PDF version of the entry
Punishment
https://plato.stanford.edu/archives/win2017/entries/punishment/

from the Winter 2017 Edition of the

Stanford Encyclopedia of Philosophy

Edward N. Zalta, Principal Editor
Uri Nodelman, Senior Editor
Colin Allen, Associate Editor
R. Lanier Anderson, Faculty Sponsor

Editorial Board
https://plato.stanford.edu/board.html

Library of Congress Catalog Data
ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit https://leibniz.stanford.edu/friends/ .

Stanford Encyclopedia of Philosophy
Copyright © 2017 by the publisher
The Metaphysics Research Lab
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305

Punishment
Copyright © 2017 by the authors
Hugo Adam Bedau and Erin Kelly

All rights reserved.

Copyright policy: https://leibniz.stanford.edu/friends/info/copyright/

Punishment

First published Fri Jun 13, 2003; substantive revision Fri Jul 31, 2015

The concept of punishment—its definition—and its practical application and justification during the past half-century have shown a marked drift away from efforts to reform and rehabilitate offenders in favor of retribution and incarceration. Punishment in its very conception is now acknowledged to be an inherently retributive practice, whatever may be the further role of retribution as a (or the) justification or goal of punishment. A liberal justification of punishment would proceed by showing that society needs the threat and the practice of punishment, because the goal of social order cannot be achieved otherwise and because it is unfair to expect victims of criminal aggression to bear the cost of their victimization. Constraints on the use of threatened punishments (such as due process of law) are of course necessary, given the ways in which authority and power can be abused. Such a justification involves both deontological as well as consequentialist considerations.

---

1. Background

Philosophical reflection on punishment has helped cause, and is itself partially an effect of, developments in the understanding of punishment that have taken place outside the academy in the real world of political life. A generation ago sociologists, criminologists, and penologists became disenchanted with the rehabilitative effects (as measured by reductions in offender recidivism) of programs conducted in prisons aimed at this end (Martinson 1974). This disenchantment led to skepticism about the feasibility of the very aim of rehabilitation within the framework of existing penal philosophy. To these were added skepticism over the deterrent effects of punishment (whether special, aimed at the offender, or general, aimed at the public) and as an effective goal to pursue in punishment. That left, apparently, only two possible rational aims to pursue in the practice of punishment under law: Social defense through incarceration, and retributivism. Public policy advocates insisted that the best thing to do with convicted offenders was to imprison them, in the belief that the most economical way to reduce crime was to incapacitate known recidivists via incarceration, or even death (Wilson 1975). Whatever else may be true, this aim at least has been achieved on a breathtaking scale, as the enormous growth in the number of state and federal prisoners in the United States (some 2.3 million in year 2015, including over 3,000 on "death row") attests.

At the same time that enthusiasm for incarceration and incapacitation was growing as the preferred methods of punishment, dissatisfaction with the indeterminate prison sentence—crucial to any rehabilitative scheme because of the discretion it grants to penal officials—on grounds of fairness led policy analysts to search for another approach. Fairness in sentencing seemed most likely to be achievable if a criminal sentence was of a determinate rather than indeterminate duration (Allen 1981). But even determinate sentencing would not be fair unless the sentences so authorized were the punishments that convicted offenders *deserved*. Thus was born the doctrine of "just deserts" in sentencing, which effectively

combined the two ideas.[1] By this route the goals of incapacitation and retribution came to dominate, and in some quarters completely supersede, the goals of rehabilitation and deterrence in the minds of politicians and social theorists.

Concurrently with these broadly socio-legal developments (to which might be added the despair of practitioners that reached its peak with the police assault on rioting prisoners in New York's Attica prison in 1972) philosophers were crafting their own arguments, reviving classic views associated with the names of Kant and Hegel to establish two principal ideas that fit surprisingly well with those reviewed above. First, philosophers urged that reformation of convicted offenders (especially in its more medically inspired modes, vividly depicted in fictionalized form in Anthony Burgess's *Clockwork Orange*[1962]), is not the aim, or even a subsidiary aim among several, of the practice of punishment. Aside from being an impractical goal, it is morally defective for two reasons: It fails to respect the convicted offenders' autonomy, and it flouts the offenders' *right* to be punished for the wrongdoing he intentionally caused (Morris 1968). (The oddity of a theory that affirms having and exercising a right to be punished has not escaped notice.) Second, justice or fairness in punishment is the essential task of sentencing, and a just sentence takes its character from the culpability of the offender and the harm the crime caused the victim and society (Card 1973, von Hirsch 1985, Nozick 1981: 366–74). In short, just punishment is retributive punishment. Philosophers reached these conclusions because they argued that there were irreducible retributive aspects to punishment—in the very definition of the practice, in the norms governing justice in punishment, and in the purpose of the practice as well.

As a result, the ground was cut out from under the dominant penal policy of mid-century, the indeterminate sentence in the service of the rehabilitative ideal for offenders behind bars. Probation as the essential nonincarcerative alternative sanction received an expanded role, but release on parole came to a virtual end. In its place (but as it turned out, only in theory) was uniform determinate sentencing, which would avoid the follies of unachievable rehabilitative goals and ensure both incapacitation and even-handed justice for all offenders. (This was, of course, before the political process distorted these aims. Not all admirers of justice in punishment supported determinate sentencing.) The culmination of this trend appears in the Sentencing Reform Act of 1984, which spawned the United States Sentencing Commission and its Federal Sentencing Guidelines. The doctrine has not been without its critics, both in theory and in practice (Zimring 1977). But to date, no alternative approach shows any signs of supplementing the just deserts sentencing philosophy—no matter how preposterous in practice the claim that a given punitive sentence is justly deserved may be in most cases.

