Summer 1984

Retributivism and Justice

Thomas Morawetz

University of Connecticut School of Law

Follow this and additional works at: http://digitalcommons.uconn.edu/law_papers

Recommended Citation

Morawetz, Thomas, "Retributivism and Justice" (1984). Faculty Articles and Papers. 68.
http://digitalcommons.uconn.edu/law_papers/68
RETRIBUTIVISM AND JUSTICE

by Thomas H. Morawetz*

I. THE JUSTIFICATION OF PUNISHMENT

In two respects legal scholarship mirrors legal pedagogy. First, effective teaching demands description of how things are, explanation of why they are as they are and consideration of how they can be justified. Similarly, scholarship is often designed to explain as well as describe and to explain in such a way as to provoke normative questions about how things should be.

For some legal topics the relevant normative questions are patterned and predictable. The justification of punishment is one such topic. The competing theories are familiar and are the subject of most students' first immersion in the subject. Familiarity may however be a disguise for confusion. This observation brings us to the second shared goal of scholarship and teaching. Both may provoke the student and the reader to look behind the familiar forms of justification and ask whether these familiar forms really make sense.

The simplest theory of the justification of punishment is met in the law student's first brush with criminal law. This theory says that criminal punishment serves four goals: to incapacitate offenders (special deterrence), to give a general disincentive for crime (general deterrence), to rehabilitate, and to achieve retribution. The first three goals are identified as forward-looking in contrast with the problematic fourth. The first three goals are forward-looking in that special deterrence, general deterrence, and rehabilitation describe goals that are part of a picture of a future in which persons collectively are better off. The de-

* Professor of Law, University of Connecticut School of Law. A.B. 1963, Harvard; J.D. 1968; M.Phil. 1968; Ph.D 1969, Yale. Professor Morawetz teaches Criminal Law, Criminal Procedure, Jurisprudence, and Legal Profession.

My seven years at the University of Connecticut have been years of transition from scholarship and teaching in pure philosophy to the application of philosophy to law. I am indebted in countless ways to Dean Phillip Blumberg for his support and for easing that transition. This article is an attempt to illustrate as coherently as possible the role of philosophy in contemporary thinking about criminal law.

1. See, e.g., W. LaFave & A. Scott, CRIMINAL LAW 21-25 (1972).
terrence of crime and the rehabilitation of offenders help create a future in which overall happiness is maximized and the pursuit of individual goals is facilitated. This form of justification is normally called utilitarian.  

Retribution is problematic because it is not forward-looking and therefore seems to certain critics an irrational goal. Some of these critics contend that retribution is revenge, that it reflects the domination of emotion over reason. This argument can be met by saying that reasonable persons have shared standards for determining the just deserts of offenders, and that retribution is really another term for justice. Although this argument may establish the rationality of justice, it does not make it a forward-looking consideration. Justice is the test of the thing that should be done regardless of consequences. It is a rectification demanded by the past rather than a “making better” determined by the future. This apparent opposition between two modes of moral justification is familiar and deep-seated. To the forward-looking theorist, only anticipations of a desirable future can justify punishment. To the theorist who invokes justice (the retributivist), any theory that does not entail just treatment is morally groundless. The very idea of a compromise that embraces all four goals is unacceptable to either side. A middle way is suspect to the extent that these methodological standpoints seem irreconcilable.

One can, of course, try to shrug the problem away. One can move to the level of intuitions about acceptable results in particular cases and embrace a kind of eclecticism in method, a middle way by default. This consists in middle-level generalizations from particular intuitions. For example, if deterrence is most effective when a minor but widespread offense like shoplifting is punished severely, justice may still require that punishments be scaled to what is fair even if deterrence is compromised. But a series of such observations may not yield much of

2. The term “utilitarianism” originated in the work of Jeremy Bentham. The forms of utilitarianism are described superbly well by J.J.C. Smart in Smart, Utilitarianism, 8 THE ENCYCLOPEDIA OF PHILOSOPHY 206-12 (1967). In the text accompanying note 54, infra, I discuss the relationship between the terms “utilitarian” and “forward-looking” as characterizations of theories.
a general theory even if it describes intuitions particularly well. It is the nature of theoretical investigation to view eclecticism with suspicion, to view it as capitulation in the face of difficulty. Eclecticism is the theorist's admission of failure unless the theorist's job is reconceived not as the job of finding a theory but of finding out whether it is possible to find a theory.6

