere means in order to treat their victimizers (criminal offenders) as ends in themselves. Retributivism exploits crime victims as, in Karl Marx’s inimitable phrase, “slave[s] of justice.”

A. Means and Ends

Though the nexus for the distinction between using persons as mere means and treating persons as ends in themselves is generally traced to Kant, a utilitarian, Cesare Beccaria, anticipated by some twenty years Kant's maxim against disrespecting persons’ humanity and autonomy by using them as mere means. In arguing against impermissible forms of punishment, Beccaria declares: “There is no liberty whenever the law in some cases permits a man to cease to become a person and to become a thing.” Perhaps influenced by Beccaria, Kant announces that one must, “Act so that you treat humanity, whether in your own person or that of another, always as an end and never as a means only.” Applying this to punishment, Kant contends:

Punishment by a court ... can never be inflicted merely as a means to promote

400 Duff, supra note 6, at 10.

401 See infra notes 407–408 and accompanying text.

402 Karl Marx, Capital Punishment, N.Y. DAILY TRIB., Feb. 18, 1853 (praising retributivism, as embodied in the views of Kant and Hegel, for not “looking upon the criminal as the mere object, the slave of justice,” but rather, raising the criminal “to the position of a free and self-determined being”). Impliedly, Marx is criticizing consequentialist theories for using criminals as mere objects and slaves of justice for consequentialist purposes. But Marx goes on to question whether any approach to punishment, including retributivism (which he terms “specious”), under existing social conditions, is morally adequate. As Marx puts it: “Is there not a necessity for deeply reflecting upon an alteration of a system that breeds these crimes, instead of glorifying the hangman who executes a lot of criminals to make room only for the supply of new ones?” Id.

403 See, e.g., Dolinko, supra note 2, at 1631 (footnote omitted) (“Kant, whose writings are the principal source of the idea that using people, or treating them as ‘means’ (solely as means, as Kant was careful to put it), is uniquely immoral . . . .”).

404 Beccaria’s On Crimes and Punishments was published in 1764, see supra note 102, over 20 years prior to Kant’s first expression of the maxim in 1785, see infra note 407.

405 BECCARIA, supra note 102, at 38 (Ch. XX) (translator’s footnote omitted); cf. KANT, supra note 22, at 140–41 (a person “can never be treated merely as a means to the purposes of another or be put among the objects of rights to things”).

406 David Young, the translator, suggests that Kant might have been influenced by this passage of Beccaria in developing his theory of punishment. BECCARIA, supra note 102, at 93 n.2 (Ch. XX).

some other good for the criminal himself or for civil society.... For a man can never be treated merely as a means to the purposes of another or be put among the objects of rights to things: His innate personality protects him from this . . . .

1. The Retributivist Charge: Consequentialism Uses Offenders as Mere Means.—Retributivists insist that their theory respects and honors offenders as ends, but that consequentialist theories disrespect offenders by using them as mere means or animals. As Hegel explains, by retributivism conceiving punishment as "embodying the criminal's own right, the criminal is honoured as a rational being." In turn, the failure to so honor criminals leads Hegel to reject consequentialist theories of punishment. The criminal is denied being treated as a rational being "if he is regarded simply as a harmful animal which must be rendered harmless, or punished with a view to deterring or reforming him." Speaking of deterrent theories based on threats, Hegel declares that "[t]o justify punishment in this way is like raising one's stick at a dog; it means treating a human being like a dog instead of respecting his honour and freedom." As a contemporary philosopher makes the argument:

[T]he retributivist seeks, not primarily for the socially useful punishment, but for the just punishment, the punishment that the criminal (given his wrongdoing) deserves or merits, the punishment that the society has a right to inflict and the criminal a right to demand. Only a theory built on these values, so a common argument goes, will respect persons as individuals of special worth—a worth that is compromised if we feel free simply to use them (as utilitarian deterrence theory appears willing to use them) for the social good.

KANT, supra note 22, at 140-141. This famous passage has also been translated as follows: "Judicial punishment can never be administered merely as a means for promoting another Good either with regard to the Criminal himself or to Civil Society.... For one man ought never to be dealt with merely as a means subservient to the purpose of another...." IMMANUEL KANT, THE PHILOSOPHY OF LAW 195 (W. Hastie trans., 1887).

There is some debate as to the matter whether Kant meant that persons must always be treated exclusively as ends or whether they may be treated as means as long as they are not exclusively treated as means. For agreement with the latter interpretation, see HART, supra note 11, at 244 ("Kant never made the mistake of saying that we must never treat men as means. He insisted that we should never treat them only as means 'but in every case as ends also.'").

See infra notes 412-413, 435 and accompanying text.

HEGEL, supra note 22, at 126 (§ 100). As one commentator construes Hegel's view, punishment is both an honor as well as a disgrace. J.E. McTaggart, Hegel's Theory of Punishment, 6 INT'L. J. ETHICS 482 (1896), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 40, 41.

Id. at 125–26 (§ 99). One commentator argues that for Hegel the deterrent effect of punishment is not undesirable but "it is beside the point. Mere deterrence fails to 'erase' the crime that has already been committed and thus fails to negate the negation." Steinberger, supra note 89, at 861.

Because, under retributivism, punishment is intrinsically good and the consequences of punishment are irrelevant so long as deserving wrongdoers receive their just deserts, retributivism is perceived as not using offenders as mere means or instruments to attain some other goal. In contrast, under a consequentialist theory justifying punishment by deterrence, for example, punishment is not intrinsically good but only instrumentally good in attaining some other goal (deterrence). The claim is that deterrence-based theories use the punished offenders as mere instrumental means in order to obtain the intrinsic goods of reducing the incidence of crime and the victimization of crime victims.

2. Accounts of Using Persons as Mere Means.—Before turning to consequentialist responses to this criticism, let us briefly consider some general conceptions of using another person as a mere means to an end. Although its importance as a moral criticism is recognized to be substantial, there is less consensus as to what it means to use another as a mere means. In general, using another as a mere means is thought to belong in the same family as the concepts of "manipulation, dehumanization, exploitation, and disrespect." The notion of using another as a mere means may also be understood in the negative—as failing to respect another’s inviolability, rationality, or moral status.

But let us consider some various accounts of using another as a mere means. Kantian scholar Onora O’Neill explains that to use something as a mere means is to treat it as a thing, instrument, or tool. Robert Nozick

---

415 See supra note 107 and accompanying text; infra notes 609–610 and accompanying text.
416 See supra note 400 and accompanying text.
417 See supra notes 108–109 and accompanying text; infra note 612 and accompanying text.
418 See supra notes 397–399 and accompanying text.
420 The belief that it is wrong to use persons is one that is well entrenched in our moral thinking. In one form or another, it underlies a wide range of criticisms and analyses.
421 Onora O’Neill, Between Consenting Adults, 14 PHIL. & PUB. AFF. 252, 252 (1985) ("Few moral criticisms strike deeper than the allegation that somebody has used another. . . . But this consensus is often shallow, since there is little agreement about what it takes to use others in morally problematic ways."); Dolein, supra note 2, at 1631 (footnote omitted) ("And although the immorality of using people is today widely accepted in ordinary discourse, even this commonsense usage is both complex and elusive.").
422 Davis, supra note 419, at 387 (footnote omitted).
423 Id. at 403.
424 As O’Neill puts it:

To treat something as a mere means is to treat it in ways that are appropriate to things. Things, unlike persons, are neither free nor rational; they lack the capacities required for agency. They can only be props or implements, never sharers or collaborators, in any projects. Things cannot act . . . so cannot consent to or dissent from the ways in which they are used.
finds the metaphor of a tool to be the paradigmatic example of using another as a mere means because there are no moral constraints on how we may treat a tool. In addition to using another as a tool or commodity, using another without their consent is another common conception of using someone as a mere means. Declaring that "[i]ndividuals are inviolable," Nozick interprets Kant's maxim as requiring that individuals "may not be sacrificed or used for the achieving of other ends [than their own] without their consent." Affirming the importance of treating persons as ends in themselves and never as mere means, John Rawls declares that consent is central to not only explicating the meaning of Kant's maxim but also its realization. O'Neill also sees consent as crucial to an adequate understanding of the distinction between treating a person as an end and using the person as a


O'Neill, supra note 420, at 252 ("Making another into a tool or instrument in my project is one way of failing to treat that other as a person . . . ").

As Nozick explains:

There is no side constraint on how we may use a tool, other than the moral constraints on how we may use it upon others. There are procedures to be followed to preserve it for future use ("don't leave it out in the rain"), and there are more and less efficient ways of using it. But there is no limit on what we may do to it to best achieve our goals.

NOZICK, supra note 36, at 31.

Id. ("If we add constraints on its use that may not be overridden, then the object may not be used as a tool in those ways. In those respects, it is not a tool at all.").

Davis, supra note 419, at 392 ("When we object to one person's using another, what we are saying (at least some of the time) is that something that is not a mere good or commodity is being treated as one.").

NOZICK, supra note 36, at 30-31.

JOHN RAWLS, A THEORY OF JUSTICE 179 (1971) ("[T]he principles of justice manifest in the basic structure of society men's desire to treat one another not as means only but as ends in themselves.").

More generally, Rawls expresses the importance of respecting persons as ends in this famous passage starting on the first page of his classic work:

Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advances enjoyed by many.

Id. at 3-4.

As Rawls puts it:

[T]reating men as ends in themselves implies at the very least treating them in accordance with the principles to which they would consent in an original position of equality.

. . . By contrast, to regard persons as means is to be prepared to impose upon them [without their consent] lower prospects of life for the sake of the higher expectations of others.

Id. at 180.
means, but in a way different than Rawls. Rather than the possible or rational consent of Rawls’s original position, O’Neill argues for the relevance of actual consent.

Perhaps the most elegant conception of the distinction is Herbert Morris’s account of treating someone as a person. That is, to treat someone as an end is nothing more than treating him as a person, and, using someone as a mere means fails to treat someone as a person.

Now that we have some sense of what is involved in treating someone as an end rather than as a mere means, the next step is to determine to whom we owe this duty. For Kant, this duty is owed to all of humanity: “Act so that you treat humanity, whether in your own person or that of another, always as an end and never as a means only.” O’Neill, however, offers a narrower scope of the duty. Rather than all of humanity, those who are mere spectators to, or “wholly unaffected by,” a course of action, are not treated as mere means even if they do not actually consent to the action. But consent must be at least possible (even if there is no actual consent), in

---

431 O’Neill, supra note 420, at 253 (“[A]n adequate understanding of what it is to treat others as persons must view them not abstractly as possibly consenting adults, but as particular men and women with limited and determinate capacities to understand or to consent to proposals for action.”).

432 As O’Neill explains:

“It is morally objectionable to treat others in ways to which they do not consent. To do so treats another as a thing or a tool, which cannot, so does not, consent to the ways in which it is used; such so fails to treat others as persons, who can choose, so may withhold consent from actions which affect them.” Id. at 254.

433 Morris suggests that “the primary reason” reason to prefer a retributivist to a consequentialist system of punishment might be that the former does, and the latter does not, treat others as persons. Morris, supra note 6, at 46.

434 Though Morris does not specifically track the means/ends distinction, his account of treating another as a person may eloquently capture what is involved in treating another as an end rather than as a mere means. Id. at 46–50. For Morris, two conditions must be satisfied in order to treat someone as a person: “We treat a human being as a person provided, first, we permit the person to make the choices that will determine what happens to him and, second, when our responses to the person are responses respecting the person’s choices.” Id. at 48–49.

435 Morris explains what constitutes not treating another as a person:

When we talk of not treating a human being as a person or “showing no respect for one as a person” what we imply by our words is a contrast between the manner in which one acceptably responds to human beings and the manner in which one acceptably responds to animals and inanimate objects. When we treat a human being merely as an animal or some inanimate object our responses to the human being are determined, not by his choices, but ours in disregard of or with indifference to his. And when we “look upon” a person as less than a person or not a person, we consider the person as incapable of rational choice. In cases of not treating a human being as a person we interfere with a person in such a way that what is done, even if the person is involved in the doing, is done not by the person but by the user of the person.

Id. at 46.

436 KANT, supra note 407, at 46 (emphasis added).

437 O’NEILL, supra note 423, at 139 n.12. O’Neill further explains that “[t]here is not much difficulty in ensuring that those who will in any case be no more than spectators have a genuine possibility of dissent.” Id. at 110.
order to treat others as ends and avoid using others as mere means, for “those closely involved in or affected by” a course of action.\footnote{O'Neill, supra note 420, at 259.} We will return to O’Neill’s narrower test as to who may, and who may not, be treated as a mere means in making the argument that retributivism uses persons as mere means.

3. **Consequentialist Responses.**—Against the criticism by retributivists that consequentialist theories use those punished as mere means, consequentialists have pitched a number of counterarguments. First, they argue, the concept of using someone as a mere means is “unclear,”\footnote{Dolinko, supra note 2, at 1631.} “notoriously obscure,”\footnote{DUFF, supra note 1, at 178.} “vague,”\footnote{HONDERICH, supra note 89, at 49. For three different possible interpretations of the means/ends distinction, see id. at 49–50.} “elusive,”\footnote{Davis, supra note 419, at 389. Davis expresses skepticism that the criterion of using persons as mere means can, consistently with our common sense, bear the weight which is placed on it: I am inclined to doubt that our commonsense views about using persons can play an important role in philosophical argument, either in the construction or in the criticism of moral theories. Two considerations underlie this skepticism: first, the things that philosophers have said about using persons do not happily characterize our commonsense notions, and second, a closer look at these notions reveals them to be elusive in important ways, ones that make their philosophical application problematic. Id. at 388–89.} and enjoys “little agreement” among commentators.\footnote{O’Neill, supra note 420, at 252.} As such, the criticism is too imprecise to warrant, or to form, a response. But since the criticism has come to play such a crucial and central role in the punishment debate,\footnote{DUFF, supra note 1, at 178 (explaining that though the concept of using persons as mere means is “notoriously obscure . . . it generates a crucial objection” to consequentialist theories of punishment). On the importance of this criticism of consequentialist theories, see also supra notes 382–383.} it cannot be brushed away this easily.\footnote{That a number of notable contemporary philosophers have regarded the inability of consequentialism to adequately explain our intuition that using persons as mere means is wrong as sufficient reason to reject consequentialism should give us pause in too hastily rejecting the using persons as mere means criterion. See generally RONALD DWORKIN, TAKING RIGHTS SERIOUSLY (1977); CHARLES FRIED, RIGHT AND WRONG (1978); NOZICK, supra note 36; RAWLS, supra note 429; Nagel, supra note 184.} Second, a consequentialist might concede that consequentialist theories of punishment use offenders as mere means but might nonetheless deny that it is problematic.\footnote{C.L. Ten, supra note 29: [I]f we punish those who voluntarily breached the law in order to prevent them from repeating their offences, or to deter potential offenders, we are not using them unfairly. The failure to punish in these cases would result in there being additional innocent victims of crime. These victims could not reasonably have avoided being harmed by criminal acts in the way that those who voluntarily broke the law could have refrained from criminal acts and therefore avoided the resultant punishment. Id. at 369–70.} Responding to Kant’s imperative not to use
people as mere means, Oliver Wendell Holmes proclaimed, "[i]f a man lives in society, he is liable to find himself so treated." Furthermore, Holmes added, "No society has ever admitted that it could not sacrifice individual welfare to its own existence." In other words, the treating of people as mere means is normal, necessary, inevitable, and thus permissible, and we should not pretend otherwise.