There has been a third development concurrent with the two outlined above, far less influential in the formation of actual penalty policy even if it is of equal theoretical importance (Harding 1989). We refer to the reconceptualization of the practice of punishment arising from the work of Michel Foucault in the mid-1970s. Foucault invited us to view the practice of punishment under law as subject to general forces in society that reflect the dominant forms of social and political power—the power to threaten, coerce, suppress, destroy, transform—that prevail in any given epoch. And he also cultivated a deep suspicion toward the claims that contemporary society had significantly humanized the forms of punishment by abandoning the savage corporal brutality that prevailed in the bad old days, in favor of the hidden concrete-and-steel carceral system of the modern era (Foucault 1977).

Foucault's insights arose from a historical, socioeconomic, and psychodynamic approach to punishment. Professed goals of punishment, norms constraining the use of power in the pursuit of these goals, the

aspiration for justice in punishment—all these, if Foucault is right, turn out to mask other (not necessarily conscious) intentions among reformers that belie the ostensible rationality (not to say rationalization) of their aims since the Enlightenment. Thus, the movement against capital punishment in the late eighteenth century is not to be explained (or, presumably, justified) by the influence of conscious, rational utilitarian calculations of the sort that Beccaria and Bentham argued had persuaded them to oppose the death penalty (Bedau 1983, Maestro 1973). It is explained instead by disenchantment with the theatrical, dramaturgical, aspects of public executions and a self-deceiving humanitarian impulse that merely shifted but otherwise left unaltered the nature and locus of the power wielded over criminals by society—perfectly embodied in Bentham's visionary carceral scheme, the notorious Panopticon prison (Semple 1993).

Two features at least of Foucault's explorations into the practice of punishment in Western society deserve mention here. First, he ignored the analytical distinctions that philosophers in the Anglo-American tradition had made familiar (to be discussed below). None plays any visible role in his account of the theory or practice of punishment. Some interpreters might not only acknowledge this, they would go further and argue that Foucault offers no philosophical views about punishment at all—because conceptual and normative analysis and the search for principles on which to rest policy are at best obscurely and indirectly pursued in his writings. Instead, so this interpretation declares, he is just a social commentator (or some other form of critical humanist) (Garland 1990). But this interpretation fails to do him justice. Foucault's views are, at least in part, unmistakably philosophical. Not only do they issue in claims that are not obviously testable empirical hypotheses, they involve large-scale reflections on and reinterpretations of human nature, public institutions, and the point of our punitive practices.

Second, Foucault implicitly challenges the very idea of any form of justification of the practice of punishment. He is, in his way, a paradigmatic thinker whose views about punishment can be called anti-foundationalist. What emerges from his account is the view that what passes for the justification of punishment (as with any other social practice) is inextricably tied up with assumptions, beliefs—in short, with ideology—that have no independent rational foundation. The very idea that penal institutions can be justified is suspect, self-delusive. Foucault more than any other recent thinker who has reflected on the institutions of punishment in western society, has brought historicist, anti-analytic, and anti-foundationalist convictions together, thus sowing deep uncertainty over how and even whether to address the task of justifying punishment.

In all these respects, Foucault must be seen as the modern successor to Friedrich Nietzsche—Foucault's great albeit unacknowledged predecessor in the philosophy of punishment. More than any thinker before or since, Nietzsche understood the way punishment is "overdetermined by utilities of every sort" and survives now under this, now under that interpretation of its purposes—because the desire to punish (and thereby subordinate, coerce, transform) other persons is so deeply rooted in human nature (Nietzsche 1887).

The cumulative effect of these forces, political and intellectual, has been to undermine confidence in the classic Enlightenment or liberal view of punishment found, for example, in Hobbes, Locke, Bentham, and Mill. Perhaps this is an exaggeration; one might argue that since it is unclear just what a liberal view of punishment really is, successfully undermining it is equally uncertain. Liberalism in punishment, it is true, has no canonical formulation; instead, it has been multiply ambiguous during its career of more three centuries, as scrutiny of Beccaria's influential proposals for reform at the zenith of the Enlightenment show (Beccaria 1764). What is needed is a reassertion, reformulation, and redeployment of

recognizably liberal ideas in the theory of punishment (see the discussion below).

2. Theory of Punishment

The prevailing features in the modern theory of punishment were developed by analytic philosophers half a century ago. The theory in the Anglo-American philosophical world was and still is governed by a small handful of basic conceptual distinctions, self-consciously deployed by virtually all theorists no matter what substantive views they also hold about punishment. The terminus a quo of these ideas are the influential writings of H.L.A. Hart (1959) in England and John Rawls (1955) in the United States. Though both Hart and Rawls pass muster as centrist liberals, they believed these analytic distinctions to be ideologically neutral.

- *Defining* the concept of punishment must be kept distinct from *justifying* punishment. A definition of punishment is, or ought to be, value-neutral, at least to the extent of not incorporating any norms or principles that surreptitiously tend to justify whatever falls under the definition itself. To put this another way, punishment is not supposed to be justified, or even partly justified, by packing its definition in a manner that virtually guarantees that whatever counts as punishment is automatically justified. (Conversely, its definition ought not to preclude its justification.)

- Justifying the *practice* or institution of punishment must be kept distinct from justifying any given *act* of punishment. For one thing, it is possible to have a practice of punishment—an authorized and legitimate threat system—ready and waiting without having any occasion to inflict its threatened punishment on anyone (because, for example, there are no crimes or no convicted and sentenced criminals). For another, allowance must be made for the possibility that the practice of punishment might be justified even though a given act of punishment—an application of the practice—is not.