The merits of eclecticism are peripheral to my main concerns. My questions will be methodological as well as substantive. They will draw upon a remark made by H.L.A. Hart in a context having nothing to do with the theory of punishment.7 In his recent criticism of Dworkin, Hart says that "a satisfactory foundation for a theory . . . will [not] be found as long as the search is conducted in the shadow of utilitarianism."8 Debates about punishment continue to be conducted in this shadow whether they involve its acceptance, its rejection, or the quest for a middle way between utilitarianism and retributivism. My critique of these debates has two parts. I shall first show how theoreticians gravitate to one or the other pole as a consequence of taking the opposition seriously. I shall then question the very idea that there is an irreconcilable opposition between the two polar modes of justification and give reasons for thinking it a myth.

II. FOUR THEORIES OF PUNISHMENT

Four influential theories of punishment are attempts to come to terms with the classical goals of criminal law. Three of the four are trapped on the horns of the methodological dilemma. They take seriously only one of the modes of justification and distort or disregard the other pole of debate.

A. Quinton: The Logic of Retributivism

In his important article, On Punishment,9 Anthony Quinton argues that retributivism has been misconceived. In his view, retributivism is reconcilable with forward-looking theories because it is simply making a logical point while forward looking theories make an ethical point, a point about justification. First, Quinton interprets retributivism as de-
manding that a person be guilty before he may be punished. He acknowledges that a pure forward-looking theory could dispense with this limitation. The overall economy of deterrence might be such that punishment is most effective when those who are generally believed to be guilty are punished (as scapegoats) along with those who are really guilty. For Quinton, any conflict is illusory because the retributivist does not intend his limitation to have ethical force. Quinton states that

the necessity of not punishing the innocent is not moral but logical. It is not, as some retributivists think, that we may not punish the innocent and ought only to punish the guilty, but that we cannot punish the innocent and must only punish the guilty. Of course, the suffering or harm in which punishment consists can be and is inflicted on innocent people but this is not punishment, it is judicial error or terrorism or . . . 'social surgery.' 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'.

Quinton's theory, however, fails because it misses the retributivist's point and confuses a description of the conditions of punishment with a justification of punishment. It is correct descriptively to say that guilt is a condition of punishment. The institution or practice of punishing has this as a pervasive and unexceptionable feature. But this is not the retributivist's point. The retributivist claims that to be justified morally, punishment may be imposed only on the guilty. He identifies justice with this feature and says that any practice that lacks it a fortiori lacks justification.

The same point can be made somewhat differently. Quinton writes as if restricting punishment to the guilty were merely a feature internal to the practice of punishment. This characterization, however, does not answer the utilitarian's challenge that punishment need not be limited to the guilty because the challenge is made from outside the practice and has the form, why this practice and not another? Quinton notwithstanding, the retributivist, like the utilitarian, makes a normative and not a logical point.

Quinton cannot succeed in defusing the retributivist claim or the utilitarian challenge because the retributivist's purpose cannot be to de-

10. Id. at 58-59.
scribe a contingent internal feature of the system of punishment. This point can be seen in terms of the link between retribution and justice. If the retributivist merely identified retribution with punishment conditioned on guilt, he could be interpreted as offering a description without normative intent. He could be uncommitted on whether retribution in this sense is desirable. But to describe the connection between punishment and guilt as a requirement of justice is to make a normative claim. If the connection is an essential feature of justice, then a system that lacks this feature is unjust. For the retributivist, the aspirations of a justificatory theory framed in nonretributive terms are incoherent. The retributivist, thus, leaves us not with the task of finding a justification other than justice, but with the task of understanding what it could possibly mean to justify a system of punishment while conceding the irrelevance of justice.

Quinton avoids the choice between the two modes of justification only by declining to identify the retributivist claim as one of justice. Once the retributivist claim is so identified, the abyss between the two modes seems unbridgeable. Quinton defends utilitarianism by interpreting retributivism as a noncompetitor.

