

19 Model Penal Code (1962), § 3.07 (2) (b) (iv).

20 *Ibid.*, 3.07 (2) (b) (iii). See the comments to the Code concerning this limitation. "Some members of the Advisory Committee opposed the statement of this limitation of the privilege, relying on Section 3.09 (3) to sustain a prosecution based on recklessness or negligence towards innocent persons for whatever crime such recklessness or negligence may involve. The Council was disposed, however, to retain the limitation as a means of emphasizing and articulating the priority that law enforcement personnel ought to accord to safeguarding innocent persons against injury from deadly force directed against persons fleeing from arrest. To leave the matter to an assessment of recklessness or negligence, with the opportunity that this entails to urge that greater risk to innocent persons is justifiable, was deemed to leave the rules that ought to govern law enforcement in too vague a state." Model Penal Code, *Tentative Draft No. 8* (1958), p. 58.

21 There may be considerable difficulty in obtaining the information necessary to resolve these issues if firearm discharge and wounding reports are not maintained. The establishment of state-wide data banks on firearm discharges and on wounding incidents would, of course, remove much of the present difficulty in obtaining information. See the

recommendations in *Police Handguns and Deadly Force*, pp. 22-24. For a summary of recent articles analyzing available data on deadly force, see Cynthia Sulton and Phillip Cooper, "Summary of Research on the Police Use of Deadly Force," in National Institute of Law Enforcement and Criminal Justice, *A Community Concern: Police Use of Deadly Force* (Washington, D.C.: United States Government Printing Office, 1979), pp. 69-94.

22 Unless, of course, the decision is preempted by the courts. Only two appellate cases concerning the justifiability of the hollow-point have been found. See *Lambert v. State of Indiana*, 249 N.E. 2d 502 (Ind. 1969). In *Lambert*, only a procedural question was before the court. By way of dicta the court said, "We note, without approval, that the police officer who wounded appellant was using ammunition of unusually high velocity and with a hollow-nose (hollow-point) bullet designed to inflict the maximum amount of injury. . . . Under all circumstances here existent we can only come to the conclusion the officer was not justified in the excessive use of force displayed in this matter, nor can the use of this particular type of ammunition for general police work be justified." 249 N.E. 2d 502, 508 (1969). (The person wounded was a fleeing, nonresisting suspect.)

Understanding Retribution

ROGER WERTHEIMER

For two centuries now, a great debate, principally focused on criminal punishment, has been carried on within the chambers of government, the community of theorists, and the consciences of ordinary people. As with any large subject, the issues have been many and diverse, and the positions taken have been equally var-

ious, but at the center of concern and pervading throughout there has been some sense that the propriety of retribution is problematic, especially when exacted by a state. And as with any anxiety over our fundamental beliefs, precisely what the problem is has hardly been clear.

Distribution

Understandably, people with a market mentality prefer that punishment be of some profit. So, we weigh

Roger Wertheimer, author of The Significance of Sense, is Associate Professor of Philosophy at the University of Houston.

the costs and benefits, the risks and uncertainties; yet the closer we come to operational detail, the more our calculations of cause and consequence become confounded or controversial. But our difficulties with the data and projections are familiar problems in social science and social policymaking. With punishment we seem more muddled about the conception of the activ-

ity and the relevance of its causal relations to its evaluation and justification.

Among theorists, this frequently finds expression in misgivings about the possibility that a legal system engineered to generate the largest social good might permit or prescribe (overtly or covertly) the punishment of innocent or incompetent persons, or punishments of a severity disproportionate to the heinousness of the crime, or some other acts whose injustice we immediately sense without consulting a prodigious cost-benefit analysis. Principles of retribution are here relegated to the status of a species of distributional principles that may challenge aggregative ambitions. As such, among the theorist's uncertainties is whether the bare possibility or a rare realization of a mismatch between optimal utility and our common sense of justice renders the system unconscionable, or whether the acceptability of the system depends on the seriousness of the shortfall.

Now, all these problems about the distribution of retribution draw us away from the distinctive character of our preoccupation with punishment. The dis-

tributional principles and problems are much the same for any system of penalties or other disincentives, punitive and nonpunitive alike; and insofar as they differ, they are much the same for our assessments of debts of gratitude and other deserved benefits. Of course, reasonable doubts may be raised about any element or aspect of such matters, and at any level of generality and specificity—doubts about proportionality, responsibility, requisite certainty, and so on. Necessarily, if punishment and principles of retribution are systematically interrelated with, analogous to, implying or implied by, our other conduct and convictions, then many of these reasonable doubts can be directed at particular punishments and principles or punishment in general. The converse is equally true.