Third, it might be conceded that consequentialist theories use offenders as means but not as mere means. Consequentialist punishment of the guilty might use them as means to attain the good consequences of punishment but also treats the guilty as ends by giving them what they deserve. Furthermore, that one goal of deterrence-based theories is special deterrence means that offenders who are punished are not being used solely as a means to deter others. In sentencing an offender, one judge had occasion to comment on Kant's maxim finding both that using others as mere means is not problematic and that punishment for deterrence reasons fails to use offenders as mere means. Faced with defense counsel's invocation of Kant's maxim in arguing against the need for the defendant to be punished for reasons of deterrence, Judge Frankel in United States v. Bergman concluded:

[Kant's] humane principle is not offended here. Each of us is served by the enforcement of the law—not least a person like the defendant in this case, whose wealth and privileges, so long enjoyed, are so much founded upon law. More broadly, we are driven regularly in our ultimate interests as members of the community to use ourselves and each other, in war and peace, for social ends.

A retributivist might reply that a person must not be used merely as a means even if it was to benefit that person. As Kant argued, a person must not even use himself as a mere means. However, the distinction between using a person as a mere means to his own benefit and treating a person as an

---

447 Holmes, supra note 88, at 44.
448 Id. at 43. Holmes continues by giving examples of the ways in which society uses its members as mere means:

If conscripts are necessary for its army, it seizes them, and marches them, with bayonets in their rear, to death. It runs highways and railroads through old family places in spite of the owner's protest, paying in this instance market value, to be sure, because no civilized government sacrifices the citizen more than it can help, but still sacrificing his will and his welfare to that of the rest.

Id. Holmes concludes that using offenders as mere means is "perfectly proper." Id. at 47.

449 For general discussion as to the unpersuasiveness of this claim see Duff, supra note 1, at 188–92.
450 Duff, supra note 6, at 10.
451 Dolinko, supra note 2, at 1631–32. For a possible objection to this claim see id. at 1632.
453 Id. at 499.
454 Kant, supra note 22, at 255 (emphasis added) ("[A] man cannot be used merely as a means by any man (either by others or even by himself) but must always be used at the same time as an end.").
end may be elusive.

Consequentialists also claim that consequentialism does not use offenders as mere means because it treats offenders as ends by taking into account their interests. Bentham understood punishment as an evil because of the pain it caused those punished. In determining whether punishment is justified, the pain caused to the person punished is taken into account; if the good consequences of punishment do not outweigh the pain caused to the person punished, the punishment is not justified. The offender is thereby treated as an end.

Fourth, consequentialists might argue that consequentialism not only treats offenders as ends, but does so to a greater extent than retributivism. David Dolinko suggests that it is consequentialism, not retributivism, which truly respects the dignity of the offender. Rather than maintain, as retributivists do, that an offender has a right to be punished (a right that the offender will gladly waive) and that punishment is a “noble and uplifting enterprise that attests to the richness and depth of our moral character,” a deterrence-based approach respects the offender by presenting a more straightforward view of punishment as a “dirty business” and a “sad necessity.” “[I]t is the deterrence theorists . . . who truly ‘respect’ the criminal by acknowledging that inflicting pain on him is, in itself, bad, and not to be done unless it can be outweighed by its good consequences.”

Fifth, it might be conceded that consequentialism uses offenders solely as mere means, but the criticism might be defused by arguing that retributivism also uses offenders solely as mere means. Dolinko argues that by punishing offenders, retributivism uses them as mere means in differing ways. Retributivism uses offenders as mere means by implementing punishment institutions that invariably, mistakenly punish innocents. The innocents mistakenly punished are used as mere means or sacrificed in order that we may maintain our fallible punishment institutions and punish those who are actually guilty. Retributivists could, of course, invoke some variant on the doctrine of double effect and reply that no particular, identifiable innocent is being used as a mere means intentionally. The validity of this reply has been extensively addressed in Part IV. In addition, retribu-

---

455 See supra note 68.
456 See supra notes 69–71 and accompanying text.
457 S.I. Benn, An Approach to the Problems of Punishment, 33 PHILOSOPHY 325, 329 (1958) (“To look to the consequences does not entail treating the criminal merely as a means to a social end, as critics have asserted; for in weighing advantages and disadvantages, the criminal, too, must ‘count for one.’”).
458 See supra text accompanying note 115.
459 Dolinko, supra note 2, at 1656 (attributing this view to retributivists).
460 Id.
461 Id.
462 Id. at 1632–33. For a similar claim, see Schedler, supra note 296, at 189.
463 See supra notes 305–308 and accompanying text.
Deterring Retributivism

tivism uses offenders as mere means by justifying the imposition of punish-
ishment despite being unable to know whether the actual amount of punish-
ishment imposed is what the offender actually deserves. Since retributivism’s treatment of offenders as ends depends on giving offenders what they deserve, the inevitable failure of our punishment institutions to determine the correct amount of deserved punishment leads to most offend-
ers not being treated as ends. Again, the retributivist response might be to invoke some variant of the doctrine of double effect: that the amount of punishment imposed is with the intention that it reflects an offender’s just deserts.

Sixth, not only is it permissible to use offenders as mere means, but the principles of retributivism might affirmatively require it. Intriguingly, a leading retributivist, Robert Nozick, seems to suggest that retributivism properly should justify the intentional use of offenders as mere means. Nozick maintains that “to be used as a means may be part of his [an offender’s] retributive matching desert, since that is what he has done to another.” By “matching desert,” Nozick is referring to retributive punishment’s attempt to turn back the crime on to the criminal or “bringing home to the offender the nature of what he has done.” In other words, under the lex talionis, retributivism seeks to match what the offender did to his victim by doing the same to the offender. Since the offender has used his victim as a mere means by which the offender gained the fruits of his crime, the offender’s matching desert also entails being used as a mere means. Thus, intentionally using an offender as a mere means in punishing him might not only be unobjectionable under retributivism, but affirmatively required. Since under this view intentionally using an offender as a mere means is justified and required by retributivism, retributivists can hardly object that consequentialist theories also justify using offenders as mere means.

B. Retributivism Uses Crime Victims as Mere Means

The previous subpart considered the retributivist charge that consequentialist theories immorally use as mere means the offenders they punish. This subpart will make the novel and seemingly radical argument that retributivism intentionally uses crime victims as mere means in order to treat their victimizers (that is, the criminals) as ends in themselves. Although this claim might seem preposterous at the outset, other commentators have made related claims. Louis Kaplow and Steven Shavell note that it is

[464] Dolinko, supra note 2, at 1635.
[465] See supra notes 400, 411 and accompanying text.
[466] See supra notes 264-269 and accompanying text.
[467] NOZICK, supra note 104, at 372.
[468] Id.
[469] See supra note 79.
criminals who use society as a mere means. And Jeffrie Murphy hints that victims might be used by the state for its own ends.

Before attempting to defend this radical claim, let us briefly set out the steps of the argument: 1) Victim-relative norm violation is necessary for the imposition of retributive punishment and for retributivism to treat culpable wrongdoers as ends; 2) The interests of crime victims as ends are irrelevant in retributivism; 3) Therefore, the necessary use of victims' being victimized constitutes using the victims merely as means in order to treat culpable wrongdoers as ends and to attain appropriate retributive punishment. Steps 1 and 2 should be fairly uncontroversial. What is critical is step 3. That is, whether retributivism's failure to treat crime victims as ends, despite the necessary use of victims' being victimized, constitutes using victims merely as means.

1. The Necessary Use of Victims as Means.—As to step 1, I do not mean to imply that victim-relative norm violation is only necessary for retributive punishment but not for other, including consequentialist, theories of punishment. Regardless of whether it is true with respect to other theo-

470 Kaplow & Shavell briefly suggest that, under a retributivist system, criminals use society, which must bear the various costs of crime and maintenance of the administration of criminal justice, as mere means:

[T]he main effect of the retributive view is to preserve the profitability of crime to some potential criminals at a greater expense to their victims and to all who must finance the costs of punishment. It is not clear how this treatment respects the dictum not to use people merely as a means to others' ends, for it could be said that everyone else in society is thus used by criminals when fair [retributive] punishment is insisted on.

Kaplow & Shavell, supra note 28, at 1265–66 (footnote omitted). That is, criminals use society as a mere means to a greater extent under retributivism than consequentialism because the former's lack of emphasis on deterrence leads to a greater incidence of crime.

In the corresponding footnote to the above quotation, Kaplow & Shavell note that because criminals voluntarily choose to commit crime, they may waive their rights (against being used as mere means). Id. at 1266 n.742. While criminals choose to commit crime, crime victims do not choose to be victimized, and thus do not waive any rights (against being used as mere means). As a result, "there seems to be a stronger sense in which, from this Kantian perspective, the victims are being used as mere means, independently of the course they would will for themselves." Id.

This Part will argue that it is not criminals that use society as a mere means, as Kaplow and Shavell suggest, but rather that retributivism uses crime victims as mere means.

471 Murphy ventures that victims' interests might be respected less than that of their victimizers and less than that of the state:

[V]ictims often feel that their particular injuries are ignored while the [criminal justice] system addresses itself to some abstract injury to the state or to the rule of law itself—a focus that appears to result in wrongdoers being treated with much greater solicitation and respect than their victims receive. If the actual victims are noticed at all (other than to alert the state to a violation of its interests) . . . .


472 The view that step 2, by itself, is sufficient to establish that retributivism uses crime victims as mere means, and not even partially as ends in themselves, will be considered infra section V.C.4, particularly infra text accompanying notes 546–553.
ries, I merely mean in step 1 that it is true for retributive theories of punishment. Though not all criminal offenses involve identifiable victims, surely the vast majority defines the prohibited conduct with respect to a victim. Imagine the difficulty of drafting an offense of murder, robbery, or assault without reference to a human victim. Let us now consider some specific ways in which victims are used in our penal system.

We often hear the claim that such and such an offense is a “victimless crime.” And so it is claimed that the offense deserves either no punishment or less punishment. By implication, a crime that did harm a victim deserves either some or more punishment. Either way, crime victims are necessarily used as a means in order to gauge the proper level of deserved punishment. For Hegel, “punishment should affect the offender as much as his offense has affected the victim.” Thus, the degree to which the victim is affected by the crime is used as a means to gauge the degree of desert of the offender and the amount of punishment to be imposed under retributivism.

A victim’s consent is used not only to delineate legal from illegal conduct, but also to differentiate among various forms of illegal conduct. Whether a victim consents to the offender’s conduct in many cases will be dispositive as to whether conduct is criminal or not. Examples include rape, larceny, and battery. The difference between blackmail and a routine business transaction may depend on whether the victim or the offender initiated the transaction. The difference between the crimes of embezzlement and larceny depends on whether the victim initially consented to entrust the property to the offender.

Bias crimes depend on features of their victims. An offense committed against victims with certain religious, ethnic, racial, or sexual features, with the offender’s motive to commit the offense based on one of these features, transforms the underlying offense into an elevated offense carrying a greater level of deserved punishment.

---

473 MPC, supra note 247, at § 210.1(1) (Criminal Homicide) (“A person is guilty of criminal homicide if he purposely, knowingly, recklessly or negligently causes the death of another human being.” (emphasis added)).

474 Id. at § 222.1 (“1) Robbery Defined. A person is guilty of robbery if, in the course of committing a theft, he: (a) inflicts serious bodily injury upon another; or (b) threatens another with or purposely puts him in fear of immediate serious bodily injury . . . .” (emphasis added)).

475 Id. at § 211.1(1) (“Simple Assault: A person is guilty of assault if he: attempts to cause or purposely, knowingly or recklessly causes bodily injury to another; or negligently causes bodily injury to another with a deadly weapon; or attempts by physical menace to put another in fear of imminent serious bodily injury.” (emphasis added)).

476 PRIMORATZ, supra note 1, at 80–81 (construing Hegel’s view that punishment must be comparable in value and character to the crime).


478 DRESSLER, supra note 222, at 524; FLETCHER, supra note 1, at 7.

479 See generally Anthony M. Dillot, Punishing Bias: An Examination of the Theoretical Founda-
In virtually all jurisdictions, an offender who completes or consummates the attempted harm against a victim will deserve greater punishment than one who fails to complete the attempted harm.\(^{481}\) The amount of deserved punishment is a function of the amount of harm sustained by the victim.

In response to these examples of retributivism using crime victims as means in order to give offenders their just deserts, a retributivist might argue that criminal offenses are conceptualized as crimes against the state and not against a human victim. It is tort law, not criminal law, where individual human victims seek redress under the law. The victims of criminal offenses are, and crimes are committed against, the state or community or society.\(^ {482}\) Nonetheless, such crimes against the state are defined with reference to a human victim. Again, imagine the difficulty of drafting an offense of rape, murder, or battery without reference to a human victim. Though such crimes may theoretically be committed against the state, they necessarily may only be committed on a human victim. For the vast majority of criminal offenses, it is necessary to use individual victims in order to formulate the offense.

Despite the necessity of using individual victims as means to define most criminal offenses, retributivists might still maintain it is the state that is formally considered the victim against which crime is committed. But if so, by considering the interests of the state, it is the state (and not the individual victim) that is treated as an end. Moreover, the very conception of the state as the official victim of a crime only serves to underscore that individual human crime victims are used as means so the state and the offenders may be treated as ends.

The only point here, from all these examples, is that human victims are essential for, and integral to, the imposition of retributive punishment. As one prominent retributivist, Michael Moore, recognized, "[v]ictims have a lot to do with the justification of punishment for a retributivist."\(^ {483}\) Moore

\(^{481}\) Moore, supra note 26, at 69.

\(^{482}\) DRESSLER, supra note 222, at 1; Fletcher, supra note 1, at 616. Contra MPC, supra note 247, at § 5.05.

\(^{483}\) Moore, supra note 26, at 69.
explains that human victims are central to retributivism because the desert of an offender is based on culpable wrongdoing; wrongdoing is not merely an act with bad consequences but also must be a violation of a norm.\textsuperscript{484} “Any kind of retributivist needs a norm violation to justify punishment. It is at this point that victims come in substantively. Victims come in as part of the content of those norms.”\textsuperscript{485} Referring to the most important norms prohibiting murder, torture, rape, and the like, Moore declares that “[v]ictims are important substantively because each of these serious moral norms prohibits actions causing harms to certain victims.”\textsuperscript{486} Because of the importance of these “victim-relative” norms,\textsuperscript{487} victims “are at the center of the norms whose violation is at the core of the criminal law.”\textsuperscript{488} Victims have a “large place”\textsuperscript{489} in retributive punishment and “are central to the norms whose violation justifies punishment for a retributivist.”\textsuperscript{490} Therefore, retributivism necessarily uses crime victims as the means to attain retributive punishment and the treatment of offenders as ends in themselves.