B. Rawls: Rule-Utilitarianism

"Two Concepts of Rules,"11 John Rawls' early and influential article, directly addresses punishment. It is concerned with utilitarian justification and puts forth an ingenious model called "rule-utilitarianism."12 Under this model, "one must distinguish between justifying a practice as a system of rules to be applied and enforced, and justifying a particular action which falls under these rules; utilitarian arguments are appropriate with regard to questions about practices, while retributive arguments fit the application of particular rules to particular cases."13

Utilitarian justification has this limited role because the system itself is found to work better when actors within the system may consult only the rules and need not make utilitarian judgments. The obvious analogy is to games: the game of baseball serves utilitarian ends, but

13. Rawls, supra note 11, at 62.
only when the players follow the rules without claiming exceptions (take a fourth strike) whenever utility would prescribe the exception in the individual case.\textsuperscript{14} Rawls' theory suggests rules of justice are internal rules like the three strike limit. They cannot be waived for reasons of utility in the individual case, but only because the package of rules serves utility effectively as a package.\textsuperscript{16}

Rawls further explains this theory by comparing the role of judge and legislator.

[T]he judge and the legislator stand in different positions and look in different directions: one to the past, the other to the future. The justification of what the judge does, \textit{qua} judge, sounds like the retributive view; the justification of what the (ideal) legislator does, \textit{qua} legislator, sounds like the utilitarian view. Thus both views have a point . . . and one's initial confusion disappears once one sees that these views apply to persons holding different offices with different duties, and situated differently with respect to the system of rules that make up the criminal law . . . . One reconciles the two views by the time-honored device of making them apply to different situations.\textsuperscript{16}

Rawls, like Quinton, offers two levels of analysis, internal and external. Unlike Quinton, however, he does not confuse this distinction with the distinction between description and normative justification.\textsuperscript{17} By distinguishing the internal move of justifying a particular action within a practice (a particular application of a rule) from the external move of justifying the practice itself as a system of rules, Rawls allocates the two modes of justification to different domains. Justice (retribution) becomes a virtue of the application of rules, and forward-looking (utilitarian) considerations justify the choice of rules.

It is important to see why Rawls' theory fails. The theory makes the existence of just treatment a provisional consideration of a fully justified practice. Just treatment is a feature of a system that, \textit{qua} system, is justified by considerations other than justice, and these other

\textsuperscript{14} I discuss this analogy to games in Morawetz, \textit{The Concept of a Practice}, in \textit{Philosophical Studies} 209-26 (1973) and in T. Morawetz, \textit{The Philosophy of Law} 46-51 (1980).

\textsuperscript{15} See Rawls, \textit{supra} note 11, at 76-77. See also Urmson, \textit{supra} note 12. Urmson's discussion is full and clear.

\textsuperscript{16} Rawls, \textit{supra} note 11, at 63-65.

\textsuperscript{17} Rawls' distinction is between two compatible kinds of justification. He concedes the normative force of each kind.
(forward-looking) considerations have conceptual priority. As Rawls realizes by the time he writes *A Theory of Justice*,\(^{18}\) justice cannot be purely an intrasystemic criterion of normative criticism. If it makes sense to assess an application of a rule as just, it also makes sense to refer to the rule itself as just (in its general application). If there are just rules, there are also just systems of rules. The notion of moral justification is not severable, and the moral standpoints of intrasystemic participant and extrasystemic observer (judge and legislator) are not distinct.\(^{19}\) Justice makes the same claim on both.

Many theorists offer a different criticism of rule-utilitarianism. They complain that the rule-utilitarian could countenance a general practice of secretly punishing the innocent.\(^{20}\) According to this claim, Rawls could get counterintuitive results because punishing the innocent violates an intuition that we all accept: the innocent should not be punished. A rule-utilitarian like Rawls may reject the suggestion that he would countenance a system in which the innocent are officially sanctioned and the public is kept in ignorance of this in four ways. As a preliminary matter he might say that such examples are farfetched and unpersuasive and that bad examples make bad theory.\(^{21}\) He may then attack the example on its facts. He may say there is an ineradicable risk that such a policy would become known and that public confidence and obedience would be dramatically threatened. This reply, however, is pragmatic rather than principled, and seems to concede the objection. Under this defense, the rule-utilitarian would approve of the situation if in theory the risk could be eliminated.

The utilitarian's third and fourth responses are more interesting. The third response is that the weight that must be assigned to harming the innocent—the subjective disappointment, the pain of unfair treatment—is so great that it necessarily outweighs whatever benefits the practice may yield. The suggestion here is that even if the utilitarian adheres to the criteria of evaluation most often associated with utilitarian forms of justification and puts forth an account of justice as a concept to be deployed in criticism from an external standpoint. *Id.* at 22-53.

18. J. Rawls, *A Theory of Justice* (1971). By this time, Rawls has explicitly rejected utilitarian forms of justification and puts forth an account of justice as a concept to be deployed in criticism from an external standpoint. *Id.* at 22-53.