However, we seem to sense something distinctively dubious about punishment or retribution. Our anxiety is not free-floating, but our conception of its proper object and cause is inchoate. It's even unclear whether the doubt is directed at punishment or retribution, due to confusion about the nature of each and their interrelation.

Response Norms

To locate the peculiarly problematic character of punishment and retribution, we need to understand some essential characteristics of norms in general. A norm is a violable thing that, if justified, justifies a response to violations of it and to conformity as well. Norms can be cast in the form of conditionals: If p then q . For example: If A and B are persons, then neither intentionally kills the other except in circumstances C .¹ The conditional is (or is regarded as) stating a norm, rather than a factual claim, when a *negative response* to the situation, p & not- q , is (or is regarded as) justified. The norm does not itself specify a response, so a justified negative response, r , presupposes some *negative response norm* of the form: If p & not- q , then r . The response, r , is negative in the sense of providing those governed by the first-order norm some reason to conform to it. This abstract analysis applies to any norm at all.

Consider a sample of negative responses, noting the categorical differences between them: Voiding a contract, voiding a law, nullifying a touchdown, awarding

a free throw, denying the title *grammatical* to a word sequence, criticizing (qua judgmental act), voicing criticism, complaining, filing a complaint, suing, pursuing, apprehending, self-condemnation, feelings of guilt, remorse, self-punishment, confession, contrition, repenting, apologizing, explaining, atoning, making amends, compensating, rectifying, begging forgiveness, forgiving, excusing, condemning, feeling indignation, resenting, contemning, deriding, reprimanding, castigating, chartising, repudiating, denouncing, discrediting, defending against an unlawful attack, denying title to ill-gotten gains, denying, disavowing, disassociating. Many of these acts and attitudes are *defined* as responses justified by negative response norms. None of these (aside from self-punishment) is, in and by itself, a punishment.

Response norms may serve a causal role by motivating conformity to a norm and thus may be assessed by risk and cost-benefit analyses of their efficacy. However, conformity may be motivated by factors other than fear of the negative responses, and it generally is

in any well-functioning normative system. Moreover, while conformity to the norms may be a desideratum if those norms are justified, it is neither essential nor enough for the norm to be operative or valid. Conformity alone can evidence no more than a regularity which might be caused without the operation of a

The need for response norms is internal to the nature of norms, however inefficacious the responses may be.

norm. The essential constituent of the operation of a norm lies in our second- and higher-order attitudes and activities, our responses to violations and conformity to the norm, and our attitudes and activities regarding those responses and so on. The need for response norms is internal to the nature of norms, however inefficacious the responses may be. Even with conventional norms, such as statutes, which exist in virtue of our patterns of action, the first-order norms may continue to exist and be valid, despite being commonly violated and unresponded to, as long as the response norms are regarded as valid by the agents of the normative system.

When evaluating alternative response norms for a given norm, considerations of causal efficacy may seem paramount, but that is only because the intelligible options are severely restricted by a matrix of tacit conceptual constraints that are taken for granted. Consider, for example, constitutive norms—norms that define requirements for being a such-and-such (a con-

tract, touchdown, sentence). A violation of such a norm must justify nullification of the violation (or its product); to refuse to nullify (void, refuse normative recognition to) a violation of a constitutive norm is only to operate with an altered norm. What effects nullification are primarily normative, not causal, relations.

The specific normative character of any norm is defined by and related to the character of its response norms. For a norm to define a right, coercion to protect the right must be justified (Kant), and for a norm to define a duty, exacting performance must be justified (Mill).

So too, kinds of normative systems are distinguished principally by differences in their response norms. The same conditional proposition (for example, the one about killing) could, in theory, be a norm in essentially different kinds of normative systems—a moral code, legal code, game, or even a language. But, for example, linguistic rules (as distinct from social rules governing speech) cannot assign penalties since linguistic rules do not distribute benefits and burdens. Games can have penalties, but no punishments. Legal systems can punish, and some theorists claim that they must have the requisite response norms (though whether the necessity is conceptual or causal is controversial). When we ask which, if any, punitive response norms they should have, we look to our morality, primarily its principles of justice. When we ask that question about our moral code itself, we must look in the very same place. A morality is an inherently self-referential system of norms and their response norms. We cannot evaluate our principles except by means of these principles themselves.

Rule Retributivism

A kind of immanent justice is implicit in the logic of norms, for we may say, "You broke the rules. Now you must suffer the consequences. That's the rule." Perhaps such a thing is normally said only by someone assuming that the rules are justified, yet it is not the justice of the rules that is asserted or denied, but only their inexorability. (A functionary might add: "And don't give me a hard time. I'm only doing my job.")