- Justification of any act of punishment is to be done by reference to the *norms* (rules, standards, principles) defining the institutional practice—such as the classic norms of Roman law, *nullum crimen sine lege* and *nulla poena sine lege* (no crime without law, no punishment without law). Justification of the practice itself, however, necessarily has reference to very different considerations—social purposes, values, or goals of the community in which the practice is rooted. The values and considerations appropriate to justifying acts are often assimilated to those that define judicial responsibility, whereas the values that bear on justifying the punitive institution are akin to those that govern statutory enactments by a legislature.

- The practice of punishment must be justified by reference either to forward-looking or to backward-looking considerations. If the former prevail, then the theory is likely to be *consequentialist* and probably some version of utilitarianism, according to which the point of the practice of punishment is to increase overall net social welfare by reducing (ideally, preventing) crime. If the latter prevail, the theory is *deontological*; on this approach, punishment is seen either as a good in itself or as a practice required by justice, thus making a direct claim on our allegiance. A deontological justification of punishment is likely to be a retributive justification. Or, as a third alternative, the justification of the practice may be found in some hybrid combination of these two independent alternatives. Attempts to avoid this duality in favor of a completely different approach have yet to meet with much success (Goldman 1982, Hoekema 1986, Hampton 1984, Ten 1987, von Hirsch 1993, Tadros 2013).

Acknowledgment of these distinctions seems to be essential to anything that might be regarded as a tolerably adequate theory of punishment.

Two substantive conclusions have been reached by most philosophers based in part on these considerations. First, although it is possible to criticize the legitimacy or appropriateness of various individual punitive acts—many are no doubt excessive, brutal, and undeserved—the practice of punishment itself is clearly justified, and in particular justified by the norms of a liberal constitutional democracy. Second, this justification requires some accommodation to consequentialist as well as to deontological considerations. A strait-laced purely retributive theory of punishment is as unsatisfactory as a purely consequentialist theory with its counter-intuitive conclusions (especially as regards punishing the innocent). The practice of punishment, to put the point another way, rests on a plurality of values, not on some one value to the exclusion of all others.

So much by way of review of the recent past as a stage setting for what follows—a sketch of what we take to be the best general approach to the problem of defining and justifying punishment.

Justifications of Punishment As a first step we need a definition of punishment in light of the considerations mentioned above. Can a definition be proposed that meets the test of neutrality (that is, does not prejudge any policy question)? Consider this: Punishment under law (punishment of children in the home, of students in schools, etc., being marginal rather than paradigmatic) is *the authorized imposition of deprivations—of freedom or privacy or other goods to which the person otherwise has a right, or the imposition of special burdens—because the person has been found guilty of some criminal violation, typically (though not invariably) involving harm to the innocent*. (The classical formulation, conspicuous in Hobbes, for example, defines punishment by reference to imposing pain rather than to deprivations.) This definition, although imperfect because of its brevity, does allow us to bring out several essential points. First, punishment is an authorized act, not an incidental or accidental harm. It is an act of the political authority having jurisdiction in the community where the harmful wrong occurred.

Second, punishment is constituted by imposing some burden or by some form of deprivation or by withholding some benefit. Specifying the deprivation as a deprivation of *rights* (which rights is controversial but that controversy does not affect the main point) is a helpful reminder that a crime is (among other things) a violation of the victim's rights, and the harm thus done is akin to the kind of harm a punishment does. Deprivation has no covert or subjective reference; punishment is an objectively judged loss or burden imposed on a convicted offender.

Third, punishment is a human institution, not a natural event outside human purposes, intentions, and acts. Its practice requires persons to be cast in various socially defined roles according to public rules. Harms of various sorts may befall a wrong-doer, but they do not count as punishment except in an extended sense unless they are inflicted by personal agency.

Fourth, punishment is imposed on persons who are believed to have acted wrongly (the basis and adequacy of such belief in any given case may be open to dispute). Being *found* guilty by persons authorized to make such a finding, and based on their belief in the person's guilt, is a necessary condition of justified punishment. Actually *being* guilty is not. (For this reason it is possible to punish the innocent and undeserving without being unjust.)

Fifth, no single explicit purpose or aim is built by definition into the practice of punishment. The practice, as Nietzsche was the first to notice,

is consistent with several functions or purposes (it is not consistent with having no purposes or functions whatever).

Sixth, not all socially authorized deprivations count as punishments; the only deprivations inflicted on a person that count are those imposed in consequence of a finding of criminal guilt (rather than guilt only of a tort or a contract violation, or being subject to a licensing charge or to a tax). What marks out nonpunitive deprivations from the punitive ones is that they do not express social condemnation (Feinberg 1965, Bedau 2001). This expression is internal, not external, to the practice of punishment.

Finally, although the practice of punishment under law may be the very perfection of punishment in human experience, most of us learn about punishment well before any encounters with the law. Thus, "authorized deprivation" must not be so narrowly interpreted as to rule out parental or other forms of "punishment" familiar to children, even though those deprivations are often ambiguous in ways that punishment under law is not.

It is helpful in assessing various candidate justifications of punishment to keep in mind the reasons why punishment needs to be justified.

- Punishment—especially punishment under law, by officers of the government—is (as noted above) a human institution, not a natural fact. It is deliberately and intentionally organized and practiced. Yet it is not a basic social institution that every conceivable society must have. It is a testimony to human frailty, not to the conditions necessary to implement human social cooperation. It also has no more than an historical or biological affinity with retaliatory harm or other aggressive acts to be found among nonhuman animals or (despite thinkers from Bishop Joseph Butler (1723) to Sir Peter

Strawson (1962) to the contrary) with the natural resentment that unprovoked aggression characteristically elicits.