19. The inapplicability of Rawls' model to justification as it occurs in "open" practices like law (as opposed to "closed" practices like games) is the main subject of my paper, *The Concept of a Practice*, supra note 14.


arianism and speaks of maximizing happiness or satisfaction, it is open to him to make interpersonal comparisons in various ways. It is open, that is, to say that the disappointment of the punished innocents has such intensity and such qualities that it has overwhelming, decisive weight. The final response is for the utilitarian to adopt nonstandard criteria of evaluation and to speak not of happiness and subjective satisfaction, but of well-being and the common good. He is then in position to conclude that a regime of injustice cannot serve the common good.

The third and fourth responses are opening shots in a theoretical dialogue. The third response, it will be objected, throws justice out the front door only to admit it through the back by saying that special and decisive weight should be assigned by the utilitarian to the sense of justice outraged. The final response merely puts in question what terms of utility determination or forward-looking evaluation are satisfactory.

C. Hart: A Multiplicity of Values

Both Quinton and Rawls seek to accommodate the notion of justice within a utilitarian framework. Although Rawls' theory fares better than Quinton's, both undervalue and misrepresent the apparent role of justice in normative reasoning. H.L.A. Hart, in his extensive writings on punishment and criminal responsibility, offers a theory superficially similar to Rawls' but one that is more sensitive to the imperatives of the competing modes of justification.

Hart says that a "morally tolerable" theory of punishment must be "a compromise between distinct and partly conflicting principles." Hart, therefore, distinguishes the "general justifying aim" of punishment—crime minimization through general as well as specific deterrence—from its secondary aim, the distribution of punishment. Justice is relevant to only the second of these problems while deterrence is

22. Utilitarianism, as propounded by Mill and Bentham, speaks of the greatest happiness for the greatest number. In the twentieth century, utilitarians have frequently refined these references to ones about "want satisfaction" and "preference satisfaction." See, e.g., Gauthier, On the Refutation of Utilitarianism, in The Limits of Utilitarianism 144-63 (H. Miller & W. Williams eds. 1982).

23. See T. Morawetz, supra note 14, at 101-06.


25. Id. at 1.

26. Id. at 8-13. Hart is usually careful to be agnostic about whether the general justifying aim of criminal law goes beyond deterrence and crime minimization to include other goals and values.
relevant to both. Thus, in particular cases, the degree of punishment must be determined by a variety of criteria so that it is exemplary, proportionate, and fair all at the same time.

Commentators have found Hart's theory ambiguous. Certain passages make Hart sound much like a rule-utilitarian. He says, for example, that

\[
\text{[f]airness between different offenders expressed in terms of different punishments is not an end in itself, but a method of pursuing other aims which has indeed a moral claim on our attention; and we should give effect to it where it does not impede the pursuit of the main aims of punishment.}^{27}
\]

Hart adds that "justice is a method of doing other things, not a substantive end."\(^{28}\) These passages suggest that deterrence is the end while justice is the means. They also suggest that deterrence is the external justification while justice remains a counter used internally by players of the game. Just treatment is justified only as long as, and to the extent that, it serves the external end.

This Rawlsian reading, however, is not the most plausible reading of Hart, who also argues time and again that the unwillingness to be unjust, for example to punish scapegoats in the interest of forward-looking considerations like deterrence, is based on independent values. This unwillingness "would still remain even if we were certain that in the case of the 'punishment' of one who had not broken the law the fact of his innocence would not get out or would not cause great alarm if it did."\(^{29}\)

Hart's theory is distinguishable from Rawls' view because it makes use of what Nozick calls "side constraints."\(^{30}\) Under this view, utilitarian goals cannot be pursued in ways that violate justice, not because rules of justice are provisionally or contingently the best rules for pursuing utility, but because they represent an independent value that limits such pursuit. Hart concludes that "the pursuit of a single social aim . . . has its restrictive qualifier" and that "our main social institutions always possess a plurality of features which can only be understood as

\[\text{27. Id. at 172-73.}^{27}\]
\[\text{29. H.L.A. Hart, supra note 24, at 77-78.}^{29}\]
\text{constraint view forbids you to violate . . . moral constraints in the pursuit of your goals." Id. at}\
\text{29.}^{30}\]
a compromise between partly discrepant principles." Most simply, this means that the distinctive aim of criminal law, the control and minimization of crime, is pursued in the context of other values. Although these values are affected by what we do with criminal law, they are not unique to or distinctive of criminal law.