The rules are here regarded as facts, like natural laws. As with any factual claim, interpretative and epistemological questions may be raised about just what the facts are—whether these really are the rules and whether *this* ruling is necessitated by them—but they do not question the metaphysical truths, that a fact is a fact and that a law necessitates. There may be some play in the rules, room for an official to bend the rules

without breaking them. An authority may define spaces in which personality can play, where excuse, mitigation, extenuation, and discretion operate. But once a final ruling has been rendered so that it can be said without qualification, "This is required by the

Justice is a self-referential system of requirements: justice justifies negative responses to injustices.

rules," then a sufficient explanation of the consequences has been given. We appeal to authorities to settle disputes, both practical and intellectual. An authority can say, "Listen, we're not going to discuss this any longer. We will do as I say, and that's final."

This feature of norms might be called *rule retributivism*. There is an element of justice in it and an element of it in justice. The justice in it is an absence of one form of injustice, arbitrariness, infidelity to lawfulness.

Capriciousness gives way to conscientiousness, even-handedness. Similar cases are reliably treated in similar ways. Expectations are legitimated and satisfied. For those who lack any loyalty to the governing rules, a justification of a negative response is not implied by a description of the logic inherent in norms and the content of the normative system. How much of a constraint is imposed on the content of the norms by this internal logic of a normative system is debatable. Even if it is little, the facts here are not devoid of moral meaning. Gratitude, loyalty, or the like may be owed even to rules that are unjust or work injustices. Such obligations may be transmitted by the rules in the form of a commitment to stand behind their consequences, and those commitments might merit respect from everyone, including persons with no respect for or obligation to the rules themselves.

So too, the element of this in justice—that justice requires that justice be done—does not itself tell us much about which rules are just and which are the principles of justice. But it does remind us that justice is a self-referential system of requirements: Justice justifies negative responses to injustices.

Locating Moral Retributivism

The response norms of a morality define a norm as moral (as opposed to nonmoral and not, or not only, as opposed to immoral). Moral response norms are distinctive in being explanatory and justificatory reasons for our emotions and attitudes, not just our actions. Our motives are regulated and often defined, if not generated, by reference to them. What defines a feeling of guilt, remorse, outrage, or indignation is that it is partly explained by the person's having a pair of moral beliefs: A moral norm has been violated, and its moral response norm justifies this feeling. We refer to such feelings to explain and justify our behavior. The behavioral response norms of morality are, in this regard, derivatives of the attitudinal ones.

Now, one thing bedeviling our doubts about punishment is that the debaters describe themselves either as *retributivists* or *consequentialists* and conceive the contrast as comparable to that between, say, libertarians and Marxists. So, whether all my readers are retributivists at heart will not be proven by polling.

The closet retributivists among us, who regard themselves—their better half, their reason—as rid of "real" belief in retributivist principles, would not without reluctance confess publicly to the surges of retributivist sentiment they find welling within themselves when properly provoked. Some are not so much ashamed as amused by their hypocrisy, passing off their penchants as some harmless atavism they (almost) never let get the better of themselves. And beyond all that, retributivism may be closeted even from one's consciousness.

One's retributivist consciousness can be raised by considering a whole matrix of common, confident convictions. For example, generosity deserves gratitude, and, more generally, those who do good deserve to receive good. Wrongdoing *can* justify denying the wrongdoer claims and entitlements—to protection against injury during the wrongdoing, to the profits of wrongdoing, to goods needed to compensate victims of the wrongdoing. Further, we may properly deny

loyalty to the disloyal, deny equal opportunity to the treacherous, and deny to the disaffecting title to a claim on our affections. We may justly deny a wrongdoer any claim to our generosity, benevolence, esteem, trust, and virtually any of our goods. We may desert those who desert us, abandon them to a desert of their own making, a lifeless, desolate wasteland bereft of the benefits of our community and society.

Further, wrongdoing can give us a right, sometimes *every* right, to be angry; and we may be right to be angry at a wrongdoer. Wrongdoing can justify feelings of resentment, indignation, outrage, remorse, guilt, and the like. Consequentialists should consider whether such sentiments are essential signs of an operative justice and whether such sentiments or their propriety can be conceptualized in nonretributivist terms. Sociopaths aside, the consequentialists I know of subject themselves to self-flagellation for their faults (real and imagined) at roughly the rate the rest of us do, and they do not decide when they have suffered enough for some wrongdoing by determining whether it is now sufficiently less likely that they will not sin in the same way again, or anything of that consequentialist kind.