- The practice or institution of punishment is not necessary, conceptually or empirically, to human society. It is conceivable even if impracticable that society should not have the practice of punishment, and it is possible—given the pains of punishment—that we might even rationally decide to do without it. Not surprisingly, some radical social thinkers from time to time (and even today) have advocated its abolition (Skinner 1948, Bedau 1991, A. Davis, 2003).
- Punishment under law, and especially in a liberal constitutional democracy, incurs considerable costs for persons involved in carrying it out, whatever the benefits may be. Some rationale must be provided by any society that deliberately chooses to continue to incur these costs. The matter is aggravated to the extent that society prefers to incur these costs rather than those of alternative social interventions with personal liberty that might result in preventing crime in the first place and healing the wounds of its victims (Currie 1985).

By way of expansion on some of the considerations alluded to above, we must not forget or obscure the importance of the fact that punishment by its very nature involves some persons (those who carry out punitive acts) having dominant coercive power over others (those being punished). To seek to be punished because one likes it, is pathological, a perversion of the normal response, which is to shun or endure one's punishment as one might other pains, burdens, deprivations, and discomforts. (Only among the Raskolnikovs of the world is one's deserved punishment welcomed as a penance.) To try to punish another without first establishing control over the would-be punishee is doomed to failure. But the power to punish—as distinct from merely inflicting harm on others—cannot be adventitious; it

must be authoritative and institutionalized under the prevailing political regime.

Finally, because the infliction of punishment is normally intended to cause, and usually does cause, some form of deprivation for the person being punished, the infliction of punishment provides unparalleled opportunity for abuse of power. To distinguish such abuses both from the legitimate deprivations that are essential to punishment and from the excesses of punitive sentences that embody cruel and inhumane punishments, one must rely on the way the former are connected to (and the latter disconnected from) whatever constitutes the sentence as such and whatever justifies it (Bedau 1972). This is especially true where punishment through the legal system is concerned, since the punishments at the system's disposal—as well as the abuses—are typically so severe.

The general form of any possible justification of punishment involves several steps. They start with realizing that punishing people is not intelligibly done entirely or solely for its own sake, as are, say, playing cards or music, writing poetry or philosophy, or other acts of intrinsic worth to their participants. Nietzsche and Foucault are among those who would dispute this claim, and they may have history on their side. They think that human nature is such that we do get intrinsic even if disguised satisfactions out of inflicting authorized harm on others, as punishment necessarily does. Others will regard this satisfaction, such as it is, as a perversity of human nature, and will say that we retain the practice of punishment because it enables us to achieve certain goals or results.

Although punishment can be defined without reference to any purposes, it cannot be justified without such reference. Accordingly, to justify punishment we must specify, first, what our goals are in establishing (or perpetuating) the practice itself. Second, we must show that when we punish we actually achieve these goals. Third, we must show that we cannot achieve these goals unless we punish (and punish in certain ways and not in others) and that we cannot achieve them with comparable or superior efficiency and fairness by nonpunitive interventions. Fourth, we must show that striving to achieve these goals by way of the imposition of deprivations is itself justified. Justification is thus closed over these four steps; roughly, to justify a practice of punishment—if not everywhere then at least in a liberal constitutional democracy—it is necessary and sufficient to carry out these four tasks.

Unsurprisingly, no matter what actual society we find ourselves in, we can contest each of these four steps, especially the last. Just as there is no theoretical limit to the demands that can be made in the name of any or all of these tasks, there is also no bedrock on which to stand as one undertakes either a critique of existing systems of punishment or the design of an ideal system. As a result, the foundations of punishment imitate the topology of a Moebius strip—if any path is pursued far enough, it will return to itself and one loses one's grip on what is inside and what outside the justification. Metaphor apart, the inescapable forensic quality of justification defeats all forms of what might be called linear—whether top-down or bottom-up—foundationalism.

3. Consequentialist or Deontological Justification

For several decades philosophers have (over-) simplified the picture of possible forms of normative justification in ethics, policy formation, and law into two alternatives: consequentialist and deontological. They have also undertaken to apply this distinction to the justification of punishment. By a purely consequentialist theory, we mean a theory that imposes no constraints on what counts as the fourth step in justification (see above). The pure consequentialist views punishment as justified to the extent that its practice achieves (or is reasonably believed to achieve) whatever end-state the theorist specifies (such as the public interest, the general welfare,

the common good). Most philosophers would reject this view in favor of introducing various constraints, whether or not they can in turn be justified by their consequences. Thus, a most important part of the theory of punishment is the careful articulation of the norms that provide these constraints on the practice and their rationale.

As for individual acts of punishment—typically, the sentence a court metes out to a convicted offender and the infliction of that sentence on the offender—their justification falls within the justification of the practice itself. In any case they could not be reasonably be justified purely on consequentialist grounds (as an act-utilitarian might wish to do). Sentencers lack sufficient information about all the actual or probable effects of inflicting one rather than another punishment on a given offender at a given time. They lack as well the opportunity and time to secure such information and to use it to inform their sentences. As a result sentencers must content themselves with a largely procedural justification of most of the punishments they impose. Insofar as the system of punishment on which they rely is essentially just, none of the sentencing acts that the institution warrants are unjust (they may, of course, be unwise).

The best justification of punishment is also not purely retributivist. The retributive justification of punishment is founded on two a priori norms (the guilty deserve to be punished, and no moral consideration relevant to punishment outweighs the offender's criminal desert) and an epistemological claim (we know with reasonable certainty what punishment the guilty deserve) (Primoratz 1989, M. Moore 1987). It is arguable, however, whether the guilty always do deserve to be punished; it is also arguable whether, even when they do they ought always to get what they deserve; and it is further arguable whether when they ought to be punished as they deserve, the punisher always knows what it is they deserve (except in the purely procedural sense alluded to above; see also below) (Bedau 1978). We cannot meet these challenges to the deontological retributivist by insisting that punishment is nothing more than a necessary conceptual consequence of living under the rule of law (Fingarette 1978).