D. An Expanded Version of Retribution

The three theories considered so far all identify one theme with retributivism: punishment must be fair or just, and therefore limited to those who are guilty. This theme is offered as a constraint on the pursuit of forward-looking goals. In other words, under these theories fairness is a necessary condition for punishing, but it is not a sufficient condition. Forward-looking goals must also be served.

A true retributivist, on the other hand, argues that considerations of desert (or justice) are sufficient to justify punishing. Thus, while Hart clearly invokes justice to safeguard the rights of potential subjects of punishment, a true retributivist invokes retributivism to emphasize the right of "society" to punish offenders with a suitable degree of severity. As we have seen, Hart sees guilt as a threshold condition for punishment. For him, the application of punishment must not only be justly deserved (i.e., only the guilty may be punished and only with severity no greater than the seriousness of the offense), but it must also serve the "general justifying aim" of furthering common ends.

In order to evaluate retributivism we must try to separate the principle of retribution from the principle of justice. There is no need to do so if we are considering only the requirement that punishment be limited to the guilty; this may be called either a requirement of retribution or one of justice. But we must consider two additional requirements often identified with the notion of retribution, namely (a) that the state be seen as having a duty to punish the guilty and (b) that intuitions of proportionality set a lower limit as well as an upper limit to the severity of punishment. Although these requirements are implicit in the notion of retribution, they are, as we shall see, not necessarily requirements of justice.

Three aspects of this expanded version of retributivism are imme-

31. H.L.A. HART, PUNISHMENT AND RESPONSIBILITY, supra note 24, at 10 (emphasis added).
33. See, e.g., T. HONDERICH, supra note 3, at 22-23.
diately apparent.

(1) This scheme makes no appeal to forward-looking considerations as justification. It may be said to have benefits from a forward-looking standpoint, but it is not justified in terms of those benefits.

(2) These principles are echoed in statutory schemes of sentencing, in the announced aims of sentencers, and in the expectations of laypersons.\(^3\) It is often said, virtually as a truism, that the guilty should be punished and punished with sentences they deserve.

(3) These principles do not match legal practice, where many norms flout retributive guidelines. Defendants may buy immunity from prosecution by choosing to incriminate their associates. Prosecutions are dropped and convictions overturned when the state violates the statutory or constitutional rights of "guilty" defendants.\(^3\) The first of these practices may be justified only by expediency, but the latter is typical of a large class of practices justified by accepted principle.

The expanded version of retributivism carries a long tail of controversy; it is easy to see why. Although the processes of pleading and trial are designed to separate the guilty from the innocent, other aspects of criminal law show the artificiality and the idealization involved in this separation. Criminal law everywhere recognizes a range of excusing and mitigating conditions, and it recognizes that each condition expresses a heterogeneous mix of situations and dispositions.\(^3\) The awareness that many offenders are at the cusp of excusability should make us wary of claiming that there is a duty to punish the guilty as a discrete and isolable class. The same awareness of heterogeneity should make us see how unclear it is to claim that the punishment should fit the seriousness of the offense. Is seriousness measured by the amount of harm caused, the amount of harm intended, the atrocious character of the offense as opposed to the degree of harm,\(^3\) or the degree of commitment, ambivalence, or remorse of the offender?

These problems are not fatal to the retributivist view. We all in

34. Consider, for example, § 1170 of the CAL. PENAL CODE (1976 amendment) (West Supp. 1984): "The Legislature finds and declares that the purpose of imprisonment for crime is punishment. This purpose is best served by terms proportionate to the seriousness of the offense with provision for uniformity in the sentences of offenders committing the same offense under similar circumstances."

35. The development of this practice under the U.S. Constitution is documented and explained in MODERN CRIMINAL PROCEDURE chs. 5, 10, & 11 (Y. Kamisar, W. LaFave & J. Israel eds., 5th ed. 1980).


37. See J. KLEINIG, supra note 4, at 123-29.
fact rank instances of offensive conduct using the language of comparative seriousness. We sympathize with the duty to punish the guilty as long as we understand it to refer to those who are morally as well as legally guilty. Jeffrie Murphy has argued that both utilitarianism and retributivism are based on evident principle and not merely disputable intuitions or emotions. "That the maximization of social utility is important is no more obviously true than that a man should not unfairly profit from his own criminal wrongdoing. . . ." Murphy goes on to remark that retributivism draws its theoretical underpinnings from that aspect of justice that deals with reciprocity. Kant also identified retribution with reciprocity when he argued that criminal law exists to prevent those who disobey the rules of conduct from gaining an unfair advantage over those who obey.