2. The Irrelevance of the Interests of Victims as Ends.—Like the first step of the argument that retributivism uses victims as mere means, the second step—that the interests of crime victims as ends in themselves are irrelevant to retributivism—should also be fairly uncontroversial. Because retributivism is said to focus exclusively on what offenders deserve (that is, offenders’ just deserts), consideration of the interests of victims is irrelevant. Moore proclaims the interests of crime victims to be extraneous in the enterprise of retributive punishment:

Retributivism is not the view that punishment of offenders satisfies the desires for vengeance of their victims. In this view the harm that is punishment is justified by the good it does psychologically to the victims of crime, whose suffering is thought to have a special claim on the structuring of the criminal justice system. This is not retributivism. A retributivist can justify punishment as deserved even if the criminal’s victims are indifferent (or even opposed) to punishing the one who hurt them. Indeed, a retributivist should urge punishment on all offenders who deserve it, even if no victims wanted it.\textsuperscript{491}

Although retributivism uses victims to define the core prohibitions of the criminal law (victim-relative norm violations) and to gauge the degree of offenders’ desert, the interests of victims are irrelevant. As Moore plainly states, “I think victims should and must be ignored if you are claim-

\textsuperscript{484} Id.
\textsuperscript{485} Id.
\textsuperscript{486} Id. at 69–70.
\textsuperscript{487} Id. at 71.
\textsuperscript{488} Id. at 72.
\textsuperscript{489} Id.
\textsuperscript{490} Id. at 73.
\textsuperscript{491} MOORE, supra note 73, at 89 (footnote omitted).
The victim’s interest in if, or how much, an offender should be punished is simply irrelevant to, and inconsistent with, retributive punishment. As Moore explains:

[V]ictims have no say when we leave the general punishment grading scheme set up by the democratic legislature for a court’s application of that scheme to a particular offender. What the victim wants by way of amount or kind of punishment, whether a certain process makes the victim happy or sad, is simply irrelevant to how a court should proceed in a criminal case.

Regardless of whether accommodating the interests of victims involves victims seeking increased punishment or leniency for the offender, accommodating such interests is irrelevant and inconsistent with retributive punishment. Taking into account the interests of victims who want to see their victimizers suffer is not retributive justice. According to Moore, that is "corrective justice and not retributive justice." Implementing such a scheme of victim vengeance, "doesn’t look retributive; it looks compensatory to the victim. Punishment in such a scheme turns on the victims deciding what they want, not on what justice demands." Similarly for Moore, accommodating victims’ interest in leniency is incompatible with retributivism. As Moore proclaims, “The propensity of a victim to forgive her transgressor is irrelevant to retributive desert.” Moore further adds that “retributive justice demands that culpable wrongdoers suffer, irrespective of whether or not those they wrong wish it.” As Richard Murphy notes, “With retributivism, any considerations of the victim or of society, at least after the crime has been committed, are irrelevant.” According to Moore, “doing justice is the essence of retributive punishment and . . . victims have neither any moral right nor expertise to say how our legal institutions should achieve such justice.”

Even theorists advocating that the interests of victims should be taken into account in formulating a theory of punishment have recognized that the resulting theory of punishment cannot be retributive. Though previously

492 Moore, supra note 26, at 67.
493 Id. at 75.
494 Id. at 75–76.
495 Id. at 76.
496 Id. at 77.
497 Id. at 78.
499 Moore, supra note 26, at 89.
500 Id.
501 Moore cites this as further support for his position that taking into account the interests of victims is incompatible with retributivism. Id. at 66–67, 75–76, 76 n.23. But see George Fletcher, The Place of Victims in the Theory of Retribution, 3 BUFF. CRIM. L. REV. 51, 63 (1999) (advocating that acceding to victims’ interests as to the punishment of the offender and giving the victim a special role at
Deterring Retributivism

a retributivist, Jeffrie Murphy acknowledges that in advocating for the inclusion of victim preference as to whether and to what extent an offender should be punished, he is no longer a retributivist. Randy Barnett has proposed an alternative theory of punishment based on restitution to the victim. Barnett acknowledges that such a scheme would not be compatible with retributivism but would instead replace it.

In short, there is purportedly no conceptual place within retributive punishment to accommodate the interests of victims, to treat victims as ends in themselves.

3. The Use of Victims as Mere Means.—Demonstrating the conclusion in step 3, the critical step in the argument that retributivism uses victims as mere means, will be comparatively more difficult. Steps 1 and 2 are presumably proven to satisfaction because retributivists themselves expressly make the arguments comprising those steps. Though I will argue that step 3—retributivism treats victims as mere means—follows from steps 1 and 2, I will be unable to use retributivists’ express arguments to establish it. Naturally, given retributivism’s reliance on Kant’s maxim as a principal criticism of consequentialist theories of punishment, retributivists would be loath to expressly admit to my conclusion in step 3. I will argue that retributivism’s express agreement with steps 1 and 2 entails retributivism’s (however reluctant) acceptance of the conclusion of step 3. If successful, the argument will demonstrate that one of retributivism’s principal arguments against consequential theories of punishment applies to itself.

To summarize the argument thus far, victims are used as the necessary means to determine the desert and punishment of an offender (step 1), but their interests as ends in themselves are ignored, are irrelevant, and are inconsistent with determining the desert and punishment of an offender (step 2). Combining steps 1 and 2, retributive punishment uses victims as means in determining whether and to what extent an offender should be punished without treating the victims as ends in themselves. If retributivism uses victims as means and fails to treat them as ends, then retributivism uses victims as mere means in order to determine if and to what extent offenders should be punished. As a result, retributivism uses victims as mere means, in violation of Kant’s maxim, in order to treat their victimizers as ends in

---

502 Murphy, supra note 471, at 221–24.
505 The argument that retributivism uses crime victims as mere means could conceivably apply to some consequentialist theories as well. But there is significantly greater conceptual room within consequentialism to take into account, as one of the many good consequences to be promoted by punishment, the satisfaction of victims’ interests. My main point, however, is not so much to defend consequentialism as it is to demonstrate that the retributivist criticism—that consequentialism uses persons as mere means—applies equally to retributivism.
themselves.

Kant’s categorical maxim to always treat persons as ends in themselves rather than mere means is not only one of retributivism’s principal justifications but also a principal criticism of consequential theories of punishment. That retributivism may violate Kant’s maxim therefore undercuts retributivism’s justification as well as demonstrates that retributivism fares no better than consequential theories. Since treating persons as ends in themselves is a purported feature of retributivist, but not of consequentialist, theories, the criticism’s application to retributivism is even more damaging.

If the above argument is true, then retributivism treats victimizers as ends but victims as mere means whereas the converse may hold with respect to consequentialist theories: they treat victimizers as mere means and victims as ends. Perhaps from the perspective of Kantian morality it is just as wrong to treat victimizers as mere means as it is to treat victims as mere means. Nonetheless, our intuitions may inform us that treating victims as mere means, as retributivism does, is worse than treating victimizers as mere means (as consequentialism arguably does). Arguably, if faced with a choice between using one or the other as mere means it would be preferable to treat culpable wrongdoers rather than innocent victims as mere means. It seems morally worse to sacrifice the interests of innocent victims in order to respect culpable wrongdoers as ends than to sacrifice the interests of culpable wrongdoers in order to respect innocent victims. Otherwise, victims of crime are being victimized twice—once by the culpable wrongdoers and then again by retributivism.

4. Can Retributivism Avoid the Claim That It Uses Victims as Mere Means?.—How would retributivism avoid the conclusion, in step 3, that it uses victims as mere means? There are two obvious ways, but both are problematic. If either the premise in step 1 or the premise in step 2 did not apply to retributivism, then, of course, the conclusion in step 3 would not follow. Let us first consider the elimination of victim-relative norms from retributivism (step 1). As Moore argued above, the core criminal prohibitions involve the violations of victim-relative norms. Victims are “central,” “are at the center,” and “have a large place” in the formulation of such norms. As Moore states, “[v]ictims … are important to retributive justice because the desert (that triggers a just punishment) is a function of the violations of the rights of victims by offenders.”

506 See supra text accompanying notes 407–408.
507 See supra notes 18, 414 and accompanying text.
508 See supra notes 397–398 and accompanying text.
509 A deterrence-based consequentialist theory might be said to treat victims as ends by punishing offenders in order to reduce the incidence of their, and others’, future crime victimization.
510 See supra text accompanying notes 488–490.
511 Moore, supra note 26, at 89.
Deterring Retributivism

ing a moral norm or drafting a criminal statute prohibiting murder or rape without making reference to a human victim.

Moore gives a specific example in which victims are used to determine the proper amount of retributive punishment.512 Suppose a culpable wrongdoer detonates one bomb that kills five people. Is the culpable wrongdoer to be punished for one homicide or five? Even though the culpable wrongdoer only committed one act, he will be punished for five homicides. Moore argues that this result only makes sense by reference to the centrality of victims in our victim-relative norms.513 In other words, victims are used as a device or means to determine the desert of an offender who with one act causes multiple homicides. Without using victims, how would we be able to make sense of the determination that the offender's desert includes five homicides rather than one? Moreover, without using victims it would be difficult to punish for homicides at all.

In light of the centrality of victim-relative norms to retributive punishment, the elimination of step 1 does not seem to be a promising avenue for retributivism to avoid the criticism that it uses crime victims as mere means. Though it would no longer be using victims as mere means (by foregoing victim-relative norms), retributivism would be unable to properly determine the appropriate desert for violations of victim-relative norms. Moreover, the very formulation of norms that comprise the core of criminal law, victim-relative norms, would be jeopardized if not impossible.

Let us consider the elimination of step 2. According to the arguments of Moore and Murphy above, accommodating the interests of victims as to if, and how much, an offender should be punished is inconsistent with retributivism. Therefore, by accommodating the interests of victims, one would no longer be doing retributive punishment. As a result, it seems that retributivism can avoid the criticism that it uses crime victims as mere means but at the cost of no longer having a retributive theory of punishment. For a retributivist then, the cure would be worse than the malady. Therefore, elimination of step 2 is likewise an unpromising avenue for retributivism to avoid the conclusion in step 3.

It would seem that retributivists must look elsewhere to avoid the claim that retributivism uses crime victims as mere means. Let us now consider some arguments that retributivists might make to resist the conclusion in step 3.

C. Objections

Since the claim that retributivism treats crime victims as mere means is quite radical, let us try to anticipate some possible objections. Four possible objections to the argument that retributivism violates the Kantian

512 Id. at 71.
513 Id. at 71–72.
maxim will be presented and countered. First, Kant’s injunction against using persons as mere means does not, and should not, apply to crime victims. Second, though retributivism may use crime victims as mere means, it does not do so intentionally. Third, though retributivism may use crime victims as means, it does not use them as mere means. Fourth, retributivism does not use crime victims in any way at all.

1. Kant’s Maxim Does Not, and Should Not, Apply to Retributivism’s Use of Crime Victims.—One could object that although retributivism might use crime victims as mere means, Kant’s injunction does not (and should not) apply to a punishment authority’s use of crime victims in punishing offenders. In other words, while Kant’s maxim applies to a punishment system’s treatment of offenders, who are directly involved, the victims of crime are too remote to the process and only peripherally involved.

The objection, however, is unpersuasive. Kant does not limit the application of his maxim to certain persons. Those not to be used as mere means include all of humanity. As Kant, in perhaps the most famous formulation of the maxim, declares: “Act so that you treat humanity, whether in your own person or that of another, always as an end and never as a means only.” Therefore, by Kant’s plain language, crime victims, as part of humanity, would seem to be included among those who must not be used as mere means. But perhaps a retributivist might argue that Kant’s plain language should not be dispositive.

As discussed above, the scope of Kant’s maxim might be interpreted more narrowly. O’Neill allows that Kant’s maxim might not apply quite so strictly with respect to “those who will in any case be no more than spectators.” But according to O’Neill, Kant’s maxim does apply to those who are “closely involved in or affected by” a course of action. Under this view, the objection might gain some traction if crime victims might be conceived of as “spectators” to the punishment of their victimizers rather than “closely involved in or affected by” the punishment. While the degree to which crime victims are affected by the punishment of their victimizers (or lack thereof) will surely vary, clearly there are many crime victims who are greatly affected by the quantity and quality of the punishment that the victimizer receives. These victims, even under this narrower interpre-

---

514 KANT, supra note 407, at 46 (emphasis added).
515 See supra text accompanying notes 437–438.
516 O’NEILL, supra note 423, at 110.
517 See, e.g., Jim Yardley, In Oklahoma City, Delay Brings Even More Anguish, N.Y. TIMES, May 12, 2001, at A12 (describing the “anguish,” “frustration,” and “anger” felt by survivors and victims of the Oklahoma City federal building bombing over the delay in Timothy McVeigh’s execution).
518 See generally GEORGE FLETCHER, WITH JUSTICE FOR SOME: VICTIMS’ RIGHTS IN CRIMINAL TRIALS (1995) (discussing the variety of ways that victims are affected by the trials and punishments of their victimizers). See particularly id. at 194, discussing how particular victims would be greatly affected by a more active role at trial.
tation of the scope of Kant's maxim, would clearly be included among the
group of persons that a system of punishment should not use as mere
means.

2. Retributivism Does Not Intentionally Use Crime Victims as Mere
Means.—Retributivism's use of victims as mere means might be insulated
from criticism by invocation of some variant on the doctrine of double ef-
fect.520 Perhaps, retributivists might argue, even though it is foreseen that
some unknown victims will be used as mere means, that they are so used is
not intentional. Unlike the claim that retributivism only unintentionally
uses offenders as mere means,521 retributivists cannot plausibly claim that
the theory's use of crime victims as mere means in order to treat offenders
as ends is unintentional. Retributivists expressly claim that, after the com-
mission of a crime, the interests, preferences, and concerns of the crime's
victim are to be ignored and are irrelevant.522 Moreover, to give any say to
victims as to the punishment of the offender is inconsistent, and incompati-
ble, with retributivism.523 Given these explicit rejections of a conceptual
space within retributivism for victims as ends in themselves, retributivists
could hardly claim it to be unintentional.

519 To provide an admittedly atypical example of how a victim may be affected by the punishment
of her victimizer consider the following hypothetical based on a Law & Order television episode: Sup-
pose that a borderline mentally retarded 18-year old girl is in a state-sponsored program to "mainstream"
herself into the general population of a high school. The rationale behind mainstreaming her is that if she is
placed into an environment where she is expected to be "normal" it will foster and encourage her to be-
come more "normal." As part of the mainstreaming process, other students are not informed of her con-
dition. In the high school, she has sexual relations with a nonretarded 18-year old boy from the school.
School administrators find out, report it to the police and the boy is put on trial for rape. The girl testi-
fies that she liked the boy, that she did want to have sexual relations with him, that she enjoyed it, and
that she would like to have sexual relations with him again. The prosecuting attorney nonetheless
claims that the girl's mental retardation precluded her ability to consent as a matter of law. The boy
eventually admits that he realized that she was not entirely "normal" and thus was at least reckless as to
whether she could properly consent. In a statement to the jury, the girl pleads with the jury not to con-
vict the defendant, not out of concern for him, but for herself. She says that she understands that she is
considered to be partially mentally retarded, but that in her own mind's eye she considers herself to be
"normal." If the boy is convicted and punished she understands that in the view of the community she
would be considered abnormal and a freak. It would be a violation of her image of herself to be judged
lacking the capacity to consent. Nonetheless, the boy is convicted and punished. See Law & Order:
Damaged (NBC television broadcast, May 6, 1998).

Though, with respect to the boy, justice may well have been done, are we so sure that justice was
done with respect to the girl? Is the girl, the putative victim of the crime, treated with respect and digni-
ity? Are her interests vindicated? Is she treated as an end-in-herself? Or is she used as a mere means
in order for the defendant to be given his just deserts?

520 See supra notes 305–308 and accompanying text.

521 See supra text accompanying notes 462–466 for a possible retributivist response to Dolinko's
claim that retributivism uses offenders as mere means by maintaining necessarily fallible punishment
institutions.

522 See supra text accompanying notes 491–493.

523 See supra text accompanying note 500.
3. **Retributivism Uses Crime Victims as Means But Not as Mere Means.**—Though retributivism uses victims as means because of the necessity of victim-relative norms, and does not treat victims as ends because it ignores their interests, nonetheless it might be argued that retributivism does not use victims as *mere* means. That is, retributivism treats victims, in part, as ends. Retributivism might be said to treat victims as ends in view of some of the metaphoric justifications of retributivism. Retributivism is said to express solidarity with victims,\(^{524}\) restore a relation of equality between the victim and offender,\(^{525}\) counteract a dominance-submission relationship between offender and victim,\(^{526}\) and free a victim's blood or life spirit from the control of the offender.\(^{527}\) All of this rhetoric, it might be argued, suggests that retributivism is very much treating victims as ends.