Even apart from the problems above, retributivists have yet to construct a nonarbitrary way of deciding what sentence the guilty offender deserves as punishment. Retributivists, ancient and modern, have always been lured by one or another form of *lex talionis* (Davis 1992), despite objections dating from post-biblical times to the present (Walker 1991). Nor does it suffice to abandon like-for-like retaliation in punishment in favor of restating the basic retributive principle in nontalionic form: Severity in punishment must be proportional to the gravity of the offense. Few will argue against this principle, but it still leaves us with a spectrum of alternatives among which to choose, marked at one end by a positivistic legalism (offenders deserve whatever the penal code provides as their punishment) and at the other end by an inchoate moralism (offenders deserve whatever accords with their moral culpability and the harm they have caused).

All retributive attempts to specify the penalty schedule linking crimes to their punishments fail because the proportionality principle underdetermines the schedule. There is no nonarbitrary way to locate either the end points of maximum and minimum severity defining the penalty schedule or the intervals between adjacent punishments (Pincoffs 1977). Without more information it is impossible to calculate which crimes deserve which punishments; an infinite number of different penalty schedules are equally consistent with the retributivist's proportionality principle. And retribution cannot supply the further information needed. As a result, every penalty schedule purporting to incorporate retributive principles exclusively fails to the extent that any given punishment cannot be justified by those principles alone.

But the basic insights of retributivism cannot be merely brushed aside. There is a role for desert in a liberal theory of punishment, but its scope needs careful restriction. The retributivist relies on the assumption that the criminal laws whose violation makes one *eligible* for punishment protect genuine individual rights. Were this not so, the retributivist could not claim that justice requires punishment for the violation of the law. Nor could the retributivist claim that the resentment or indignation directed toward offenders is fitting, rather than merely ill-disguised anger. Retributivism, whether in law or morals, without an appeal, tacit or express, to the justice of punishment is inconceivable—or inconceivably distinct from mere retaliation or revenge (Nozick 1981, Henberg 1990).

Once this is acknowledged there emerges an unmistakable forward-looking, nonretributive point to introducing liability to punishment for law violation, publication of this liability so that it works as a threat, and expectation of increased compliance with the law because of dislike of the perceived punitive threat by most people and their unwillingness to risk incurring what is threatened for noncompliance. Risk of punishment provides an incentive for any normal person to comply with just laws protecting individual rights. No purely backward-looking conception of the practice of punishment, focused exclusively on the desert of the offender, can accommodate provisions for this incentive.

On the view sketched so far, a system of punishment under law is fundamentally a technique of social control (Gibbs 1975), and its employment is justified to the extent that it actually protects such social justice as society through its laws has achieved. This purpose is external, not internal, to the practice of punishment. To accept this conception of punishment is to concede the central claim of the consequentialist, not that of the retributivist. The institution of punishment so conceived is thus not justified on purely deontological or on purely consequential grounds, because punishment manifests some features of each line of consideration, even though the principles justifying it are nonretributive. Nevertheless, punishment retains some retributive elements, conceptually and normatively. Any given act of punishment may look starkly retributive to the one who undergoes it—the sentence imposed is a deprivation inflicted on someone found guilty, and not on anyone else, and it is imposed solely because of that finding.

Against this background we can now consider a step-by-step argument for a liberal justification of punishment. The general idea has been presented in various forms and fragments over the past half century by many writers. [2]

4. Liberal Justification

We can begin with an empirical generalization of unimpeachable reliability: Some kinds of intentional human conduct are harmful to others, and it is inappropriate to expect (teach, require) people who have been victimized by such harm either to forgive those who harmed them or to suffer the harm in silence. (Private retaliation must also be pre-empted by general confidence that offenders will be arrested, tried, convicted, and sentenced by the authorities.) In a just society, undeserved victimization is understood to violate individual rights and is therefore prohibited by law and is punishable. Thus the color and texture of any possible justification for punishment will depend upon more general political and moral theory, consistent with the responsibilities for legal protection afforded by a just society. Justification for punishment under law thus emerges as a contingent matter, inescapably dependent on other and deeper normative considerations that only a theory of social justice can provide.[3]

To repeat, in a society that takes justice seriously, such intentionally harmful conduct will be prohibited by law and, and if and when it occurs, condemned under the law. To do otherwise would be to fail to protect and

vindicate the rights of individuals that the criminal law is principally designed to protect. The central instrument of such condemnation is the penal sanction attached to the law that defines certain harmful acts as crimes.

In a just society that is also a rational society, unlawful harmful conduct is preferably prevented before the fact rather than punished after the fact. From society's point of view, compliance under threat is much to be preferred to noncompliance followed by arrest, trial, conviction, sentence, and punishment. (There are exceptions, of course; justified civil disobedience is one of them.) But compliance is not so valuable that it is worth trying to increase it at any price, especially at the price of irreparable invasions of personal liberty. Thus, a person's willing compliance with the law as a consequence of having internalized the norms of a just society is preferable to one's unwilling compliance or intentional noncompliance. But if willing compliance is not forthcoming, then society must settle for second-best—unwilling compliance—since it is preferable to noncompliance. Prohibition by law plays an essential role in securing grudging compliance, and the principal vehicle for such prohibition is the punitive sanction attached to violation of the criminal law. No doubt, non-deterrent effects of the sanction system, such as the expressive affirmation of shared values, are more important for general compliance than are the deterrent effects. Still, once such sanctions are in place, they create public *liability* to authorized punishment.

Even in a just society, not every person will comply with the law, and not everyone who does comply will do so out of respect for the rights of others, that is, out of recognition of others as persons with rights deserving mutual respect. Here we encounter in another form the fundamental rights-protecting principle on which the system of punishment is built: It is better to increase law compliance by liability to sanctions of those who would otherwise violate the law than it is to permit them to act on their perverse autonomy without any socially imposed cost to themselves, since that would require us to tolerate the victimization of the innocent. Such toleration would be at odds with the moral urgency of protecting rights. For this reason, rational self-interested persons acting behind a veil of ignorance would choose to impose on themselves and on others a liability to criminal sanctions for certain law violations.