The Kantian view has not stood unchallenged. We have seen that retributivism and the underlying theory of reciprocity underestimate the mix of offenders. Not all offenders seek to profit from crime, and some who seek to profit do not achieve any gain, even when punishment is left wholly out of account. The notion of reciprocity between the criminal and society or between the criminal and her victim is a metaphor drawn from a relationship that is both simpler and different. Although justice requires reciprocity in the case of an actual compact or within a small group with voluntary and mutual undertakings, it is much less clear what justice requires under the conditions of an actual criminal offense.

In this section I have indicated the gravitational force of the utilitarian and retributive modes of moral justification, and the difficulty of imagining and defending a middle way. Quinton and Rawls exemplify the utilitarian or forward-looking mode while Hart offers an account that gives sound attention to both modes without trying to reconcile them at the theoretical level. Quinton and Rawls give insufficient attention to the extrasystemic role of justice. Retributivism, in the form considered here, may be at odds with justice. Although justice is congruent with the weak form of retributivism which merely asserts that only the guilty should be eligible for punishment, it is not obviously congruent

38. Id. at 115-29.
40. I. Kant, Metaphysical Elements of Justice 36-39 (J. Ladd trans. 1965) (translation of Metaphysische Anfangsgrunde der Rechtslehre (1797)).
41. The use of such metaphors and of this style of reasoning is criticized persuasively by M.B.E. Smith in Smith, Is There a Prima Facie Obligation to Obey the Law?, 82 Yale L.J. 950 (1973).
with a stronger form of retributivism, which asserts that there is a duty to punish the guilty with fitting severity.

III. **Toward a New Methodology**

The opposition between forward-looking justifications and justice as justification can be bridged not by eclecticism but by reconceiving the terms of a forward-looking theory. Only in this way can reasoning about punishment emerge from "the shadow of utilitarianism," from a context in which familiar forms of utilitarianism are swallowed whole, rejected, or adopted with demarcated constraints.

A forward-looking theory will be incomplete as long as it fails to consider every effect the practice of punishing has on the common good. The effects of criminal adjudication and punishment on the incidence of crime are surely the first effects that come to mind, but there are others. Hart reminds us that the public process of trial and conviction affirms values that bind society. A society in which stable standards of punishment generally reflect and reinforce moral values is one in which persons rest secure in their capacity to act and plan, and in which they experience little discontinuity between public events and private expectations. The notion of the common good is not severable from the notion of a society with these latter characteristics.

The best discussion of the social benefits of punishment is in Joel Feinberg's essays, "Justice and Personal Desert" and "The Expressive Function of Punishment." Feinberg draws attention to what he calls natural responsive attitudes to the actions of others. For example, he notes the congruity—logical as well as emotive—between gratitude and kindness, resentment and intentional injury. Feinberg asserts that there is a natural fit between "one person's actions or qualities and another person's responsive attitudes." These responsive attitudes are the foundations of a shared sense of what treatment persons deserve; they owe nothing to forward-looking goals. But, Feinberg notes, consideration of the common good "give[s] a reason (in addition to natural inclination) for expressing . . . our attitudes and appraisals" in such institutions as criminal adjudication and punishment. Individual responsive attitudes and institutional criteria for decision-making mutually reinforce and condition one another. Feinberg notes that whether

---

44. *Id.* at 82.
45. *Id.* at 83.
or not symbolic public disapproval helps or hinders deterrence and reform, it serves other functions that are inherent in such expression. Among these functions are authoritative disavowal of certain kinds of harmful conduct, vindication of the law as rooted in moral values, and formal absolution of those who are free of blameworthy conduct. It is easy but wrong to think of these functions as means to an end and as needing vindication in the light of some aspect of the common good. The following argument demonstrates why no such vindication is needed.