Nevertheless, the objection that retributivism treats victims, at least in part, as ends is unpersuasive. First, as Moore notes, this metaphoric rhetoric embodies a slide from retributivism into consequentialism.\(^{528}\) For retributivism to be justified by counteracting dominance or restoring equality or vindicating the victim in some way suggests that it is attempting to promote some future state of affairs or good consequences. Punishment of deserving wrongdoers is thereby not intrinsically good but rather an instrumental good by which the victim may be vindicated. It might be replied that the good consequences to be promoted (vindication of the victim) are not factually or contingently consequential but only conceptually consequential.\(^{529}\) As such, vindicating the victim is a conceptual consequence of retributive punishment but does not constitute a slide into the factual or contingent consequentialism of deterrence theories of punishment. Whether or not the promotion of merely conceptual consequences fails to slide into consequentialism will be discussed in Part VI.

But even if we accept that justifying punishment by its promotion of conceptual consequences does not slide into consequentialism, the very fact that the vindication of victims is only conceptual or abstract underscores that victims are not treated as ends in themselves. Some abstract concept is promoted, rather than the actual interests of flesh and blood victims.\(^{530}\) The conceptual consequence, and not the actual victim, is treated as an end. As a result, retributivism fails to treat the victim as an end.

Second, if the claim is that retributivism treats victims as ends, despite

---

\(^{524}\) Fletcher, *supra* note 518, at 201–05.

\(^{525}\) Hegel, *supra* note 22, at 124 (§ 99).


\(^{528}\) Moore, *supra* note 26, at 82–83.

\(^{529}\) See infra text accompanying notes 577–578.

\(^{530}\) See *supra* note 471.
ignoring their interests, then retributivism is subject to the same criticism retributivists have leveled against rehabilitative theories of punishment. Retributivists have criticized rehabilitative approaches as paternalistic, demeaning, and disrespectful of offenders by giving offenders not what they deserve but rather the therapeutic treatment that they need. By implementing treatment on the offender's behalf but against the offender's preferences, rehabilitative punishment makes the offender not a rational subject or moral person, but rather an object. Rehabilitative punishment theories, retributivists maintain, treat offenders as less than rational, choosing beings. Herbert Morris, in an influential article, explains what is entailed by the rehabilitative theory treating offenders as things or animals:

When we treat a human being merely as an animal or some inanimate object our responses to the human being are determined, not by his choices, but ours in disregard of or with indifference to his. And when we "look upon" a person as less than a person or not a person, we consider the person as incapable of rational choice.

In contrast, Morris describes how retributivism treats offenders as persons or ends: "We treat a human being as a person provided, first, we permit the person to make the choices that will determine what happens to him and, second, when our responses to the person are responses respecting the person's choices."

If retributivism may be said to treat victims as ends (despite ignoring, and finding irrelevant, their interests) by punishing offenders on their behalf, retributivism is treating victims in the same way—paternalistically and as less than persons—that retributivists criticize rehabilitative theories for treating offenders. Suppose a victim does not want her victimizer punished at all or wants leniency or is opposed to the type of punishment the offender will receive, for example, capital punishment or chemical castration. Suppose further that a system of retributive punishment nonetheless determines that the offender shall be punished against the victim’s wishes. If it is still maintained that the punishment is on behalf of the victim or to vindicate the victim’s interests, then is this not paternalism? Is this not a claim

---

532 See MURPHY, supra note 113, at 134–35.
533 MORGAN, supra note 6, at 46.
534 Id. at 48–49.
535 See, e.g., Sara Rimer, Victims Not of One Voice on Execution of McVeigh, N.Y. TIMES, Apr. 25, 2001, at A1 (noting that some victims and survivors of Timothy McVeigh’s 1995 bombing of the federal building in Oklahoma City oppose McVeigh receiving the death penalty); Candidate Who Murdered Opponent Gets Life, No Parole, USA TODAY, Aug. 24, 2000, at A3 (reporting that family members of Tennessee State Senator Tommy Burks, who was murdered by his opponent in the 1998 senate election, asked prosecutors not to seek the death penalty).
that the victim is an object to be vindicated, but not a rational human being whose choices are to be respected? Is not the victim being treated with, as Morris puts it, "disregard of or with indifference to" her choices?

Consider also Kant's suggestion that only young children and the insane may be treated as ends or persons when we act in violation of their expressed interests: "I cannot do good to anyone in accordance with my concepts of happiness (except to young children and the insane), thinking to benefit him by forcing a gift upon him; rather, I can benefit him only in accordance with his concepts of happiness." The Kantian principles of beneficence and respect for others as persons require that we recognize and share the ends of others as their ends and not paternalistically take them over as our ends, as we see fit. In other words, retributivism cannot persuasively claim to be treating victims as ends in themselves by claiming to promote aims or ends that the victims themselves do not share.

The claim that retributivism treats victims as ends by punishing offenders to vindicate the victims in some way, despite ignoring and finding irrelevant the victims' interests, subjects retributivism to the same criticism retributivists have made against rehabilitative theories of punishment. Retributivists have found the criticism to be a sufficient basis to reject the rehabilitative theory of punishment. Unless retributivists wished the same criticism to be applied to retributivism, presumably they would not wish to make the claim that retributivism treats victims as ends (while ignoring their interests). As a result, retributivism may be forced to concede that it fails to treat victims as ends.

To summarize, retributivists might argue that the various metaphorical justifications of retributivism, amounting to a vindication of the victim in some way or another, constitute treating crime victims, at least in part, as ends in themselves. Two responses demonstrate the failings of this argument. First, justifying punishment by some form of conceptual or abstract vindication of the victim only serves to underscore that the actual flesh-and-blood victims are not being treated as ends—abstract ends are being promoted, rather than the interests of actual victims. Second, retributivism's vindicating a victim against the victim's interests subjects retributivism to the charge of paternalistically treating crime victims as objects rather than rational subjects—the same criticism retributivists have directed at rehabilitative theories of punishment. Therefore, the claim of conceptual vindication of the victim fails to alter the conclusion that retributivism uses crime

---

536 Morris, supra note 6, at 46.
537 Kant, supra note 22, at 248.
538 O'Neill, supra note 423, at 115–16. As to the duty of respect, "[P]olicies of respect must recognize that other's maxims and projects are their maxims and projects. They must avoid merely taking over or achieving the aims of these maxims and projects, and must allow others the 'space' in which to pursue them for themselves." Id. at 115. As to the duty of beneficence, "it is from the start antipaternalistic. The duty to seek others' happiness is always a duty to promote and share others' ends without taking them over . . . . " Id. at 116.
victims as mere means.

4. Retributivism Does Not Use Crime Victims at All.—Retributivists might claim that retributivism, in its use of victim-relative norms, does not use victims themselves in any way at all, but instead uses facts about victims.539 If so, it might be argued, then retributivism is not using victims as mere means in violation of Kant's maxim. For example, the use of a victim's consent to define what is and what is not a crime is not using the victim but using a fact about the victim—that she did or did not consent. The definition of bias crimes based on features of a victim is not using the victim but using a fact about the victim—his religious, ethnic, racial, gender, or sexual orientation features. Using the amount of harm sustained by a victim to determine the amount of an offender's deserved punishment is not using the victim but using a fact about the victim—the amount of harm sustained by the victim. In all of these instances, it might be argued, retributivism is not using victims as mere means but merely using facts about victims to treat offenders as ends in themselves.

The objection is problematic for a number of reasons. First, while the means/ends distinction is notoriously elusive and ambiguous,540 distinguishing between the use of a person as a means and the use of a fact about that person as a means may be even less clear. What criteria are we to apply to determine whether what is being used as a mere means is a person or a fact about that person? Without this being sufficiently elucidated, it may be merely a matter of intuition as to whether a particular instance of "using as a mere means" falls into the category of person or fact about a person. Since it is retributivism that touts it does not use persons as mere means, and severely criticizes consequentialist theories of punishment for doing so, the burden would seem to be on retributivists to sufficiently define the distinction upon which the claimed superiority of their theory, and the claimed inferiority of consequentialist theories, depends. After all, since it is retributivists who have wielded the argument as a sword against consequentialist theories of punishment, retributivists cannot use the dullness of the sword as a shield now that its point is facing retributivism.

Second, the objection is belied by the express language of a leading retributivist, Michael Moore. Moore and others claim that retributivism depends on victims themselves, not on facts about victims. For example, it is victims, not facts about victims, who "are central to the norms whose violation justifies punishment for a retributivist;"541 it is victims who have a "large place" in retributive punishment;542 and it is victims "who are at the center of the norms whose violation is at the core of criminal law."543

539 I am indebted to Kent Greenawalt for suggesting this point.
540 See supra text accompanying notes 439-443.
541 Moore, supra note 26, at 73.
542 Id. at 72.
543 Id.
Though there are a number of other examples, consider finally the following statement by Moore: "Any kind of retributivist needs a norm violation to justify punishment. It is at this point that victims come in substantively. Victims come in as part of the content of those norms." In all of these examples, it is the use of victims, not facts about victims, that is necessary for retributivism.

It might still be claimed that one retributivist's framing of the issue in terms of victims does not preclude other retributivists from claiming that what is being used is facts about victims. While this is true, the cited language at least demonstrates the plausibility of the claim that it is victims, and not facts about victims, which are being used by retributivism.

Third, let us assume arguendo that retributivism does not use victims at all, either because retributivism instead uses facts about victims or for some other reason. Even if retributivism does not use crime victims, it nonetheless fails to treat crime victims as ends in themselves. To satisfy Kant's maxim it is not enough to avoid using persons as means; persons must also be affirmatively treated as ends. As Kant tells us, "a man cannot be used merely as a means by any man (either by others or even by himself) but must always be used at the same time as an end." The duty is, therefore, not only negative but positive as well. The negative component is satisfied by not using someone merely as a means; the positive component is satisfied by always treating someone, at least in part, as an end in herself. As Kantian scholar O'Neill, in interpreting Kant’s maxim, explains, "merely not to be used is not enough for being treated as a person."

Because after the commission of the crime, victims' interests, preferences, wants, needs, and concerns are "ignored" and treated as "irrelevant," retributivism fails to treat crime victims as ends in themselves. If victims' interests are ignored and irrelevant, then retributivism is entirely indifferent to the interests of victims. But treating others with indifference is, according to one conventional view of Kant's maxim, failing to treat them as persons or ends in themselves. Or as Kant himself explains, treating others with mere indifference is not enough to avoid using others as mere means: "it is not enough that he is not authorized to use either himself

---

544 See generally supra section V.B.1.
545 Moore, supra note 26, at 69 (emphasis added).
546 KANT, supra note 22, at 255 (emphasis added).
547 O'NEILL, supra note 423, at 113 ("[T]here are two separate aspects to treating others as persons: The maxim must not use them (negatively) as mere means, but must also (positively) treat them as ends-in-themselves." (citation omitted)).
548 Id. at 105.
549 Id. at 67.
550 Id. at 75; Murphy, supra note 498, at 1308; see SADURSKI, supra note 270, at 241.
551 See generally supra section V.B.3.
552 O'NEILL, supra note 423, at 106 (viewing this criterion as a relevant but not fundamental conception of failing to treat persons as ends in themselves).
or others merely as means (since he could then still be indifferent to them); it is in itself his duty to make man in general his end.\textsuperscript{553}

In other words, the validity of step 2 in our argument—that retributivism finds the interests of crime victims as ends to be irrelevant—is, by itself, sufficient to establish that retributivism runs afoul of Kant's maxim even if retributivism does not use crime victims in any way.

Fourth, let us assume arguendo that the distinction between using persons and using facts about persons as mere means is tenable and that retributivism only uses facts about victims, but not victims themselves, as mere means. If so, then consequentialists can use the same distinction to avoid the criticism by retributivists that deterrence-based theories of punishment use offenders as mere means to promote the reduction of crime.

Retributivists have criticized deterrence-based theories for not justifying actual punishment but merely the perception or appearance of punishment.\textsuperscript{554} This is because if an offender is punished but the public is unaware, the good of general deterrence is lost.\textsuperscript{555} In addition, the benefit of deterrence may be obtained by merely pretending to punish rather than by actually punishing an offender.\textsuperscript{556} Therefore, what does the real work of deterrence is not actual punishment but the publicity of punishing.\textsuperscript{557} Thus, a deterrence-based theory does not justify actual punishment but merely the perception of punishment.\textsuperscript{558} (Of course, to make the perception of punishment credible, offenders might, at least occasionally, have to be actually punished.\textsuperscript{559})

Let us assume arguendo that the above retributivist critique of the deterrence-based theory is valid. Could not consequentialists claim that it is not offenders themselves who are being used as mere means to attain deterrence but instead a fact about offenders, or rather, a perceived fact? Essentially, the fact, or perceived fact, that the offenders are being punished, and not the offenders themselves, is what deterrence-based theories use to attain the end of crime reduction. Thus, not only do deterrence-based consequentialist theories not use offenders as mere means, but they do not use them as means at all. In this way, consequentialism could avoid the retributivists' criticism that it uses offenders as mere means.

So, where does this leave us? Against the charge that retributivism uses crime victims as mere means to attain the end of retributive punishment, I have suggested that retributivism might employ a distinction between using persons as mere means and using facts about persons as mere means.

\textsuperscript{553} KANT, supra note 22, at 198.
\textsuperscript{554} See supra note 66 and accompanying text.
\textsuperscript{555} See supra note 63 and accompanying text.
\textsuperscript{556} See supra note 64 and accompanying text.
\textsuperscript{557} See supra note 65 and accompanying text.
\textsuperscript{558} Id.
\textsuperscript{559} See supra note 66.
means. Retributivists might claim that retributivism only uses facts about crime victims as mere means to attain the end of retributive punishment of deserving wrongdoers. Thus, it does not use crime victims themselves as mere means. But if so, the deterrence-based theory of punishment may employ the same distinction to escape criticism that it uses offenders as mere means to attain the end of deterrence. If the utilization of the distinction is valid, deterrence-based theories equally do not use offenders as mere means but only facts, or perceived facts, about offenders.

In order to escape the same criticism that retributivists have leveled against consequentialism—using persons as mere means—retributivism must introduce a distinction that not only absolves itself, but also consequentialism, from the criticism. The resulting situation is that either both retributivism and consequentialism use persons as mere means or neither theory uses persons as mere means. Either way, with respect to the criteria of using persons as mere means, retributivism fares no better than consequentialist theories of punishment.

In fact, retributivism might fare worse. After all, it is retributivism that has touted treating persons as ends as a virtue of its theory and pointed to consequentialism’s failure to do so as a defect of that approach. If using persons as mere means is a valid criticism of deterrence theories, it is still only an external criticism. But if retributivism uses persons as mere means, it is an internal criticism that contradicts one of the tenets of its theory. Moreover, if any class of persons should be treated as mere means, it should be victimizers and not victims. But retributivism gets it backward. It uses victims as mere means in order to treat victimizers as ends in themselves. Though consequentialism may use persons as mere means, at least it uses victimizers as mere means in order to treat (future) innocent crime victims as ends in themselves.