Can we care about justice without believing that a homicide, if unjustifiable, unmitigated and inexcusable, is due cause for justified anger? And do we not endorse the wrongdoer's recognizing what he has done and reforming, not by drugs or surgery, but by passing through the pain of guilt and contrition? Though we may differ about the proper degree and duration of discomfort, don't we think he ought to feel guilty or remorseful, to have a hostile attitude toward

himself that accepts the propriety of his suffering? And surely, to endorse his having that attitude toward himself is to express that attitude toward him.

Forget how and by whom desert is awarded, and focus on whether it *should* be awarded at all. Consider, say, a state lottery, or some other impeccably random procedure. Could anyone with a sense of justice be without any sense that this would be a better, more fittingly ordered world if the winners were the honest, the kind, and the decent rather than the ne'er-dowells? And if wrongdoing can justify a desire that the

Moral response norms are distinctive in being explanatory and justificatory reasons for our emotions and attitudes, not just our actions.

wrongdoer suffer and a belief that he deserves to suffer, why can't the desert justify intentionally providing it?

Again, any response norm, moral or not, presupposes that its first-order norm is justified, and thus cannot justify it. It is the meaning of our moral principles, not their truth, that our principles of retribution explain. Ultimately, it is only through our retributive acts and attitudes that we can give moral significance to human actions or manifest our conception of their significance.

Condemnation

Punishment seems problematic in principle not because of the pain imposed per se but because of condemnatory hostility that propels its imposition. We have no comparable qualms about any close cousins of punishment such as nonpunitive penalties, the suffering incurred through arrest and custody, or the thwarting of an unlawful attack. Any of these and other harms we confidently impose may be as great, or greater, than punitive pains, measured on the scales that measure the *price* of a penalty. Yet, for example,

we require that we have no reasonable doubt—a *moral* certainty—before we punish, but only probable cause or a preponderance of evidence to inflict other harms.

A punishment is an act defined and explained by a sentence of condemnation. For an act to be a punishment, the agent need not inflict an evil upon the condemned and do so intentionally in the execution of a condemnation for disobedience. A punishment fulfills the demand of a condemnation, an act explained by and expressing hostility. Virtually everything con-

ceptually interesting in punishment is in the condemnation.

Condemn has a succession of sense, but throughout it preserves the flavor if not the full force of its cognates *damn* and *damage*. Unlike the broad categories of criticizing and disapproving, condemning is essentially hostile. It expresses anger or hatred, not mere sorrow or fear or disgust—an *aggressive* disposition or wish not just to avoid, prevent, oppose, or reform but

Punishment seems problematic in principle due not to the pain imposed per se but to the condemnatory hostility that propels its imposition.

to damage, deprive, or even destroy the object or to have it suffer that fate. To condemn a building is to declare it dangerous and shut it off from society until it is razed or renovated. When a court condemns a criminal, it expresses authority's ill will and officially authorizes the community to bear that hostility toward the convict. When referring to a mild punishment of a child, *condemn* may seem too forceful; too much of the violence of *damn* comes through. However, the critical distinctions here are matters of kind, not of degree.

The meaning of condemnation (not the word, but rather its referent) is manifested in many ways. Though the pain inflicted in a wrongful punishment may be less than that in a legally proper but mistaken arrest of an innocent man, mistakenly punishing an innocent man has a more terrible character than pre-*condemnatory* mistakes. The arrest is not false; the mistake is excusable. But the punishment of the wrong man wrongs the man. Justice itself has miscarried. However faultless, misdirected hostility is inexcusable. Only forgiveness can restore the possibility of amicability: Trust, not in the competence, but in the basic benevolence of the erring antagonist must be reaffirmed.²

Condemning something *to* something is an act (often only a commanding) intended to cause the condemned to suffer an evil.³ The intention need not be fulfilled. But in punishment the condemning-to must succeed, and, further, the condemning-to must be due to a judgmental condemning. We can judgmentally

condemn without condemning-to or trying to, but we must in some sense desire and approve—regard it as a good—that the condemned suffer.

That desire is characteristic of anger, hatred, motives for hostility. However, the attitude expressed in condemnation is normative and self-referential, not a mere brute feeling. A condemner must regard his hostility toward the condemned as justified by a pair of justified beliefs: The condemned is evil (bad or the like), and that justifies my hostility. The first belief is the content of the condemnatory judgment, what the condemned is condemned *for*; the second is the content or consequence of a negative response norm, and, if the latter directs a condemning-to, it is expressed in a sentence of condemnation. Naturally, when we are angry or hate for personal reasons we are prone to believe (or wish) that our feelings are justified and thus to be self-righteous. But while condemning may be rooted in and presuppose natural dispositions to anger or hatred toward what threatens harm to what we care about, proper condemnation is essentially impersonal. The condemned is thought to be evil, and not merely an evil for the condemner.