First, there is an important distinction between resentment and justified disapproval. Unlike resentment, justified disapproval carries with it a claim to be able to justify or demonstrate that the disapproved conduct violates shared norms of mutual respect and dignity. Resentment is a personal responsive attitude toward actions affecting oneself, while disapproval is a responsive attitude backed by reasons and concerned with actions affecting oneself or others. Resentment and justified disapproval are different categories of response—one emotive, one cognitive—but they are responses to the same kind of action. Second, there is a logical relationship between resentment or disapproval on one hand and certain kinds of conduct on the other. What we resent or disapprove of (and what we generally praise or blame) is not typically arbitrary or a matter of idiosyncratic preference. Rather such judgments are appeals to shared standards, to a shared sense of the common good. This is what John Rawls taps when he finds and describes the "sense of justice." Third, any forward-looking characterization of the common good will describe social institutions in which the sense of justice is realized and reinforced, not as a means to some other end but as an indispensable aspect of living well with others. Finally, the requirement that a system of punishment be just (that punishment be allocated to the guilty and that moral blameworthiness play some role in determining the severity of punishment) is an inherent requirement of any characterization of the common good.

If all this is true, one of the recurrent criticisms of forward-looking theories—that such theories fail because a system justified in forward-

46. *Id.* at 101-04.
48. *Id.* at 13-16.
looking terms could encompass conviction and punishment of the innocent if it was generally believed that those punished were guilty—misses the mark. This criticism is irrelevant because a forward-looking theory is an attempt to describe a future in which the common good is in fact secured, not one in which persons happen to believe that the common good is secured. The criticism would suggest itself only to someone who thought that there was no difference between reality and belief about reality. The difference, however, could hardly be greater. One would not confuse a situation in which persons wrongly believed that they were safe from crime with one in which potential offenders were really deterred. One should, accordingly, keep distinct a situation in which persons believed they lived in a just society from one in which they were in fact treated justly.

The point here is methodological. To understand it we must distinguish conceptual relationships from contingent ones. For example, the relationship between the common good and security from harm is conceptual, not contingent. It is part of the notion of the common good that persons be safe and secure, and the inquiry into why deterrence is a desirable goal is otiose. The relationship between the common good and just treatment is also conceptual and not contingent. The question why punishment should be limited to the guilty requires no answer. The attempt to refute the criticism on its own terms by showing that systematic injustice would not (as a contingent matter) be balanced by gains from enhanced deterrence, is misconceived ab initio. It is misconceived not because the empirical question is close or undecidable, but because the empirical question is not meaningfully raised; gains in added safety are, in principle, not the sorts of things that can justify

50. See supra notes 20-23 and accompanying text.

51. I am drawing on the familiar and highly problematic distinction between analytic (conceptual) truths and synthetic (contingent, empirical) truths. Very roughly, the distinction is between propositions whose truth is apprehended through one's understanding of the meanings and uses of the terms and to propositions whose truth is determined by inspection of empirical situations. "Ethics is a branch of philosophy" is a proposition of the first kind while "Harold Stassen is the Republican nominee for President" is an example of the second. The distinction has spawned endless debate among philosophers. Among the most influential discussions are W. Quine, Two Dogmas of Empiricism, in FROM A LOGICAL POINT OF VIEW 20-46 (1961) and Grice & Strawson, In Defense of a Dogma, 65 PHIL. REV. 141-58 (1956).

52. Of course, one can describe the ways in which persons benefit from safety and security, but those descriptions do not have explanatory value. If the value of safety and freedom from harm is not self-evident, there's little to be said to demonstrate their value. Compare Hart's discussion of the minimal content of natural law in H.L.A. HART, THE CONCEPT OF LAW 189-95 (1961).
systematic injustice.\(^53\)

The familiar opposition between maximizing the collective happiness or satisfaction and satisfying justice is a chimera. Happiness and satisfaction are understandable and realizable not in vacuo but in a social context. The kind of happiness or satisfaction realizable in a just context is not comparable or commensurable with happiness or satisfaction realizable in an unjust context. The kind of happiness that is relevant to forward-looking anticipations of the common good is necessarily the former.

The term “forward-looking theory of the common good” is preferable to “utilitarianism” because the latter is tainted. Throughout its tangled history utilitarianism has been associated with the notion that happiness or satisfaction is only contingently or accidentally related to the moral features of a social context. A context in which the greatest happiness is gained by the greatest number may or may not be a just society; its justice remains to be demonstrated.\(^54\) In Hart’s most recent work he sees, I think, the vulnerability of traditional utilitarianism. He says that we must move away from debate about “the ways in which utilitarianism has ignored certain values taken to be uncontroversial.”\(^55\) At the same time, he calls for “a more radical and detailed consideration of the ways in which rights relate to other values.”\(^56\) My remarks are intended to be in the spirit of these observations.