The ensuing “trilemma” for retributivism is that either a) it foregoes punishment of violations of the core of criminal law, that is, victim-relative norms, or b) it accommodates the interests of victims but ceases to be retributive, or c) it runs afoul of the Kantian maxim by treating crime victims as mere means, thereby violating one of retributivism’s central tenets.

D. A Brief Digression: The Tension Within Retributivism over Victims’ Rights

The existence of the trilemma perhaps explains why “card-carrying” Kantian deontologists like Jeffrie Murphy and George Fletcher feel the pull of victims’ rights and have taken, in Moore’s term, “the victim’s turn.” While Murphy has conceded that taking victims’ interests into account renders his account of punishment no longer retributive, Fletcher maintains that there is not only logical space within retributivism for victims’ interests

---

560 Moore, supra note 26, at 67.
but that a richer account of retributivism would include integrating the interests of victims and doing justice to the suffering of victims.\textsuperscript{561} Furthermore, Fletcher argues that victims should have a say in the plea-bargaining and trial stages.\textsuperscript{562}

The debate between the retributivists Moore and Fletcher\textsuperscript{563} reflects the tension within retributivism itself. On the one hand, incorporating the interests of the victim into a determination of the deserved punishment of the victimizer seems incompatible with retributivism. On the other hand, using one set of persons (crime victims)\textsuperscript{564} as mere means in order to treat another set of persons (offenders) as ends in themselves, if not incompatible with retributivism, renders retributivism subject to one of its principal criticisms of consequentialism.

Moore suggests a possible way out of the trilemma. Accommodating the interests of victims by taking into account their preferences as to if and how much an offender should be punished is not violative of retributivism if “victim preference is constitutive of justice . . . .”\textsuperscript{565} That is, rather than retributivism requiring the punishment of deserving wrongdoers according to their just deserts, a “rights-based retributivism” would give the victim the right to decide if and to what extent\textsuperscript{566} punishment should be imposed. Retributive justice would be satisfied, therefore, not when the offender is punished according to his desert but when the offender is punished (or not) according to the victim’s preferences.

Moore suggests three criticisms of this alternative form of retributivism. First, “rights-based retributivism” is no longer retributivism.\textsuperscript{567} It is corrective justice providing compensation to the victim.\textsuperscript{568} Second, it violates the principle of equality.\textsuperscript{569} Suppose culpable wrongdoers $A$ and $B$ commit the same wrongdoing with the same culpability against victims $C$ and $D$, respectively. Victim $C$ is forgiving and decides not to seek punishment of $A$, but victim $D$ seeks the maximum possible punishment of of-

\begin{itemize}
\item \textsuperscript{561} Fletcher, \textit{supra} note 501, at 55–63.
\item \textsuperscript{562} FLETCHER, \textit{supra} note 518, at 188–201.
\item \textsuperscript{563} Moore’s article, see \textit{supra} note 26, is a response to Fletcher’s article, see \textit{supra} note 501, as to whether consideration of victims’ rights is compatible with retributivism.
\item \textsuperscript{564} It might also be argued that retributivism sacrifices future crime victims in order to do justice to criminal offenders. Suppose that consequentialism does a better job of preventing crime than retributivism. Thus, while retributivism arguably does justice to the criminal, consequentialism does justice to future innocent crime victims (by decreasing their numbers). As C.L. Ten has asked of punishment under retributivism: “For whose benefit is punishment to be instituted? Surely not for the benefit of law-abiding citizens who run an increased risk of being victims of crime. Why should innocent people suffer for the sake of dispensing retributive justice?” C.L. Ten, \textit{supra} note 29, at 369.
\item \textsuperscript{565} Moore, \textit{supra} note 26, at 76.
\item \textsuperscript{566} Presumably, the victim’s right to decide the extent of punishment would be constrained by the limit of how much punishment “ordinary” retributivism would allow.
\item \textsuperscript{567} Moore, \textit{supra} note 26, at 75–77.
\item \textsuperscript{568} Id. at 75–76.
\item \textsuperscript{569} Id. at 77.
\end{itemize}
sider B. Though A and B seem equally deserving of punishment, A is not punished at all and B is punished severely. If retributivism necessarily involves desert-based punishment, and A and B have the same desert, then "rights-based retributivism" allowing differential punishment of A and B cannot be retributivism. Third, "rights-based retributivism" is simply wrong: "retributive justice demands that culpable wrongdoers suffer, irrespective of whether or not those they wrong wish it."570

If Moore is correct in these criticisms of "rights-based retributivism," then there does not appear to be a satisfactory way out of the trilemma for retributivism.

E. Conclusion

Perhaps the most fundamental criticism retributivists have leveled against consequential theories of punishment is that they violate Kant’s injunction against using persons (offenders) as mere means. Retributivists claim that unlike consequentialism, their own theory gives offenders what they deserve and thereby treats offenders as ends. But even if retributivism does treat offenders as ends, it only does so by using crime victims as mere means in violation of Kant’s maxim. Because victim-relative norm violations are necessary for the imposition of retributive punishment, and treating crime victims as ends in themselves is irrelevant to, and incompatible with, retributivism, that theory treats crime victims not as ends but as mere means. Thus, the retributivist critique of consequentialist theories of punishment applies to retributivism itself.

A number of possible retributivist objections were considered and rejected. First, it might be argued that Kant’s maxim does not (and should not) apply to retributivism’s use of crime victims as mere means. But Kant’s express language states that all humanity must be treated as an end and not a mere means. As part of humanity, crime victims clearly fall within the protection of Kant’s injunction.

Second, it might be argued that retributivism only unintentionally uses crime victims as mere means. But because of retributivists’ express claims of the necessary use of victim-relative norms, and their express claims as to the irrelevance of victims’ interests, retributivism intentionally uses crime victims as mere means.

Third, it might be conceded that retributivism uses victims as means but not as mere means. Retributivists might point to the various good abstract or conceptual consequences promoted by retributivism—amounting to a vindication of victims—reflecting retributivism’s treatment of victims as ends. That the end of punishment is some conceptual consequence, however, only serves to underscore that some abstract concept is being promoted rather than the actual interests of flesh and blood victims. In ad-

570 Id. at 78.
Deterring Retributivism

dition, by claiming to vindicate victims, despite ignoring their interests, retributivism is treating victims as objects rather than rational subjects—the same criticism by which retributivists have found rehabilitative theories of punishment untenable.

The fourth objection is that retributivism does not use victims at all; rather, it uses facts about victims as mere means. Even if true, retributivism still violates Kant’s maxim by failing to satisfy the positive duty to treat victims as ends in themselves. In addition, if retributivism only uses facts about victims (but not victims themselves), then consequentialism could equally avail itself of this claim. As a result, a deterrence-based theory does not use the punishment of offenders as mere means, but only uses the fact, or perceived fact that offenders are being punished, in order to promote the good consequence of deterrence. Either way, as to Kant’s injunction of never treating persons as mere means but always as ends in themselves, retributivism fares no better than consequential theories of punishment. To the extent that retributivism can escape from the very same criticism it has leveled against consequentialist theories, consequentialist theories equally escape.

VI. CONCEPTUALLY CONSEQUENTIAL RETRIBUTIVISM’S PUNISHMENT OF THE INNOCENT AND USE OF OFFENDERS AS MERE MEANS

For the most fundamental form of retributivism, which we have termed simple or pure retributivism, the desert of those punished constitutes the justification of punishment, and punishment is thereby an end in itself. In contrast, conceptually consequential retributivism justifies punishment based on the abstract or conceptual consequences punishment generates. These abstract or conceptual consequences are claimed to be meaningfully distinct from the factual or contingent consequences that are promoted by consequentialist theories of punishment. This Part will demonstrate that a large class of retributivist theories—conceptually consequential retributivism—succumbs to the same criticisms that retributivists make about consequentialism precisely because of its reliance on consequences to justify punishment.

After explaining why retributivists are driven to resort to consequences—albeit a special type of consequences—to justify punishment, we will consider whether conceptually consequential retributivism collapses into consequentialism. Even if it does not, the distinction between contingent and conceptual consequences will be shown to be so slender as to allow a deterrence-based theory, justifying punishment by the conceptual consequences of deterrence, to fall under the rubric of retributivism. And even if it does not, the utilization of consequences makes conceptually con-

571 See supra note 127 and accompanying text.
572 For a brief explanation of this class of retributivism, see supra text accompanying notes 124–126.
sequential retributivism succumb to the same criticisms retributivists have leveled against (contingently) consequentialist theories of punishment. Regardless of the nature of the consequences used to justify punishment, the employment of consequences in retributivism will be shown to justify the intentional punishment of particular, identifiable innocents as well as the use of offenders as mere means. In other words, while the conceptual nature of the consequences used to justify punishment in retributive theories may or may not prevent a slide into consequentialism, retributivism's use of consequences (regardless of their nature) leads to the same problems as incurred by (contingently) consequentialist theories of punishment.

A. Conceptually Consequential Retributivism

This subpart will explain conceptually consequential retributivism. After considering why retributivism might resort to consequences to justify punishment, it will examine the distinction between actual and conceptual consequences and will provide examples of conceptually consequential retributivism. Next, it will contrast conceptually consequential retributivism with both simple, or pure, retributivism as well as (contingent) consequentialism. Finally, it will consider whether conceptually consequential retributivism collapses into consequentialism.

1. Why Use Consequences? The Retributivism Dilemma: Circular-ity or Slide into Consequentialism.—The existence of both the pure and the conceptually consequential forms of retributivism may stem from the dilemma retributivism faces between circularity and collapse into consequentialism. In attempting to justify simple or pure retributivism as the preferable theory of punishment, one is limited in the types of arguments that can be elicited without sliding into consequentialism. But without eliciting some additional argument apart from “it’s right to punish criminals because doing so is right,” retributivism is widely regarded as circular or empty. Hugo Bedau succinctly states the problem for a retributivist:

\[\text{HART, supra note 11, at 9 (contending that retributivist accounts “all either avoid the question of justification altogether or are in spite of their protestations disguised forms of Utilitarianism”); Fletcher, supra note 501, at 53 (“When intuitive retributivists are challenged they often retreat into a vague consequentialism . . . .”).} \]

\[\text{See, e.g., Benn, supra note 45, at 30 (noting that Kantian justification for retributivism serves “not to justify punishment but, rather, to deny that it needs any justification”); Dolinko, supra note 12, at 517; Fletcher, supra note 501, at 53 (“[T]he intuitive [retributivist] argument that we must sanction evil deeds with punishment hardly seems like an argument at all.”); Kaplow & Shavell, supra note 28, at 1232 (“It seems that some defenders of the retributive view take it to be self-evidently correct. Most proponents of retributivist theory do little more than assert [the validity of retributivism] . . . .”); id. at 1233 (noting that retributivist justifications of punishment “seem virtually indistinguishable from restatements of the definition of the notion of retribution” (footnote omitted)); J.L. Mackie, Retributivism: A Test Case for Ethical Objectivity, in PHILOSOPHY OF LAW, supra note 55, at 677, 679 (“‘Desert’ is not a further explanation [of how punishment is justified under retributivism], but is just the general, as yet unexplained, notion of positive retributivism itself.”); Quinton, supra note 41, at 6 (retributivism consid-}
Either he [the retributivist] appeals to something else—some good end—that is accomplished by the practice of punishment, in which case he is open to the criticism that he has a nonretributivist, consequentialist justification for the practice of punishment. Or his justification does not appeal to something else, in which case it is open to the criticism that it is circular and futile.575

Retributivists have two options in countering the problem of circularity or emptiness. First, they might deny that the circularity is problematic. Moore argues that retributivism is “no worse off than any other nonutilitarian theories in ethics, each of which seek to justify an institution or practice not by the good consequences it may engender but rather by the inherent rightness of the practice.”576 This might be a persuasive argument to a fellow nonconsequentialist, but to a consequentialist it only demonstrates the problem with nonconsequentialist theories and supports the adoption of consequentialist views.

The second solution to the problem of circularity or emptiness in simple or pure retributivism is to avoid it altogether by justifying punishment in terms of the good consequences that punishment generates. Though this avoids one horn of the dilemma— circularity or emptiness—it triggers the other. If retributivism is to avoid the charge of circularity or emptiness by justifying punishment by the good consequences generated, then it is subject to the claim of collapsing into consequentialism.

Retributivists reply by invoking a distinction between actual and conceptual consequences.577 Consequentialist theories, such as deterrence or rehabilitation, justify punishment by actual or contingent consequences. These types of consequences are, in theory, subject to empirical verification; they either occur or do not occur in the real world. A given instance of punishment, or punishment in general, either promotes deterrence or it does

---

575 Bedau, supra note 83, at 616. For a similar view, see Benn, supra note 457, at 327: “[W]hat pass for retributivist justifications of punishment in general, can be shown to be either denials of the need to justify it, or mere reiterations of the principle to be justified, or disguised utilitarianism.”


577 See DUFF, supra note 1, at 7; Fletcher, supra note 14, at 516.
not. The justification of punishment, then, is contingent on the good consequence being promoted. In contrast, conceptual consequences are claimed to logically and necessarily follow punishment and therefore are not contingent. The justification of punishment by conceptual consequences is not contingent on being subject to empirical verification and whether or not they actually occur in the real world. By resorting to merely conceptual, rather than actual or contingent, consequences, conceptually consequential retributivism is claimed not to collapse into consequentialism.\(^{578}\)

Conceptually consequential retributivism, then, may be seen as retributivism’s answer to both horns of the above dilemma. Because punishment is justified by something external to itself (punishment’s consequences), the charge of circularity or emptiness is avoided. And because the consequences are conceptual, rather than actual or contingent, conceptually consequential retributivism purportedly avoids collapsing into consequentialism. This latter claim is subject to some dispute and will be considered further below.

2. Examples.—Let us consider some examples of the class of retributivist theories that I have designated conceptually consequential retributivism. A wide variety of conceptual consequences are utilized to justify punishment, including: expressing solidarity with victims,\(^ {579}\) restoring a relation of equality between the will of the victim and of the offender,\(^ {580}\) annulling or negating the crime,\(^ {581}\) avoiding the diminishment of society’s values,\(^ {582}\) demonstrating the seriousness of the criminal conduct as something that should not have been done,\(^ {583}\) voicing society’s emphatic denunciation of crime,\(^ {584}\) removing the advantage gained by the offender in freely indulging his will,\(^ {585}\) and negating the offender’s advantage gained by his renunciation of self-restraint.\(^ {586}\) Conceptually-consequential retributivists also justify punishment on grounds such as: educating the offender,\(^ {587}\) "re-
Deterring Retributivism

storing the equilibrium of benefits and burdens,\textsuperscript{588} effecting "a connection with correct values for those who have flouted them,"\textsuperscript{589} counteracting a dominance-submission relationship between offender and victim,\textsuperscript{590} "re-establish[ing] the acknowledgment of the victim's worth damaged by the wrongdoing and . . . repair[ing] the damage done to the victim's ability to realize her value,"\textsuperscript{591} "bring[ing] the criminal to understand the nature and implications of her crime,"\textsuperscript{592} and freeing a victim's blood or life spirit from the control of the offender.\textsuperscript{593}

It is not always clear, however, whether a purportedly retributive theory utilizing consequences to justify punishment is employing actual or conceptual consequences. Let us take for example Kant's theory of punishment. After posing a hypothetical of an island nation about to dissolve and whose inhabitants would disperse throughout the world, Kant asks whether the executions of the murderers still on death row should be carried out even though there will be no deterrent effect (because the society is about to disband).\textsuperscript{594} Kant answers that every murderer must be executed so that "blood guilt does not cling to the people for not having insisted upon this punishment; for otherwise the people can be regarded as collaborators in this public violation of justice."\textsuperscript{595} How should Kant's justification for executing the murderers on death row be classified? Simple or pure retributivism is purported to be retrospective and immune to the push or pull of the future consequences of punishment.\textsuperscript{596} But Kant is clearly attempting to justify retributive punishment by invoking a prospective perspective. Kant is seeking to justify punishment by recourse to the bad future consequences that would befall others if murderers were not executed. Punishment is justified because it eliminates the bad consequences of blood guilt clinging to members of society and the people being complicit in a public miscarriage of justice. So, if it is not simple or pure retributivism—because consequences are employed to justify the punishment—are the consequences actual or conceptual?