Perhaps normally to judge something to be evil is to criticize it, but this is not itself to condemn it. A verdict of guilty is not itself a sentence of condemnation. Condemnation is constrained in its objects and occasions.

Punishment comes only from condemnation for disobedience or norm violation. If the condemnation is explained by justified anger, its proper object must be conceived of as the cause of a (threatened) wrongful harm or failure to benefit something needy and valued. Some victimless acts may be wrong, perhaps shameful, disgusting or otherwise discreditable, but, absent a queer cosmology or a conception of consensual victimization, condemning them seems crazy.

However faultless, misdirected hostility is inexcusable.

Further, a condemnatory judgment is a judgment in the root sense, not a mere cognitive process or product but an *authoritative* opinion, one that requires a position of power, that of sitting or standing in judgment. In our morality, any person may judgmentally condemn; in our legal systems the governed can only make complaints. Condemning requires presuming

oneself empowered to pass judgment and being empowered to do so.

We may, in various ways, lack the power to condemn wrongdoing. Sometimes we cannot sit or stand in judgment because we are bowed before a fearsome oppressor and must *plead* with him not to wrong us or someone else. Our oppressor may later make such a plea if he is hauled before a court, either a plea before conviction or a plea before sentencing. A plea is not an accusation, verdict, or condemnation, though it may have the same propositional content. When we plead we do not state or imply that our audience's present or prospective wrongdoing calls for a hostile response. On the contrary, we do not wish to imply it; we dare not and do not. To declare, even by implication, that a hostile response is called for is to endorse that response and thereby express and make a hostile response. It is tantamount to war, which is not what we do whilst desperately praying for peace.

Sometimes we cannot condemn because we too are not without sin: We have done or would do the same or worse ourselves. But actually this by itself does not disable us, for we can condemn ourselves and treat

others in the same way. Rather, it is that condemning requires some belief that the demand to have done otherwise is reasonable, and one test of reasonableness is whether we would judge the demand to be

A verdict of guilty is not itself a sentence of condemnation.

reasonable if it were made of us in the same circumstances. That, of course, is a difficult test since its outcome depends on one's own motivational system and on which aspects of the wrongdoer's motivational system we regard as parts of the given circumstances. In any case, to condemn one must presume oneself empowered to make a demand, and one cannot make a demand without some power to enforce it and thus some power to comply when the demand applies to oneself.

Authority

In punishing we presume to operate the machinery of authority.⁴ An authority is a normative, self-referential, communicational agency, a power that claims the legitimacy of its power, an author whose word is law. Like an author of a fiction, an authority authors the world of its subjects: He makes things to be so in and by saying they are so, and that is the point of saying it.⁵ Unlike the author of a fiction, an authority is serious, and the world he commands is a normative ordering of the natural world, an authorized version of the real world.

The distinctive expression of authority is: You must (must not, may, will, are to) do this because I tell you that you must (must not, etc.) do this. Only some authorities require that you be motivated to do this by their telling you to do it, but any authority insists that, in virtue of his saying that you must do this, it is true that you must do this. An authority regards his telling you to do this as itself constituting a final reason for you to do it, but he need not require that you do this for that reason.

By contrast, if someone advises or instructs you to do something, the truth of what he says and your reason for acting accordingly are antecedent to his saying what he says. So too, if an unauthorized power—such as a gunman—tells you to do this, its telling you does not give you a new final reason for acting but only informs you that doing this is a means to attaining ends (such as preserving your health) you already have. Unauthorized imperatives are *hypothetical*; those of an authority are *categorical*.

A punishment has the structure of a communicational act with the requisite form of Gricean reflexivity.⁶ It is a penalty described in and explained by a *sentence* of condemnation. The *terms* of a punishment are given in and comprise that sentence. (It is not by a happy chance that a punishment can teach someone a lesson, even though it may fail to do so.) Though the punishee need not know that he has been condemned or that his suffering is due to it, in paradigm cases he does, and in any case his punishment presupposes a communicational relation: The authority

regards the condemned as having been commanded.

By contrast, vengeance does not presuppose a command, condemnation, authority, or communicational relation. Neither avenger nor avenged need presume to command the offender, to be an author of his world, or to bind others by their word. Revenge responds to a personal offense, not to an offense against an authoritative order, but to an insult, an assault, and paradigmatically an assault against honor. Neither avenger nor avenged need presume any superiority over the offender: They are gratified to get even. Vengeance, honor, shame are personal, prepolitical, not prior to self-consciousness and personal relations, but prior to a recognition of being ruled by authority, an independent voice that commands respect, constrains the will, restrains self-esteem.