If the punitive sanction is to function effectively as a preventive of noncompliance, then it must be perceived not only as a legitimate threat but also as a credible threat. Its legitimacy is established by its protection of individual rights, its authorization by constitutional procedures, and its administration through due process and equal protection of the law. Its credibility is established by its being generally perceived to be both reasonably severe (hence unpleasant) and effectively enforced (hence arrest and its consequences is likely for anyone who does not comply).

There are, however, constraints in the use of penal threats and coercion even to preserve a just social system. Four are particularly important for a liberal theory of punishment.

1. Punishments must not be so severe as to be inhumane or (in the familiar language of the Bill of Rights) "cruel and unusual."
2. Punishments may not be imposed in ways that violate the rights of accused and convicted offenders ("due process of law" and "equal protection of the laws").
3. Punitive severity must accord with the relative severity of the crime: The graver the crime, the more severe the deserved punishment. The severity of the crime is a function of the relative importance of the reasons we have to dissuade people from committing it, reasons that will make reference to harms done to victims, to social relationships, and to the security of our rights.
4. Punitive severity is also subject to the principle of minimalism (less is better), that is, given any two punishments not ruled out by any of the prior principles and roughly equal in retributive and preventive effects for a given offense and class of offenders, the less severe punishment is to be preferred to the more severe.

Conviction of an accused offender under laws that satisfies the foregoing criteria establishes an individual's *eligibility* for punishment. His *liability* to punishment is determined by his own acts and omissions in regard to those laws. All and only punishments that are the product of a system of law consistent with the foregoing constraints may be said to be deserved by the offender. Deserved punishment, insofar as it exists at all, thus emerges as a result of "pure procedural justice" (Rawls 1971). That is, we have only the vaguest idea of the just or deserved punishment for a given offender guilty of a given crime apart from the sentencing schedule provided by the laws of a just society (and thus laws that conform to the constraints above). The punishment deserved is the punishment authorized under a fair penalty schedule; no other conception of deserved punishment can be defended; the perennial lure of an illusory independent criterion for desert, founded ultimately on intuition, as well as of utilitarian calculations, must be resisted. Given this account of desert, anyone both liable and eligible for punishment deserves to be punished, and ceteris paribus ought to be punished.[4]

The argument for imposing deserved punishments so defined on guilty offenders is thus in part an argument from consistency. It is inconsistent to specify liability and eligibility conditions for punishment and then not apply the sanction so authorized when the facts in a given case show that it is warranted. It is unfair to the law-abiding for law-breakers to incur no socially approved cost for their misconduct; it is unfair because it would create a class of harmful free riders in the society. The socially approved costs of crime imposed on offenders consist mainly in the deprivations authorized by the punitive sanction. Fairness to the law-abiding also suggests that society ought to expend a reasonable fraction of its resources in combating crime and preventing victimization.

The creation of a punitive sanction in the name of fairness and under the circumstances specified above is justified. So is the infliction of such a sanction in the name of compliance with the law. Therefore, the practice of punishment, including creating liability to punishment, using sanctions as a threat and an incentive for compliance and actually inflicting the punishment where eligibility conditions are met, is justified.

5. Conclusion

The foregoing argument incorporates deontological and consequentialist considerations. It is better than a pure retributivism because it shows why a system of punishment is needed and how that system is to be nested into the larger political and moral concerns of a just society. It allots a clear and defensible function to punishment (social defense) without yielding to atavistic demands for retaliation or to illusory deontological demands for pure retributive justice, and without pretending that the punishments it metes out are "deserved" in any fundamental sense. The argument acknowledges the sovereign choices of the individual without invoking any awkward and paradoxical "right to be punished" (Morris 1968). It is better than a pure consequentialism, because it constrains punitive interventions with individual liberty to the bare minimum consistent with achieving the purpose of punishment and it is consistent with the rights of offenders. Through the punishment system, all are given fair warning that they put their own rights at risk if they intentionally engage in certain kinds of harmful conduct (H.L.A. Hart 1959). Furthermore, punishment coincides with an ordered hierarchy of moral norms. It has the right "expressive function" (Feinberg 1965)

The system of punishment that emerges under this theory is liberal and non-paternalistic, respects the nominal autonomy of all persons equally, and acknowledges the contingency of its justification as applied in any given case.

It is also true that the system of punishment that emerges under this argument leaves punishment in any actual individual case something of a ritual—in some cases an empty ritual, and in any case a highly formalized act whose exact expressive function and incapacitative effects are uncertain. Acts of punitive deprivation must be imposed on each convicted offender without the comfort of believing, much less knowing, that the purposes for which the system of punishment was designed and maintained will really be advanced by inflicting a given punishment. Too much punishment vs. too little punishment plagues every actual sentencing decision. Some have been led by this fact to view punishment with considerable distrust, because we cannot count on it having any beneficial effect on the punished (Duff 1986)—or on the rest of society. Others are less troubled by this because they focus on how the expressive function of punishment under law serves society by making punishment of whatever degree a "symbol of infamy," whatever its other effects may be (Feinberg 1965). Nevertheless, the stigma of punishment can go too far, in effect rendering sentences indeterminate.

Notice, finally, that the entire argument for the justification of punishment unfolds in the belief that alternative, non-punitive methods of social control have been examined and rejected (or severely limited in scope) on the ground that they will not suffice—or will not work as well as punitive methods in securing compliance with just laws.