A retributivist might reply that we should not take Kant's metaphysical

\textsuperscript{588} MÖRRISS, supra note 6, at 34. The full quote is as follows: "Justice—that is punishing—such individuals—restores the equilibrium of benefits and burdens by taking from the individual what he owes, that is, exacting the debt." \textit{Id.}

\textsuperscript{589} NOZICK, supra note 104, at 384. For the view that Nozick's theory collapses into consequentialism, see C.L. TEN, supra note 35, at 45–46.

\textsuperscript{590} Fletcher, supra note 526, at 1634–35.

\textsuperscript{591} Hampton, supra note 5, at 1686.

\textsuperscript{592} DUFF, supra note 1, at 259.

\textsuperscript{593} See DAUBE, supra note 527, at 122–23 (explaining the Biblical basis for Kant's notion of blood guilt); see also Fletcher, supra note 501, at 60.

\textsuperscript{594} KANT, supra note 22, at 142.

\textsuperscript{595} \textit{Id.}

\textsuperscript{596} See supra notes 86–87 and accompanying text.
musings on blood guilt too seriously. All that is meant is that punishment precludes other members of society from being complicit in the crime. The good to other members of society that comes from avoiding blood guilt and complicity in the crime is only an abstract, conceptual good or consequence. But it is not clear that what Kant is referring to is only abstract and conceptual. Kant’s theory has been said to carry clearly consequentialist strains, appeal to “unmistakenly consequentialist” considerations, invoke “utilitarian considerations... irreconcilable with his main doctrine,” and be consequentialist as to the threat of punishment. Perhaps beneath the fancy language of blood guilt and complicity in crime

597 See FLETCHER, supra note 46, at 37–38.
598 Id. at 38.
599 See Fletcher, supra note 14, at 516.
600 For the view that the consequences by which Kant’s theory justifies the punishment of the murderers on death row are symbolic or conceptual, see Joel Feinberg, The Expressive Function of Punishment, 49 MONIST 3 (1965), reprinted in PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT, supra note 34, at 25, 33.
601 See Murphy, supra note 249, at 509 (“It is no longer clear to me to what extent it is proper to continue to think of Kant as a paradigm retributivist in the theory of punishment.”). Murphy, a Kantian scholar and former retributivist, goes on to list 14 passages from throughout Kant’s writings that reveal a lack of commitment to retributivism and explicit leanings toward consequentialism. Id. at 513–16. In one such passage, Kant seems to declare himself a consequentialist: Punishments “‘imposed by governments are always deterrent. They are meant to deter the sinner himself or to deter others by making an example of him.’” Id. at 513 (quoting Immanuel Kant, Universal Practical Philosophy, in LECTURES ON ETHICS 55 (Louis Infield trans., 1978)). Murphy summarizes the 14 passages as follows:

To summarize, justified punishment is a deterrence system functioning to maintain a system of ordered liberty of action. To set any more morally ambitious goal for punishment would be to adopt an unacceptable theory of the role of the state and would represent an attempt to play God. . . . Punishment is a necessary evil, but we should inflict and support it with regret and without any sense of having embarked on a moral crusade.

Murphy, supra note 249, at 517–18.
602 Cottingham, supra note 116, at 243–44 (arguing that Kant is justifying punishment by “unmistakenly consequentialist” considerations, namely the avoidance of the bad consequences that would occur if crime is not punished).
603 EWING, supra note 57, at 16 n.1. Ewing refers to a passage in which Kant suggests that consequentialist concerns justify deviations from the lex talionis. Kant carves out an exception to his view that murder always requires capital punishment. If capital punishment would sufficiently depopulate the state or desensitize the subjects to violence, a different punishment could be imposed. As Kant explains:

If, however, the number of accomplices (correi) to such a deed is so great that the state, in order to have no such criminals in it, could soon find itself without subjects; and if the state still does not want to dissolve, that is, to pass over into the state of nature, which is far worse because there is no external justice at all in it (and if it especially does not want to dull the people’s feeling by the spectacle of a slaughterhouse), then the sovereign must also have it in his power . . . [to] pronounce a judgment that decrees for the criminals a sentence other than capital punishment, such as deportation, which still preserves the population.

KANT, supra note 22, at 143 (footnote omitted).
605 For other evidence from Kant’s writings of some leanings toward consequentialism, see supra note 76 and accompanying text.
is the very ordinary, concrete, and nonconceptual emotion of guilt. That is, punishment is imposed in order to avoid feeling guilty about not punishing.606

3. Does Conceptually Consequential Retributivism Collapse into Consequentialism?—Rather than attempt to ascertain whether Kant's theory of punishment, or any particular retributive theory,607 collapses into

606 Retributivists have placed great stock in our emotions regarding crime and punishment. Anger, hatred, and outrage are not merely defended and legitimized, but celebrated and glorified as appropriate retributive responses to crime. JEFFREY G. MURPHY & JEAN HAMPTON, FORGIVENESS AND MERCY 94-95 (1988); Moore, supra note 119, at 210; Samuel H. Pillsbury, Emotional Justice: Moralizing the Passions of Criminal Punishment, 74 CORNELL L. REV. 655, 671-72 (1989); Roger Wertheimer, Understanding Retribution, 2 CRIM. JUST. ETHICS, Summer-Fall 1983, at 19, 23. The intensity and appropriateness of such emotions are even used to justify the institution of retributive punishment. See Murphy, supra note 471, at 215.

Given this centrality of emotions, particularly guilt, in retributive thinking, it is not a stretch to conclude that avoiding the abstract ends of blood guilt and complicity entail avoiding the very concrete guilt feelings generated by a failure to mete out retributive justice. Thus there is some basis to argue that deserving wrongdoers, by being punished, are treated as means to prevent others from feeling guilt under a retributive theory of punishment. 607 In addition to Kant's theory being accused of collapsing into consequentialism, the other theory exerting a powerful influence on retributivism—Hegel's—is also charged with collapsing into consequentialism. Hegel's view is that punishment negates or annuls the crime or wrongdoing. But it is claimed to collapse into utilitarianism or consequentialism:

The doctrine of "annulment," however carefully wrapped up in obscure phraseology, is clearly utilitarian in principle. For it holds that the function of punishment is to bring about a state of affairs in which it is as if the wrongful act had never happened. This is to justify punishment by its effects, by the desirable future consequences which it brings about.

Quinton, supra note 41, at 7-8. Hegel's theory also appears to be consequentialist in determining the degree of desert and punishment. The degree of an offender's punishment is, at least in part, a function of the danger it poses to society at a particular time. For Hegel, crime's
danger to civil society is a determination of its magnitude . . . . This quality or magnitude varies, however, according to the condition of civil society, and this is the justification both for attaching the death penalty to a theft of a few pence or of a turnip, and for imposing a lenient punishment for a theft of a hundred and more times these amounts.

HEGEL, supra note 22, at 251 (§ 218). Hegel sets the degree of punishment for a particular crime based upon the incidence of crime in, or the stability of, a particular society:

If society is still inwardly unstable, punishments must be made to set an example, for punishment itself is a counter-example to the example of crime. But in a society which is internally stable, the positedness of crime is so weak that the cancellation [Aufhebung] of this positedness must itself assume similar proportions. Thus, harsh punishments are not unjust in and for themselves, but are proportionate to the conditions of their time . . . .

Id. at 251 (§ 218) (Addition).

That the degree of punishment for a particular offense should vary according to the particular current conditions of the state or society is similar to utilitarian theories of punishment such as Beccaria's and Bentham's. For example, Beccaria argued that the severity or mildness of the punishment for a particular offense should vary based on the current conditions of a society or state:

I conclude with the reflection that the magnitude of punishment ought to be relative to the condition of the nation itself. Stronger and more obvious impressions are required for the hardened spirits of a people who have scarcely emerged from a savage state. . . . But, to the extent that human spirits are made gentle by the social state, sensibility increases; as it increases, the severity of punishment must diminish if one wishes to maintain a constant relationship between object and
consequentialism, let us consider whether the entire class of conceptually consequential retributivist theories collapses into consequentialism. Let us assume that a plausible account may be given of the difference between abstract or conceptual consequences and contingent or factual consequences. The question, however, is whether the use of only the former type of consequence, but not the latter, is sufficient to preclude retributivist theories from collapsing into consequentialism. That is, is a purportedly retributivist theory of punishment which justifies punishment only by the good conceptual, but not factual or contingent, consequences still a retributivist theory? For one retributivist, Moore, the answer appears to be no—justifying punishment by its conceptual consequences makes punishment merely an instrumental good, rather than an intrinsic good. For Moore, feeling.

BECCARIA, supra note 102, at 81 (Ch. XLVII) (translator's footnote omitted). As to determining the degree of desert and justifying the amount of punishment, then, Hegel seems to be squarely in the consequentialist camp. For the view that Hegel is retributivist in determining if someone should be punished but consequentalist in determining how much someone should be punished, see EwING, supra note 57, at 24; Dubber, supra note 89, at 1582.

Even retributivists criticize other retributivist theories for collapsing into consequentialism. Moore criticizes George Fletcher's theory, see supra note 501, justifying punishment by the good conceptual consequences of vindicating the victim as collapsing into consequentialism:

[You cannot hold Fletcher's theory and still be a retributivist. Fletcher's actual language is here instructive; he argues that the function of the criminal law is to end dominance, to terminate dominance, to counteract dominance, or to correct dominance. The point of punishment for Fletcher, in other words, is to cause a certain state of affairs to come into being by punishing; namely, the state of affairs whereby this dominating relationship between offender and victim no longer exists. Yet to say this is not to be a retributivist any more. Fletcher recognizes that you can't have causal relations between the good that justifies punishment, and the act of punishing, and yet be a retributivist; rather, you are an instrumentalist, someone for whom punishing those who deserve punishment is not intrinsically good but only instrumentally good.

Moore, supra note 26, at 82–83 (footnote omitted). For other conceptually consequential retributivist theories which have been criticized as collapsing into consequentialism see supra notes 582 & 589.

For discussion of the distinction, see Duff, supra note 1, at 7, 10.

An intrinsic good, or something which has intrinsic value, is good or has value in and of itself regardless of whether it can generate or lead to other goods or other things of value. On the other hand, an instrumental good, or something which has instrumental value, is not necessarily a good, or does not have value, in and of itself, but only is a good or has value because it is an instrument towards, or leads to, other goods or other things of value. On the intrinsic good/instrumental good distinction, see WILLIAM K. FRANKENA, ETHICS 64–67 (1963); GERALD F. GAUS, VALUE AND JUSTIFICATION: THE FOUNDATIONS OF LIBERAL THEORY 126–130 (1990); G.E. MOORE, PRINCIPIA ETHICA 23–30 (Cambridge Univ. Press 1956) (1903); MOORE, supra note 73, at 157; NOZICK, supra note 104, at 413–15; Duff, supra note 6, at 5–7; Noah M. Lemos, Value, in THE CAMBRIDGE DICTIONARY OF PHILOSOPHY, supra note 56, at 948, 948–49.

For one retributivist, Moore, the intrinsic/instrumental distinction is precisely what marks the distinction between retributivist and consequentialist theories of punishment. MOORE, supra note 73, at 157 ("[W]hat is distinctively retributivist is the view that the guilty receiving their just deserts is an intrinsic good.").

For the view that retributive punishment is understood as an intrinsic good, see for example, EwING, supra note 57, at 13, 14 & nn.1–2; MOORE, supra note 73, at 157; NOZICK, supra note 104, at 374; Benn, supra note 45, at 30; Lawrence Davis, They Deserve to Suffer, 32 ANALYSIS 136, 136 (1972); Dolinko, supra note 12, at 515–28; Duff, supra note 6, at 6–7. But see Mundle, supra note 236, at 74.
Retributivism requires that punishment be only an intrinsic good or end in itself, the good or end in itself of punishing those who deserve it.\textsuperscript{610} A consequentialist or instrumentalist theory of punishment justifies punishment based on the good consequences that follow from punishment.\textsuperscript{611} Punishment is therefore not intrinsically good, but only instrumentally good in order to attain some other good.\textsuperscript{612} If we punish in order to attain the goods of preventing blood guilt from clinging to society, preventing society's complicity with the crime, expressing solidarity with the victim, restoring a relationship of equality between victim and victimizer, or for similar reasons, then punishment of those who deserve it is not an intrinsic good but rather an instrumental good. Punishment, then, is instrumentally good in attaining these other goods or good consequences.

If the intrinsic/instrumental distinction separates retributive from consequential theories of punishment, then theories justifying punishment by the good conceptual consequences that are attained slide into consequentialism. In other words, the claim that the consequences are merely conceptual is not enough to prevent a slide into consequentialism, because the good of punishing those who deserve it is not exclusively an intrinsic good but is also an instrumental good in achieving the various good conceptual consequences. Justifying punishment by the promotion of the consequences stemming from punishment necessarily renders punishment itself only an instrumental good in service to those other consequences. That the consequences stemming from punishment are conceptual or symbolic or intangible does not alter the classification of the good of punishment itself as instrumental.

Rather than try to reach an ultimate resolution of what may be only a terminological or definitional debate, let us cash out what it will mean for retributivism if the conceptual consequences promoted by punishment may be used to justify punishment. That is, let us consider some of the consequences of retributivism's use of conceptual consequences.

\textbf{B. Is Conceptual Deterrence a Form of Retributivism?}

Let us assume arguendo that even if the consequences used to justify punishment are only conceptual, that still suffices to prevent the slide into consequentialism. Even so, the resulting distinction between retributive and consequential theories of punishment is tenuous. If the conceptual/factual consequences distinction is to serve as the litmus test for distinguishing theories of punishment, then retributive and consequential theories will still come quite close to collapsing into each other. That is, even if conceptually consequential retributivism does not collapse into consequentialism, consequentialist theories might slide into retributivism.

\textsuperscript{610} Moore, \textit{supra} note 26, at 82–83.
\textsuperscript{611} \textit{See supra} note 109.
\textsuperscript{612} Moore, \textit{supra} note 26, at 82–83.
Could not a consequentialist theory of punishment, justifying punishment by the deterrence of crime, claim that the deterrence attained is not factual or contingent deterrence but merely conceptual or rational deterrence? By conceptual deterrence I mean that no one need actually be deterred from the commission of a crime by the punishment of another. The justification of punishment by conceptual deterrence would not be contingent on some empirically verified, or verifiable, deterrent effect. The claim would merely be that part of the nature, or conceptual meaning, of punishment is that it would be rational for people to avoid its imposition. Conceptual deterrence, it might be asserted, necessarily or logically follows from the imposition of punishment. Thus, such deterrence is not factual or contingent but only abstract or conceptual. Let us term this the conceptual deterrence theory of punishment.