Penalties per se do not presuppose an authority, a voice behind the rules. Games have penalties but no punishments. Something done in playing a game may be punishable, but not per se for breaking a rule of the game. Penalties may be imposed by a judicial authority within an activity—an umpire or referee—that makes rulings on events governed by the rules, but punishment presupposes a legislative authority that makes the rules. We do not *obey* or *disobey* the rules of games; they are not (qua game) dictated by an authority. We may play games playfully or take them seriously, but they are not in themselves serious activities. There is no authoritative attitude behind the rules dictating how they are to be taken.

Like feeling guilty or indignant, punishing is not a rule-governed activity in the sense of being performable and explicable by the following of rules alone.

Within a system of rules, punishments, like penalties, may be imposed by applying a negative response norm for violations. What makes a penalty punitive is not something in the rules, but rather an attitude behind the rules, which may be said to permeate the rules. Penal rules may be enacted for any of many purposes—to deter, teach, secure revenue, enliven a game, and so on—but, qua penalty, they do not presuppose any particular purpose, attitude, or intent.

What makes a penalty punitive is not something in the rules, but rather an attitude behind the rules, which may be said to permeate the rules.

Punitive rules *may* be enacted for many of the same purposes, but they *must* have an internal aim directed by condemnatory hostility.

Authority is hierarchical; it subordinates those subjected to it, defining a difference of status. A punishment displaces; it reduces one's social-political status. The place of the condemned is a nether world. A penalty only disadvantages, worsens only one's economic position. It is something of value *paid*—by exaction or freely, even with pleasure—for failure to meet a standard. We pay penalties; we do not *pay* punishments. We suffer punishments, are subjected to them, undergo them. We do not *undergo* penalties.

Intention

Like vengeance, punishment is a self-conscious imposition of suffering. Though we may fail in some of our intentions in punishing (such as mistakenly punishing the wrong man or accidentally punishing too severely), we cannot punish by mistake, accident, or inadvertence. We cannot punish without intending to punish.

A punisher must present himself to the punishee (if not also to himself) as an agent of a state or a moral code or God Almighty or some lesser authority. He

cannot be by happenstance conforming to a penal law and carrying it into effect. To punish, he must apply the law and act from the law, at least in the sense of representing himself as intending to be applying the law, thus realizing *its* intent, *its* will. Our penal law expresses such a will. We may call it the will of the people, and the people may say it is the will of God or their own conscience. What that means and whether it be the will of Reason have yet to be determined. In any case, though a sentencing judge need not *personally*

harbor resentment or any hostile feelings, though he may act with cold-blooded conscientiousness or conform to the law out of crass self-interest, nevertheless, insofar as he acts out of respect for the law, he identifies his personal will with the will of the law. Whatever his personal intentions and motives may be, if he is to play the role of punisher and his act is to be a punishment, his official will as punisher is identified with the will of the law. And the will of a punitive law is condemnatory, hostile.

Nonhostile intentions may be operative as well, but like vengeance, and unlike penalties, an act is not a punishment even if suffering is inflicted knowingly and willingly, if it is only as a means or side-effect. In penalizing, though we believe the penalty to be an evil and we intentionally exact the payment of it, we might do so only as a means. In punishing we intentionally exact the *suffering* of a penalty, an evil for the punishee. In vengeance we intentionally exact suffering, but not as a penalty.

For the punisher to inflict an evil *intentionally*, he must intend to inflict an evil, and, in consequence of that intent, inflict an evil. To intend to inflict an evil, he must believe and (in some sense) desire that what he inflicts is an evil for the punishee. Thus, that belief and desire must be essential parts of his reason for performing that action. But evil is not itself a causal property, so a punisher can never be intending *only* to achieve some casual consequence of the suffering be-

cause then it would be inessential that he believe and desire that he inflicts an evil. The suffering must be sought as an end itself, and sought precisely because it is an evil for the punishee. Thus the official will of the

We pay penalties; we do not pay punishments. We suffer punishments, are subjected to them, undergo them. We do not undergo penalties.

punisher must regard it as a good that the offender suffer an evil.

Punishment calls for retribution. We can penalize, but we cannot intend or be motivated to punish without a belief that the due of evildoing is an evil for the evildoer, and we cannot justifiably punish unless that belief is justifiable. While we might motivate or cause someone to punish without ourselves having that belief, he cannot be motivated to punish without his having the belief. He need not recognize that he has this belief; self-reports on such matters are notoriously suspect. And, again, he need not even have the belief if his verbal or nonverbal behavior constitutes condemnation or punishment in virtue of his position within a system of authority.