Many details remain to be specified before we have a comprehensive liberal theory of punishment in hand. Philosophy can, of course, help supply certain desiderata of the theory, such as specification of the quality and quantity of deprivations (the modes of punishment) appropriate to include in the penalty schedule; construction of the schedule coordinate with the class of crimes; identification of subordinate norms to supplement those already mentioned, which serve as constraints on the schedule and the imposition of sanctions on any given offender; and specification of the norms that make it appropriate to reduce or even waive punishment in favor of some nonpunitive alternative response in a given case (K. Moore 1989). But philosophy alone cannot provide the necessary details; philosophical argument by itself would underdetermine a penal code and has no means to administer one. Yet the heart of a liberal theory of punishment in practice lies in its code of sanctions and their fair administration. Further development of this theory, and its full policy implications, must take place in another forum.

Bibliography

Ackerman, Bruce A., 1980, *Social Justice in the Liberal State*, New Haven: Yale University Press

Alexander, Michelle, 2010, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*, New York: The New Press.

Allen, Francis A., 1981, *The Decline of the Rehabilitative Ideal*, New Haven: Yale University Press.

Barry, Brian, 1989, *Theories of Justice*, Berkeley: University of California Press.

Beccaria, Cesare, 1764, *On Crimes and Punishments*, tr. David Young, Indianapolis: Hackett, 1986.

Bedau, H. A., 2001, "Feinberg's Liberal Theory of Punishment," *Buffalo Criminal Law Review*, 5: 103–44.

–––, 1991, "Punitive Violence and Its Alternatives," in James B. Brady and Newton Garver (eds.), *Justice, Law, and Violence*, Philadelphia: Temple University Press, pp. 193–209.

–––, 1983, "Bentham's Utilitarian Critique of the Death Penalty," *Journal of Criminal Law and Criminology*, 74: 1033–65

–––, 1978, "Retribution and the Theory of Punishment," *Journal of Philosophy*, 75: 601–20.

–––, 1972, "Penal Theory and Prison Reality Today," *Juris Doctor*, 2: 40–43.

–––, Bentham, Jeremy, 1789, *The Principles of Morals and Legislation*, Laurence J. LaFleur (Intro.), New York: Hafner Publishing, 1948.

Butler, Joseph, 1723, "Sermon Upon Resentment", in Butler, *Works* (Volume 2), Oxford: Oxford University Press, 1850, pp. 87–98.

Card, Claudia, 1973, "Retributive Penal Liability," *American Philosophical Quarterly Monographs*, 7: 17–35;

Currie, Elliot, 1985, *Confronting Crime: An American Challenge*, New York: Pantheon.

Davis, Angela Y., 2003, *Are Prisons Obsolete?*, New York: Seven Stories Press.

Davis, Michael, 1992, *To Make the Punishment Fit the Crime: Essays in the Theory of Criminal Justice*, Boulder, CO: Westview, pp. 42–68.

Duff, R. A., 1986, *Trials and Punishments*, Cambridge University Press.

–––, 2007, *Answering for Crime: Responsibility and Liability in the Criminal Law*, Oxford: Hart Publishing

Dworkin, Ronald, 1986, *Law's Empire*, Cambridge, MA: Harvard University Press

Farrell, Daniel M., 1985, "The Justification of General Deterrence," *The Philosophical Review*, 94: 367–394;

Feinberg, Joel, 1965, "The Expressive Function of Punishment," *The Monist*, 49: 397–423;

Fingarette, Herbert, 1978, "Punishment and Suffering," *Proceedings of the American Philosophical Association*, 50: 499–525.

Foucault, Michel, 1977, *Discipline and Punish: The Birth of the Prison*, New York: Pantheon.

Garland, David, 1990, *Punishment and Modern Society*, Chicago: University of Chicago Press.

–––, 2001, *The Culture of Control: Crime and Social Order in Contemporary Society*, Chicago: University of Chicago Press.

Gibbs, Jack P., 1975, *Crime, Punishment, and Deterrence*, New York: Elsevier.

Goldman, Alan, 1982, "Toward a New Theory of Punishment," *Law and Philosophy*, 1: 57–76;

Gross, Hyman, 1979, *A Theory of Criminal Justice*, New York: Oxford University Press.

Hampton, Jean, 1984, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs*, 13: 208–38;

Harding, Christopher, and Richard W. Ireland, 1989, *Punishment: Rhetoric, Rule, and Practice*, London: Routledge.

Hart, Jr., Henry M., 1958, "The Aims of the Criminal Law," *Law and Contemporary Problems*, 23: 401–41.

Hart, Herbert L. A., 1968, *Punishment and Responsibility: Essays in the Philosophy of Law*, Oxford: Oxford University Press

Henberg, Marvin, 1990, *Retribution: Evil for Evil in Ethics, Law, and Literature*, Philadelphia: Temple University Press.

Hoekema, David, 1986, *Rights and Wrongs: Coercion, Punishment, and the State*, Selinsgrove, PA: Susquehanna University Press.

Honderich, Ted, 1976, *Punishment: The Supposed Justifications*, rev. ed., Harmondsworth: Penguin.

Husak, Douglas, 2008, *Overcriminalization: The Limits of the Criminal Law*, Oxford: Oxford University Press.

Kelly, Erin I., 2009, "Criminal Justice without Retribution," *Journal of Philosophy*, 106: 440–462.

Lacey, Nicola, 1988, *State Punishment: Political Principles and Community Values*, London: Routledge.

Maestro, Marcello, 1973, *Cesare Beccaria and the Origin of Penal Reform*, Philadelphia: Temple University Press.

Martinson, Robert, 1974, "What Works?—Questions and Answers About Prison Reform," *The Public Interest*, 10: 22–54.