Is this conceptual deterrence theory of punishment consequential or retributive? To answer this question let us consider some features of the theory. The theory justifies the imposition of punishment neither because punishing those who deserve it is intrinsically good, nor because it is an end in itself, but, rather, in order to attain the good theoretical consequence of rational or conceptual deterrence. Thus, punishment itself is only an instrumental good. Would the punishment of innocents be justified under this theory? Yes, if it would serve to promote the good of conceptual deterrence. That is, if by punishment of innocents it would be less rational for other members of society to commit crime, then intentional punishment of innocents would be justified. Would this conceptual deterrence theory use offenders as mere means to attain the good consequence of conceptual deterrence? Yes. Because the end of punishing is to promote conceptual deterrence, the punishment of the offenders is the mere means. Our conceptual deterrence theory certainly does not share many of the features traditionally associated with retributivism. But according to retributivists, the conceptual nature of the consequences apparently suffices to prevent our conceptual deterrence theory from being consequential. As a result, the conceptual deterrence theory seems to be a retributive theory of punishment after all.

The point of this perhaps minor reductio ad absurdum is that classifying retributive and consequential theories based on the conceptual/factual consequences distinction threatens the collapse of the wall between the theories of punishment. Even if we accept some retributivists' insistence that justifying punishment based on the good conceptual consequences to be attained does not constitute a slide into consequentialism, that does not

---

613 If intentionally punishing the innocent, perhaps by falsely claiming that they were guilty, would make it rational for others not to commit crime, then conceptual deterrence would be promoted. Because if what justifies punishment under this theory—conceptual deterrence—is promoted by the intentional punishment of the innocent, the conceptual deterrence theory justifies the intentional punishment of the innocent.
preclude consequentialist theories from sliding into retributivism. Moreover, since the conceptual deterrence theory is a form of retributivism (based on the conceptual/factual consequences distinction), then retributivism justifies the intentional punishment of the innocent and the use of offenders as mere means.

As we will see in the next subpart, the most significant effect of retributivist theories justifying punishment by the consequences (even if only conceptual) punishment promotes is that conceptually consequential retributivism incurs the same problems for which retributivists have found consequentialist theories untenable. Though the nature of the consequences may be different, the problems stemming from the use of consequences are the same. Conceptually consequential retributivist theories will be shown to justify the intentional punishment of particular, identifiable innocents and the use of offenders as mere means.

C. Punishment of the Innocent and Use of Offenders as Mere Means

The class of retributive theories we have designated conceptually consequential retributivism justifies the intentional punishment of the innocent and the use of offenders as mere means because of its use of consequences (even if only conceptual) to justify punishment. To see this, it might be helpful first to understand why consequential theories are, and simple or pure retributivism is not, generally thought to be vulnerable to these criticisms. Let us take, for an example, a general deterrence-based consequential theory of punishment, which justifies punishment by the good factual or contingent consequences that punishment produces.\(^614\) If the good consequence of deterrence may be obtained by punishing the innocent as well as the guilty, then the attainment of that consequence justifies the punishment regardless of the offender's guilt or innocence.\(^615\) Guilty or innocent, the person punished is the mere means by which that good consequence, in this case deterrence, is obtained. Rather than treating the person punished as an end in himself, the person punished is merely the means by which the end or goal of punishment (in this case deterrence) is attained. At least, this is how retributivists would explain it. Simple or pure retributivism, because it does not justify punishment by resort to its consequences, is generally thought not to be susceptible to these criticisms. As one commentator explains, simple or pure retributivism justifies punishment:

\[\text{[A]n appropriate response to past wrong-doing which has, and needs, no further purpose beyond itself: for if punishment has no further aim it cannot be accused of manipulating those on whom it is imposed or against whom it is threatened—of using them as means towards a further end.}\] \(^616\)

\(^{614}\) See supra notes 61, 62, 124 and accompanying text.

\(^{615}\) See supra note 141-143, 160-161, 167 and accompanying text.

\(^{616}\) DUFF, supra note 1, at 186.
Is the susceptibility of the (contingently) consequentialist deterrence theory of punishment to the twin criticisms of intentional punishment of the innocent and use of offenders as mere means dependent on, or a function of, the nature of the consequences sought to be promoted? That is, does it matter whether the consequences sought to be promoted are contingent or conceptual? It is generally assumed that it is not the use of consequences themselves but rather the contingent nature of the consequences that renders consequentialist theories vulnerable to criticisms of punishing the innocent and using offenders as mere means. As one commentator expresses the assumption:

The contingency of the relation between punishment and its consequentialist aims generates the most familiar objection to any purely consequentialist theory of punishment: that it would justify clearly unjust kinds of punishment (the deliberate punishment of an innocent scapegoat, the excessively harsh punishment of the guilty), if they would serve the system’s aims.

According to this argument, then, it is the contingent nature of the consequences, rather than the use of consequences themselves, that causes the problems of consequentialist theories of punishment.

This idea apparently underpins the rationale of conceptually consequential retributivism. Because the consequences necessarily or logically follow from punishment of those who deserve it, every instance of punishing those who deserve it will generate the specified conceptual consequence. Therefore, promotion of the good conceptual consequences will not justify punishment of all those (including the innocent) whose punishment would be justified by actual or contingent consequences. That is, the attainment of good contingent consequences will, but the attainment of good conceptual consequences will not, justify punishment of the innocent or the use of offenders as mere means.

But the logic of this argument is seriously flawed. That punishment of those who deserve it will invariably generate the specified good conceptual consequence does not mean that punishment of those who deserve it is the only way to attain the good conceptual consequences. Punishment of those who deserve it may be a sufficient condition for the attainment of the good conceptual consequences. But it is not a necessary condition. Therefore, the good conceptual consequences may be generated without punishing those who deserve it. This opens the door for the possibility of attaining good conceptual consequences by punishing those who do not deserve it. As a result, attainment of the good conceptual consequence could justify punishment of the innocent and the use of offenders as mere means.

---

617 See id. at 2 (explaining that the "contingent relation between punishment and its justifying aims generates [the] familiar objections" to consequentialist theories, including punishment of innocent scapegoats and, in general, imposing unjust punishment).

618 Duff, supra note 6, at 6.
We have already seen this in our discussion of the conceptual deter-
rence theory, which justified punishment not on any actual deterrent effect
realized in the world but on the promotion of rational or conceptual deter-
rence. We noted that the theory would be subject to the criticisms that it
justifies intentional punishment of the innocent and the use of offenders as
mere means. Thus, the susceptibility of a deterrence-based theory (justify-
ing punishment either by conceptual or contingent deterrence) to these cri-
cisms does not seem to depend on, or be a function of, the nature of the
consequences—contingent or conceptual—sought to be promoted.

That deterrence-based theories are subject to these twin criticisms re-
gardless of whether the consequence to be attained is conceptual or factual
suggests that the abstract nature of the consequences in conceptually conse-
quential retributivism will not preclude such retributivist theories from be-
ing subject to the same criticisms. If so, then it would seem to be not the
nature of the consequences, but merely the use of consequences per se, that
renders a theory of punishment subject to these criticisms.

1. Intentional Punishment of Particular, Identifiable Innocents.—To
see if this conjecture is true, let us consider some particular theories from
the class of conceptually consequential retributivism. The various concep-
tual consequences used to justify punishment in the class of conceptually conse-
quential retributive theories include: expression of solidarity with vic-
tims,\textsuperscript{619} restoration of a relation of equality between the will of the victim
and of the wrongdoer,\textsuperscript{620} dissolution of a dominance-submission relation-
ship between wrongdoer and victim,\textsuperscript{621} freedom from blood guilt,\textsuperscript{622} and
restoration of the balance of benefits and burdens.\textsuperscript{623} Let us apply any one
of these conceptually consequential retributive theories to the following hy-
pothetical.

Suppose that a culpable wrongdoer has culpably committed serious
wrongdoing against a victim but has evaded liability for any of a number of
reasons (such as a nonexculpatory or extrinsic defense) or that, in addition
to the serious wrongdoing, the wrongdoer has committed a minor offense
against the same victim for which he has been found guilty, but has not yet
been punished. In either case, even though the culpable wrongdoer has not
been adjudicated to be legally guilty of the serious criminal offense, all of
the above conceptual consequences would be promoted by punishing the
wrongdoer for the serious wrongdoing. Serious wrongdoing has been
committed which has produced a need to express solidarity with the victim,
a relation of equality to restore, a dominance-submission relationship to

\textsuperscript{619} FLETCHER, supra note 518, at 201–05.
\textsuperscript{620} HEGEL, supra note 22, at 124 (§ 99).
\textsuperscript{621} Fletcher, supra note 526, at 1634–35.
\textsuperscript{622} KANT, supra note 22, at 142.
\textsuperscript{623} See generally MORRIS, supra note 6; MURPHY, supra note 113.
sever, blood guilt to be released, and a balance of benefits and burdens to be allocated. Just as the good consequence of deterrence might be promoted regardless of the guilt or innocence of the person punished, so too the various good conceptual consequences of the various conceptually consequential retributive theories could be promoted by intentionally punishing a particular, identifiable legal innocent.

A retributivist might argue that even if a retributive theory would justify punishment of the wrongdoer for the serious wrongdoing of which he is legally innocent, the person would never be punished for the very reason that he is legally innocent of the serious crime. While this may be true, retributivism per se would nonetheless justify the punishment. In other words, that a retributive system of punishment would not, as a matter of practice, punish the legally innocent person does not eliminate the problem that retributivism per se does justify such punishment.

Conceptually consequential retributivism also justifies the intentional punishment of particular, identifiable moral innocents. Suppose a defendant is convicted of an offense without committing a morally culpable wrong. For example, the defendant is convicted of an offense unjustly defined to be a crime or is convicted under a strict liability standard. Though the defendant has not engaged in morally culpable wrongdoing, she has nonetheless committed a legally defined criminal offense and has violated the legal rights (even if not moral rights) of another. Despite not committing a morally culpable wrong, all of the above conceptual consequences would be promoted by punishing her. A criminal offense has been committed, involving the violation of the legal rights of another, for which solidarity with the victim needs to be expressed, a relation of equality must be restored, a dominance-submission relationship has to be severed, a blood guilt needs to be released, and a balance of benefits and burdens must be allocated. Just as the good consequence of (contingent) deterrence might be promoted regardless of the guilt or innocence of the person punished, so too the various good conceptual consequences of the various conceptually consequential retributive theories could be promoted by intentionally punishing a particular, identifiable moral innocent.

And perhaps more importantly, even if a conceptually consequential retributivestystem of punishment would not punish innocents, conceptually consequential retributivism cannot explain to us why it would be wrong to do so. Recall that consequentialists have also argued, against the criticism by retributivists that deterrence-based theories justify the intentional punishment of the innocent, that a deterrence-based system of punishment would not, in fact, intentionally punish the innocent.624 Retributivists have replied that even if a consequentialist system of punishment would not intentionally punish innocents, consequentialism cannot explain why inten-

624 See supra notes 144–147, 153–155 and accompanying text.
Deterring Retributivism

The retributivist argument against consequential deterrence theories may be applied just as powerfully to conceptually consequential retributivism. That which justifies punishment for such theories—the promotion of the various conceptual consequences—fails to include any reason why innocents should not be punished. As a result, just as theories justifying punishment by factual or contingent consequences are subject to the criticism that they justify intentional punishment of particular, identifiable innocents, so also retributive theories justifying punishment by conceptual consequences are subject to the same criticism. More importantly, conceptually consequential retributivism cannot explain why punishment of the innocent is wrong.

2. Use of Offenders as Mere Means.—Part V concluded that retributivism, in general, uses the victims of crime as mere means in violation of Kant's maxim. This section will show that a specific class of retributive theories—conceptually consequential retributivism—uses offenders themselves as mere means and thus falls victim to the precise criticism that retributivists have directed at consequentialist theories. The good (contingent) consequences of, for example, deterrence or rehabilitation, justify punishment. These consequences are the purpose or end of what is of intrinsic value for consequentialist theories; the punishment of offenders is the mere instrument or means by which the good consequences are obtained. In contrast, retributivism is claimed to never treat offenders merely as means but always, at least partially, as ends in themselves. Because punishment of deserving offenders has intrinsic value and is an end in itself for simple or pure retributivism, the punishment of deserving offenders is claimed not to constitute the use of offenders as mere means. Retributivism in general is conventionally assumed to treat offenders as ends in themselves by giving them only what they deserve. But giving offenders only what they deserve may not be enough for conceptually consequential retributivism to avoid using offenders as mere means.

By resorting to consequences to justify punishment, conceptually consequential retributivism uses the punishment of wrongdoers as a mere means to attain the various good conceptual consequences. The purpose or end of punishing is to promote good conceptual consequences, such as: an-

625 See supra notes 158–159 and accompanying text.
626 See supra notes 397–399, 409–418 and accompanying text.
627 See supra note 109 and accompanying text.
628 See supra notes 617–618 and accompanying text.
629 See supra note 400 and accompanying text.
630 See supra text accompanying note 616.
631 See supra note 400 and accompanying text.
nulling or negating the crime,632 avoiding the diminishment of society’s values,633 expressing the seriousness of the criminal conduct as something that should not have been done,634 expressing society’s emphatic denunciation of crime,635 removing the advantage gained by the offender in freely indulging his will,636 negating the offender’s advantage gained by his renunciation of self-restraint,637 and educating the offender.638 The persons being punished are merely the means by which those good conceptual consequences are to be obtained. It is the attainment of the good conceptual consequence that is of intrinsic value; the punishment of the offender is only of instrumental value. Thus, just as retributivists claim of consequentialist theories, conceptually consequential retributivism uses offenders as mere means.

A conceptually consequential retributivist might concede that offenders are used as means but still object that they are not used as mere means. While the punishment of offenders might be used as means to attain the good conceptual consequences, they are also treated as ends by punishing them in accordance with what they deserve. Therefore, conceptually consequential retributivism does not use offenders as mere means. But the retributivist critique of consequentialism suggests this objection fails.

Retributivists have argued that consequentialist theories use offenders as mere means, in violation of Kant’s maxim, even in those instances in which only the guilty are punished.639 Jeffrie Murphy points out that the consequentialist theory faces Kantian objections not only to punishment of the innocent, but also to punishment of the guilty.640 This is because it

[M]ust involve justifying punishment in terms of its social results—e.g., deterrence, incapacitation, and rehabilitation. And thus even a guilty man is, on this theory, being punished because of the instrumental value the action of the punishment will have in the future. He is being used as a means to some future good—e.g., the deterrence of others. Thus those of a Kantian persuasion, who see the importance of worrying about the treatment of persons as mere means, must, it would seem, object just as strenuously to the punishment of the guilty on utilitarian grounds as to the punishment of the innocent.641

According to Murphy, then, where punishment is justified by its consequences, even the guilty—those who receive the punishment they de-

---

632 Hegel, supra note 22, at 123 (§ 97).
633 Oldenquist, supra note 17, at 750.
634 Morris, supra note 583, at 266.
635 See supra note 584.
636 Finnis, supra note 585, at 134.
637 Murphy, supra note 6, at 14.
638 Parent, supra note 587, at 351–53.
639 Murphy, supra note 6, at 5; Wasserstrom, supra note 34, at 340–41.
640 Murphy, supra note 6, at 5.
641 Id.
serve—are used as mere means in violation of Kant's maxim.\textsuperscript{642}

Murphy’s critique of utilitarian or consequentialist theories applies equally to conceptually consequential retributivism.\textsuperscript{643} Because conceptually consequential retributivism justifies punishment based on the good conceptual consequences generated, punishment of an offender will be of instrumental value. The offender, even one who is guilty and is punished only as he deserves, is being used as a means to attain the good conceptual consequences—e.g., annulling or negating the crime. “Thus those of a Kantian persuasion, who see the importance of worrying about the treatment of persons as mere means, must, it would seem, object just as strenuously to the punishment of . . . [those who deserve it on the grounds of conceptual consequences] as to the punishment of the innocent.”\textsuperscript{654} In other words, conceptually consequential retributivism, even if it only punishes culpable offenders as they deserve, would still be objectionable on Kantian grounds of using offenders as mere means.