Consequences

The punisher must regard it as a good that the offender suffer an evil, but need not think that that good is by itself sufficient to justify his inflicting it. The belief that someone deserves to suffer does not imply that he ought to be punished, let alone who should be the punisher. And even if the punisher thinks it good in itself that he punish the offender, he need not think that good is by itself sufficient to justify his punishing the offender, for he may regard other possible benefits, harms, rights, and duties as relevant. A punitive will is essentially hostile but need not be exclusively hostile. While feeling angry at someone, we may have other motives regarding him and other things (including ourselves). These other motives may constrain our

hostility, and we may thus restrain our conduct to maximize its benefits, minimize its costs, respect someone's rights, and so on. A hostile act or attitude may, like any other, have any of many effects and may be regulated with an eye to such considerations. Still, the act or attitude ceases to be hostile if directed solely by some benevolent or irrelevant intent or wish and without any intent or wish that its object suffer.

The good effects of punishing may be good enough to justify inflicting a penalty, but they cannot motivate the hostility of a punishment. That hostility derives, not from the desires for the benefits, but from the frustration of some desires and consequent hatred or anger. For any causal effect, either the hostility is in-

essential for its attainment, or the value of the effect presupposes the propriety of the hostility.

The benefits of punishing we best understand are not effects of punishment per se. Both their causation and their value are explicable without any punishing or any conception of it. Securing our safety is as self-evident a good as any can be, and the means to it may be no less transparent. Incapacitation is effected by incarceration, execution, exile, physical deformation, and the like, and that effect and the benefits it brings are logically and causally independent of the incapacitator's intents, beliefs, and desires. Pretrial custody, preventive detention, psychiatric confinement,

and quarantine incapacitate in the same ways and to the same degree as a punitive imprisonment of the same duration. Similarly, potential offenders are deterred by *their* belief that their suffering a penalty is an evil for themselves; the deterrer can do his job just as well without himself believing or desiring that what he inflicts is a real evil, and without its being a real evil. We have good reason to seek to incapacitate or deter when we punish, and we may have insufficient reason to punish when we have little hope of succeeding, but we do not attain those goals by punishing unless we believe the unwanted behavior to be sufficiently condemnable to warrant the suffering inflicted.

Character

Like deterrence, discipline (correction, forming and reforming character, developing and maintaining respect for the law or the authority, teaching the punishee or the public a lesson, rehabilitation, and so on⁷) has no need to inflict a real evil nor any need of a hostile intent. At most it might be necessary that the *pupils believe* they are suffering or witnessing a punishment. Suffering may educate and cause character change, but its being an evil is not a causal property; and though the appearance of a hostile authoritative intent may sometimes depend on the reality of the hostility, it does not in principle and often does not in practice. The audience learns the same from experiencing or witnessing suffering whether its perception of the event as punitive is veridical or not. Punitiveness as such is inefficacious because the punitive differs

harm, but the efficacy of discipline may depend on the *semblance* of punishment because the purpose of discipline, unlike deterrence, may be something more than instilling a reliable mechanism for producing desirable behaviors. Punishers (particularly parents) and people in general may care about a punishee's character for itself. Though the state may not demand it, it matters to us—and to our morality—that people be just, be motivated by a genuine concern for justice itself, and not merely behave as though they were. Parents (and other punishers) may punish not only to keep their subordinate from being a nuisance to themselves and others or to train him to be socially successful, but also because they love him and perhaps feel pride or shame in the kind of person he is.

Though the disciplinarian need not be hostile to achieve his goals or to seek the goal (shared with deterrence and incapacitation) of inducing law-abiding behavior, he does need to have or approve of a hostile response to noncompliance if he seeks to instill moral character as an end in itself, for the character he values and seeks to instill is precisely one that regards a hostile response to wrongdoing as justifiable. A punisher might truthfully say, "This hurts me more than you," but he cannot be speaking the literal truth when he says, "I'm *only* doing this for your own good." We may punish someone out of love in the sense that we may wish that he not suffer anything which will not benefit him, or in the sense that, were it not for the love, we would be insufficiently motivated to take the trouble to punish. But we cannot punish lovingly, as we might when tending a wound, sympathetically

The good effects of punishing may be good enough to justify inflicting a penalty, but they cannot motivate the hostility of a punishment.

from the nonpunitive in its intent and motive but not in its causal consequences.