Moore, Kathleen Dean, 1989, *Pardons: Justice, Mercy, and the Public Interest*, New York: Oxford University Press

Moore, Michael S., 1987, "The Moral Worth of Retribution," in Ferdinand Schoeman (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, Cambridge: Cambridge University Press.

Morris, Herbert, 1968, "Persons and Punishment," *The Monist*, 52: 475–501.

Murphy Jeffrie G., 1973, "Marxism and Retribution," *Philosophy and Public Affairs*, 2: 217–43.

–––, 2003, *Getting Even: Forgiveness and Its Limits*, Cambridge: Cambridge University Press.

Nietzsche, Friedrich, 1887, *On the Genealogy of Morals*, tr. Walter Kaufmann, New York: Vintage, 1969.

Nozick, Robert, 1981, *Philosophical Explanations*, Cambridge, MA: Harvard University Press, pp. 366–74.

Pincoffs, Edmund, 1977, "Are Questions of Desert Decidable?" in J. B. Cederblom and William Blizek (eds.), *Justice and Punishment*, Cambridge, MA: Ballinger, pp. 75–88.

Primoratz, Igor, 1989, *Justifying Legal Punishment*, Atlantic Highlands, NJ: Humanities Press.

Quinn, Warren, 1985, "The Right to Threaten and the Right to Punish," *Philosophy and Public Affairs*, 14: 327–373.

Rawls, John, 1955, "Two Concepts of Rules," *Philosophical Review*, 64: 3–32.

–––, 1971, *A Theory of Justice*, Cambridge, MA: Harvard University Press, 1971

Reiman, Jeffrey, 1990, *Justice and Modern Moral Philosophy*, New

Haven: Yale University Press.

Richards, David A. J., 1977, *The Moral Criticism of Law*, Encino, CA: Dickenson.

Scanlon, T. M., 1998, *What We Owe To Each Other*, Cambridge, MA: Harvard University Press.

–—, 1999, "Punishment and the Rule of Law"; reprinted in Scanlon, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press.

Semple, Janet, 1993, *Bentham's Prison. A Study of the Panopticon Penitentiary*, Oxford: Clarendon Press.

Shelby, Tommie, 2007, "Justice, Deviance, and the Dark Ghetto," *Philosophy and Public Affairs*, 35: 126–160.

Singer, Richard G., 1979, *Just Deserts: Sentencing Based on Equality and Desert*, Cambridge, MA: Ballinger.

Skinner, B. F., 1948, *Walden Two*, New York: Macmillan.

Strawson, Peter F., 1962, "Freedom and Resentment"; reprinted in Strawson, *Freedom and Resentment and Other Essays*, London: Methuen, 1974, pp. 1–25.

Stuntz, William J., 2011, *The Collapse of American Criminal Justice*, Cambridge, MA: Harvard University Press.

Tadros, Victor, 2011, *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press.

Ten, C. L., 1987, *Crime, Guilt, and Punishment*, Oxford: Clarendon Press.

Twentieth Century Fund, 1976, *Fair and Certain Punishment*, New York: McGraw Hill

von Hirsch, Andrew, 1976, *Doing Justice: The Choice of Punishments*, New York: Hill & Wang

–—, 1985, *Past or Future Crimes: Deservedness and Dangerousness in the Sentencing of Criminals*, New Brunswick, NJ: Rutgers University Press.

–—, 1993, *Censure and Sanctions*, Oxford: Oxford University Press.

Walker, Nigel, 1991, *Why Punish?*, Oxford: Oxford University Press.

Wasserstrom, Richard, 1980, "Punishment," in Wasserstrom (ed.), *Philosophy and Social Justice*, Notre Dame: University of Notre Dame Press, pp. 112–51.

White, Mark D. (ed.), 2011, *Retributivism: Essays on Theory and Policy*, New York: Oxford University Press.

Wilson, James Q., 1975, *Thinking About Crime*, New York: Basic Books, 1975.

Zimring, Frank E., 1977, "Making the Punishment Fit the Crime: A Consumer's Guide to Sentencing Reform," *Occasional Papers*, No. 12, Chicago: University of Chicago Law School.

Academic Tools

- How to cite this entry.
- Preview the PDF version of this entry at the Friends of the SEP Society.
- Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
- PP Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

- The Innocence Project
- The Sentencing Project
- The Forgiveness Project
- Bureau of Justice Statistics
- Amnesty International, Death Penalty Facts

Related Entries

autonomy: personal | Bentham, Jeremy | consequentialism | desert | ethics: deontological | Foucault, Michel | Hegel, Georg Wilhelm Friedrich | Kant, Immanuel | Nietzsche, Friedrich | pain | punishment, legal | Rawls, John | rule of law and procedural fairness

Acknowledgments

The editors would like to thank Bliss Carnochan for spotting a mistake in an earlier version of this entry; the statistics concerning the number of prisoners in federal and state prisons were in error and they have now been fixed.

Notes to Punishment

1. Twentieth Century Fund, *Fair and Certain Punishment*, New York: McGraw-Hill, 1976; von Hirsch 1976; Singer 1979.

2. See, e.g., H.M. Hart, Jr. 1958, H.L.A. Hart 1959, Honderich 1976, Feinberg 1965, Wasserstrom 1980, Gross 1979, Richards 1977, Hoekema 1986, Lacey 1988, Ten 1987, and Walker 1991.

3. See, e.g., Rawls 1971, Scanlon 1998, Dworkin 1986, Ackerman 1980, Reiman 1990, and Barry 1989.

4. There seems to be a sharp contrast between the role that forgiveness of offenders may properly play in our personal lives, and the narrow or even nonexistent role it ought to play in a punishment system under law. Deserved mercy, apart from being an oxymoron, may have a role to play in the latter, but it is not the same thing as deserving to be forgiven or deserving to be excused.

Copyright © 2017 by the authors
Hugo Adam Bedau and Erin Kelly