It may be objected that any attempt to integrate justice into a forward-looking theory must fail. He who takes justice seriously is oblivious to consequences and affirms that justice be done even if the heavens fall.\(^57\) This objection fails in two ways. First, appearances notwithstanding, the objection does not set justice in opposition to future consequences, but rather sets in opposition two different sets of consequences, the effects of just treatment and the effects of unjust treatment. In this sense, the objection itself is forward-looking. Secondly, the on-going debate about justice and forward-looking theories is con-
cerned with on-going systems of rules (practices), not particular cases of injustice. The objection would have to be rewritten to refer to just rules, not to justice in a particular case. That said, one who holds the view I have defended may concede that individual cases of injustice may occur and need to be excused; but in principle they must remain exceptions within a system of justice. If the objection is that the heavens would fall were justice done systematically, i.e., were just rules adopted, then it adopts a pessimism to which there is no reasonable response.

It could also be objected that integrating justice into a forward-looking theory absorbs the retributivist’s notion that there is a duty to punish the guilty and to do so proportionately. This objection forgets the caveats laid out above. Justice, in my view, speaks uncertainly to such issues. It may counsel us to distinguish among the guilty in the light of circumstances and to weigh ameliorating or aggravating circumstances in deciding when and how to punish. The following propositions illustrate this point.

(1) A forward-looking theory of the common good will prescribe a system of just rather than unjust treatment.

This, I have argued, is a conceptual matter.

(2) A just system of punishment will restrict punishment to the guilty.

This too is a conceptual matter and not an empirical question.

(3)(a) A just system of punishment will always punish the guilty and will punish them with severity proportionate to the seriousness of their offenses.

This proposition is treated as conceptual by retributivists. Compare it with the following.

(3)(b) In a just system of punishment there are ordinarily good reasons to punish the guilty and to do so with severity proportionate to the seriousness of their offenses.

I have suggested that (3)(b), and not (3)(a), best describes the requirements of justice. It follows from (3)(b) that whether and how an offender should be punished in a particular case is a matter to be decided on the basis of all relevant reasons.

---

58. It can be argued, I think, that this best fits Hart’s position as well.
IV. QUESTIONING THE GENEALOGY OF MORALS

The paper has one unexplained premise. It is that the most familiar way of characterizing a forward-looking theory, identified over the decades with utilitarianism, is incoherent. To defend my premise would bring us deeply into the epistemology of moral reasoning and is beyond the scope of this paper. Nonetheless, an underlying characterization of forward-looking theories, shared by both most utilitarians and anti-utilitarians, can and must be described, however briefly and roughly. By its terms, individuals are seen as entertaining life plans involving identifiable ends (plans for maximizing happiness or satisfaction) atomistically. The ends reflect personal choices or dispositions, and they may have any content. For the utilitarian, the point of moral reasoning is to ascribe moral justification to the coordination of such plans to maximize collective and cumulative satisfaction. For the anti-utilitarian, the point of moral reasoning is to uncover the appropriate constraints on the pursuit of these ends so that each individual is accorded the protection of respectful treatment. These constraints are, at least in part, the constraints of justice. For the anti-utilitarian, moral justification lies in satisfaction of the conditions set by such constraints.

To begin to see the incoherence of the traditional characterization one must challenge the underlying picture of human experience and aspiration. One must sketch an alternative picture, one in which the genealogy of personal awareness of the good lies ab initio in social interaction and is describable essentially in interpersonal terms of reciprocity. It would follow that the notion of a personal good is best seen as derived from participation in the common good and that the common good inherently involves, as I have argued, a system of just rather than unjust treatment. A move away from the atomistic epistemological premise is a move away from the utilitarian dialogue, a move (as Hart would have it) out of the shadow of utilitarianism. Its elaboration is a project for another time.

59. See, e.g., J.S. Mill, Utilitarianism (1861); J. Bentham, Principles of Morals and Legislation (1789); R. Brandt, A Theory of the Good and the Right (1979); and D. Lyons, Forms and Limits of Utilitarianism (1965).

60. See, e.g., R. Dworkin, supra note 32, at 94-100; J. Rawls, supra note 18, at 22-53.

61. In other words, each person is seen as determining what shall count as her/his personal goals and as going about maximizing the satisfaction of these goals. One individual's goals may or may not include the empathetic realization of the goals of others to whom she/he has affective ties.