Of course, the obvious difference between Murphy’s critique of consequentialist theories and the critique’s attempted application to conceptually consequential retributivism is that the former theories utilize contingent consequences and the latter employ conceptual consequences. But is this difference, as it relates to using offenders as mere means, significant? It does not seem so. Regardless of the nature of the consequences, the punishment of the offender stands in the same relation to the consequences used to justify the punishment. Punishment of the offender is the mere

\textsuperscript{642} Richard Wasserstrom offers a similar argument that punishment of the guilty may trigger the same moral objections as punishment of the innocent. Wasserstrom, supra note 34, at 340. As Wasserstrom explains it, “[I]n both cases the objection is identical: a person is used as a means to benefit others—whether through deterrence in punishing the guilty; or some other real or presumed benefit to others, in punishing the innocent.” Id.

\textsuperscript{643} Wasserstrom’s argument, similar to Murphy’s, see supra text accompanying notes 640–641, could also be applied to conceptually consequential retributivism. The persons punished under conceptually consequential retributivism, whether guilty or innocent, are punished in order to attain conceptual consequences that purportedly benefit others. Thus, the persons punished (even the guilty) are used as mere means in order to benefit others. The conceptual consequences (used to justify punishment under various versions of conceptually consequential retributivism) that explicitly seek to benefit others include expressing solidarity with victims, restoring a relation of equality between the victim and offender, counteracting a dominance-submission relationship between offender and victim, and re-establishing the victim’s worth. All of these various conceptual consequences strive to benefit those other than the persons being punished. The attainment of these conceptual consequences purports to benefit the victims of crime. Thus, to the extent that conceptual consequences benefit others, Wasserstrom’s argument against the deterrence theory applies equally to conceptually consequential retributivism. As a result, conceptually consequential retributivism uses those who are punished (even if guilty) as mere means in a morally objectionable manner. Those who are punished are used as mere means in order to benefit others—for example, their victims.

\textsuperscript{654} Id. Though I am quoting from Murphy, the language inserted in the brackets constitutes a substantive alteration of Murphy’s analysis. I deleted Murphy’s language of “the guilty on utilitarian grounds” and substituted the language in brackets. This has been done to show that Murphy’s argument—that consequentialism not only uses innocents, but the guilty as well, as mere means—also would demonstrate that the promotion of conceptual consequences uses offenders as mere means.
means or instrument by which the good consequences are obtained. Attainment of the good consequences is of intrinsic value; the punishment of the offender is of instrumental value. Therefore, if punishment of those who deserve it, as justified by its contingent consequences, constitutes using offenders as mere means, then punishment of those who deserve it as justified by its conceptual consequences also constitutes using offenders as mere means. In other words, if (contingently) consequential theories use offenders as mere means, then conceptually consequential retributivism also uses offenders as mere means.

A retributivist still might argue that as long as an offender is punished in the manner that she deserves, she might be used as a means but not as a mere means. But if this is true for conceptually consequential retributivism, then it is equally true for (contingently) consequentialist theories. Regardless of whether consequentialist and conceptually consequentialist retributivism theories do or do not use offenders as mere means in violation of Kant's maxim, neither theory fares better than the other.

D. Objections

This subpart considers two objections to the above arguments against conceptually consequential retributivism. First, conceptually consequential retributivism justifies punishment not only by the good conceptual consequences to be obtained, but also because the offender deserves it. Second, the conceptual consequences elicited in support of retributivism are merely descriptive and play no part in the justification of punishment.

1. Justifies Punishment by Both Conceptual Consequences and Desert.—Adherents to conceptually consequential retributivism might point out that their theory justifies punishment not merely to promote good conceptual consequences but also to mete out the desert of the offender. So, conceptually consequential retributivism, properly understood, does not have a different justification than pure or simple retributivism but an additional one. Would such an understanding of the theory avoid the above criticisms?

Construing conceptually consequential retributivism in that way would provide a basis to distinguish the conceptual or rational deterrence theory from retributivism. Though the conceptual deterrence theory utilizes conceptual consequences to justify punishment (the conceptual consequence of rational deterrence), it does not also use the desert of the offender as a justification. Thus, the collapse of the conceptual deterrence theory into retributivism would be precluded.

The revised interpretation of conceptually consequential retributivism, however, would not avoid justifying the intentional punishment of particular, identifiable innocents. Suppose, as discussed above, one or more of the

---

645 See supra subpart VI.B.
good conceptual consequences could be promoted by punishing an innocent. The conceptual consequences component would justify such punishment. But the desert component would not. The two components of what justifies punishment would give contradictory answers. This incompatibility or instability of the two components, however, could easily be remedied in two ways.

First, the two components might operate disjunctively. That is, punishment would be justified if either good conceptual consequences might be promoted by punishing or if the person punished deserved it. Let us suppose again that the punishment of an innocent would serve to promote one or some of the good conceptual consequences. Promotion of the good conceptual consequences would then serve to justify the intentional punishment of a particular, identifiable innocent. Though the disjunctive operation of the two components yields a non-contradictory determination of whether a given instance of punishment is justified, the problem of punishing the innocent remains.

Second, the two components might operate conjunctively. That is, punishment would not be justified unless both (i) the offender deserves it, and (ii) the good conceptual consequence would be obtained. This seems more promising. With punishment justified only if both components are satisfied, the conjunctive operation of the components provides a non-contradictory determination of whether a given instance of punishment is justified. In addition, it would seem that even if the punishment of an innocent would promote a good conceptual consequence, such punishment would not be justified because the person punished does not deserve it. The conjunctive operation of the components would seem to avoid punishment of the innocent.646

And under the conjunctive approach, conceptually consequential retributivism would avoid using offenders as mere means. Though punishment of an offender would be used as a means to attain the good conceptual consequence, the offender would also be punished because he deserves it. Therefore, though the offender would be used as a means, he would not be used as a mere means. That one of the components justifies punishment because the offender deserves it means that the offender would, at least in part, be treated as an end.

646 But a more subtle version of the problem of justifying the intentional punishment of particular, identifiable innocents still remains. Under conceptually consequential retributivism, even the conjunctive version, there will be various approaches to what constitutes the status of deserving punishment. Under the moralistic conception, morally culpable wrongdoing renders one deserving of punishment. But as discussed above, see supra subpart II.B, this approach justifies the intentional punishment of particular, identifiable legal innocents. And the legalistic conception justifies the intentional punishment of particular, identifiable moral innocents. See supra subpart II.C. The combined moralistic and legalistic conception of deserving punishment avoids the problem, in part, of intentionally punishing the innocent. But as discussed above, see supra subpart III.C, the combined conception does this at the cost of giving, in some instances, contradictory determinations of the degree of an offender’s desert and punishment.
Though the conjunctive operation of the components might, in part, preclude conceptually consequential retributivism from intentionally punishing the innocent and using offenders as mere means, it creates another difficulty. Suppose that an offender deserves punishment but that for some reason the good conceptual consequence could not be attained by punishing that offender. Conceptually consequential retributivism, construed conjunctively, would not be able to justify the offender's punishment, thereby violating the retributive duty to punish the guilty.

Retributivists would, of course, reply that conceptual consequences necessarily or logically follow the punishment of a deserving offender. Whether they do or do not, however, is merely a matter of intuition. Retributivists offer no argument that punishment, in fact, does avoid society's complicity with a crime, express solidarity with the victim, or annul or negate the crime. But it may not matter so much whether conceptual consequences invariably follow deserved punishment or not.

Perhaps the more important point is that conceptually consequential retributivism could not explain why it would be wrong to forego punishing a deserving, guilty offender if the good conceptual consequences could not be attained by punishment. That is because if the good conceptual consequences could not be attained, then punishment of the guilty, culpable wrongdoer would be unjustified, and therefore, foregoing punishment would not be wrong under conceptually consequential retributivism. So, of course, conceptually consequential retributivists could not explain the wrong of not punishing deserving offenders. But presumably, not punishing the guilty, culpable wrongdoer is wrong, and conceptually consequential retributivism should be able to explain why. After all, simple or pure retributivism could offer an explanation because foregoing punishment of the guilty would be wrong under it. As Kant declares, "[t]he principle of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaemonism [the principle of utility] in order to discover something that releases the criminal from punishment . . . ." That is, punishing a deserving wrongdoer is a duty of justice, and good consequences, or the lack thereof, are never a reason to forego punishing a wrongdoer deserving punishment.

Despite their inability to explain why it is wrong to forego punishing a culpable wrongdoer if the good conceptual consequence could not be attained by punishing, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denying the "if" in our argument. That is, because of the conceptual nature of the consequences, the good conceptual consequences would always be promoted by punishing a deserving wrongdoer. As a result, a conceptually consequential retributivist might insist on denial
ally consequential retributivist would never need to explain the wrong of foregoing punishment of a deserving wrongdoer where the good conceptual consequences would not be promoted.

Even so, conceptually consequential retributivism still cannot supply a morally relevant explanation of why foregoing punishment of the guilty is wrong. Since both components are required for punishment to be justified, the wrong of not punishing must also be explained by resort to both components. Thus, the wrong of not punishing a deserving wrongdoer whose punishment would promote the good conceptual consequence must consist of both (i) the injustice of violating the duty to punish and (ii) the unattainability of the good conceptual consequence of punishment. Thus, the explanation must consist, at least in part, of the inability to attain the good conceptual consequences. But for retributivism, “[t]he state of affairs where a wrongdoer suffers punishment is morally better than the state of affairs where he does not; and it is better irrespective of any of the consequences of punishing him.” That is, the wrong of not punishing those who deserve punishment should outweigh any inability to attain beneficial conceptual consequences from the punishment. Therefore, to the extent that it relies on consequences, conceptually consequential retributivism cannot provide a satisfactory account of why it is wrong to forego punishment of the guilty.

This would seem to be a fair criticism because it is retributivists who have argued that the problem with consequentialism is not so much that it would punish the innocent but that it could not adequately explain why doing so would be wrong. That a consequentialist’s explanation might resort to violations of utility, counterproductivity, or that it would undermine deterrence has been criticized by retributivists as failing to capture the morally relevant reason why punishment of the innocent is wrong. In other words, that punishment of the innocent would generate bad consequences fails to account for our moral abhorrence to it. Equally, conceptually consequential retributivism’s reliance on abstract consequences to explain why foregoing punishment of the guilty is wrong also fails to capture our moral intuition that it is unjust.

The resulting dilemma for conceptually consequential retributivism is that (i) without the conjunctive interpretation, it justifies the intentional punishment of particular, identifiable innocents and uses offenders as mere means, but (ii) with the conjunctive interpretation, it cannot supply a morally relevant account of why foregoing punishment of guilty, deserving wrongdoers is wrong.

2. Conceptual Consequences Do Not Justify Punishment.—The ob-

---

651 Rawls, supra note 1, at 5.
652 See supra notes 158–159 and accompanying text.
653 Id.
jection might be that conceptually consequential retributivism does not utilize punishment's good conceptual consequences to justify punishment. It perhaps merely refers to the good conceptual consequences generated by punishment as a side-benefit. Or perhaps the conceptual consequences provide some answer to consequentialists who demand that punishment serve some purpose or end beyond itself. In any event, and for whatever reason retributivists might refer to punishment's conceptual consequences, the conceptual consequences are in no way part of the justification of punishment. Under such a construction of conceptually consequential retributivism, the actual and sole justification of punishment is the desert of the offender. Does this construction of conceptually consequential retributivism avoid any of the above criticisms?

The elimination of the conceptual consequences from the justification of punishment does eliminate the problem of using offenders as mere means. This is because the punishment of offenders is no longer the instrument or means by which something of intrinsic value is attained—the conceptual consequences. Since punishment of offenders is its own end, offenders are no longer used as mere means. And intentional punishment of the innocent would, in part, be avoided because the innocent would not deserve it.

But problems would remain. If the conceptual consequences play no part in the justification of punishment, then why restrict the side-benefit of punishment to only conceptual consequences? Why not refer to the side-benefit of actual consequences? If retributivism is to refer to some side-benefit of punishment that plays no part in the justification of punishment, then why would it not refer as well to the side-benefit of actual consequences such as deterrence? Moreover, if the sole justification for punishment is an offender's desert, then conceptually consequential retributivism collapses into simple or pure retributivism.

E. Conclusion

To avoid the charge of circularity or emptiness stemming from simple retributivism's formula of "it is right to punish criminals because doing so is right," a large class of retributive theories resort to the consequences of punishment to justify punishment. In turn, to avoid the criticism that the use of consequences collapses this class of retributivism into consequentialism, retributivists invoke the distinction between factual or contingent consequences and conceptual or logical consequences. By justifying

---

654 Even under this interpretation of conceptual consequential retributivism, the theory would still be subject to the problems of punishing the innocent, including under strict liability, discussed in supra Part IV. Such punishment of the innocent might be thought to constitute intentional punishment of the innocent. See text accompanying supra note 335. Moreover, both the moralistic and legalistic conceptions of conceptually consequential retributivism would justify the intentional punishment of particular, identifiable innocents. See supra subparts II.B–C.
Deterring Retributivism

punishment merely by the latter type of consequences, they claim, conceptually consequential retributivism avoids sliding into consequentialism. The distinction is revealed to be tenuous by showing that a deterrence-based theory, justifying punishment by rational or conceptual deterrence, would then be a form of retributivism. Thus, even if conceptually consequential retributivism does not collapse into consequentialism, a conceptual deterrence theory slides into retributivism.

Even if the distinction is accepted and conceptually consequential retributivism does not slide into consequentialism, the use of consequences to justify punishment for retributivism is problematic. Regardless of the nature of the consequences employed to justify punishment, the use of consequences by retributivism incurs the same problems as consequentialist theories. Because the purpose or end of punishment is the promotion of the good conceptual consequences, the punishment of wrongdoers is merely the means by which the good conceptual consequences may be attained. And because many of the good conceptual consequences may be promoted even if a defendant is innocent, the promotion of the good conceptual consequences justifies intentional punishment of particular, identifiable innocents.

Unless retributivism departs from the formula of desert as the sole justification for punishment, it is circular or empty. But once retributivism departs from desert as the sole justification for punishment by resorting to consequences, and since the consequences may be obtained by punishing an offender without desert, retributivism is subject to the very same problems of consequentialist theories—justifying intentional punishment of particular, identifiable innocents and the use of offenders as mere means.

CONCLUSION

Retributivism's claimed superiority and place on the moral high ground in the debate on the justification of punishment is based, in large part, on retributivists' claims that consequentialist, but not retributive, theories of punishment (i) justify the intentional punishment of particular, identifiable innocents, (ii) use persons as mere means in violation of Kant's maxim to always treat persons as ends in themselves, (iii) fail to give a principled account of the degree of an offender's desert and punishment, and (iv) punish without due regard to an offender's fault, culpability and guilt. This Article has held a mirror to these critiques and has demonstrated that retributivism suffers equally from each of these failings. In some instances, retributivists might be able to defend their theory from these criticisms but only by advancing arguments that allow consequentialist theories to avoid these problems as well. Either way, retributivism is no better than consequentialist approaches to punishment. And, as a theory of punishment, it might be worse. While undeniably flawed, consequentialism is true to itself. In contrast, because retributivists' own criticisms of conseque-
tialist theories may be turned back on retributivism, retributivism fails to satisfy its own criteria of just punishment.