Still, discipline may differ from deterrence in its ends and thus in its means. The efficacy of a deterrent depends only on the apparent severity and certainty of

worrying aloud: "This doesn't really hurt much, does it? I hope not. Don't fret; it will all be over soon enough. There now, that wasn't so bad after all, was it?" If punishment is to teach anything more than that

The deterrer can do his job just as well without himself believing or desiring that what he inflicts is a real evil, and without its being a real evil.

there is a connection between being caught and being constrained, then the punisher must convey an attitude of righteous hostility that insists on explaining itself through the judgment that the misbehavior is itself due cause for justified hostility. The character that punitive pain aims to instill identifies itself with that hostility, endorses the authority's conception of the hostility as justified, willingly shares it, sympathizes with the hostility rightly felt by the wronged, feels guilt upon recognizing its own wrongdoing, is pleased by justice done, and so on. Retributivism is implicit in the essential motivation of punishment and in the character it seeks as a consequence. So, the value of the consequence cannot, without circularity, justify its cause.

The same contrast between physical appearance and moral reality reappears at the other end where moral character is presupposed and punishment is valued as a vehicle of community catharsis, a means of venting righteous spleen. As with discipline, purgation is consequent upon perception of the vehicle, not on the accuracy of the perception. (However, with catharsis more commonly than with discipline, the punisher acts upon himself, and thus the possibility of pretense differs.) Moreover, though frustration and pent-up feelings of rage may be instrumentally pernicious, unless our interest in having someone suffer is itself legitimate, the satisfaction of the interest lends little license to the vehicle. As before, the value of the consequence presupposes the propriety of the hostility motivating punishment and thus cannot well justify it.

A punishment must be explained by a condemnation and hence must express it. Thus, it can have the *function* of expressing condemnation, and doubtless sometimes does.⁸ But a punishment need not serve this or any other purpose, and often does not: Consider self-punishment. The expression of condemnation can sometimes have valuable consequences—consequences whose absence would, in some cases, make punishment unjustifiable, but whose presence never by itself justifies punishment. In a theory of punishment the expressive role of condemnation is less central than its explanatory role, for the latter implies that punishment presupposes retributivism.

The Good of Deserved Evil

It is sometimes said that deserving suffering makes no sense, whereas deserving benefits does, because no one need be harmed or wronged by a wrongdoer's not getting his due, and thus no injustice would be done and nothing of value would be lost.

Yet, presumably, justice is itself a good, something we have reason to pursue and promote. It is good that justice be done. However, justice is not itself and need not involve a good for anyone.⁹ Justice is done when everyone is served his due, distributed his proper portion, his proportionate share of goods and evils. But the good of justice is not some further thing that gets distributed. It is not itself possessed by anyone. Justice can go undone though no one is wronged or harmed or made worse off.

When an injustice is done to someone by someone,

when doing wrong is wronging someone, then a right is violated. For liberalism, criminal punishment is proper only in return for the violation of a right, a violation that does some kind of violence to someone.

However, justice may fail to be done because some-

Retributivism is implicit in the essential motivation of punishment and in the character it seeks as a consequence.

one acts wrongfully, though without doing an injustice to anyone, for example, by ingratitude or disloyalty. He may do something wrong without wrong-

ing anyone, for another may have a claim to something and be entitled to claim it from him, but not be entitled to the thing itself. He deserves it, on grounds of principle and propriety, but it is not his, not by right, by rule or law. His due is owed to him and he may demand it, but he does not own it; it is not his to command. His claim may not be backed up by force if the default does not do him violence, injury, or harm. The generous and loyal are not wronged by ingratitude

For a theory of punishment the expressive role of condemnation is less central than its explanatory role.

and disloyalty, not *per se*. An injustice may be done them if their rights are violated, but though there be an additional failing and thus a greater wrong if the right be violated by a beneficiary, it is not a greater wrong to the benefactor. Still, the claim of desert may be enforced by complaint, condemnation, and desertion.

So too, wrongdoers *can* claim their due of retribution, and they *do*—with a frequency and intensity that should surprise no one. Of course, they cannot enforce their claim by force, and insofar as they have already deserted their victims, they are powerless to impose their claim except by complaint. Nevertheless, whether a debt owed be of good or evil, the recipient—and anyone else—can complain that nonpayment of dues is an injustice, though no one is wronged by the default.

Further, justice may go undone without anyone acting wrongly. Some of what is due us need not be anyone's duty to provide. We may have a claim to something we are not entitled to claim from anyone. We

may make a plea, but not a demand. A wrongdoer may deserve and thus be *owed* a punishment (or some other retributively justified suffering); perhaps by his very wrongdoing "he asked for it." This may justify someone in providing it, but often no one has a duty to do so or wrongs anyone by failing to. So too with the notion that guilt is a debt whose payment in punishment is *owed by* the guilty. Debts are obligations, liabilities, but not *per se* duties. Debts can be exacted or be forgiven and cancelled; duties can be neither. Nonpayment of a debt, unlike nonperformance of a duty, is not itself a violation, not normally punishable, not a criminal dereliction, but a civil delict. An injustice is done if someone has the duty to exact payment and fails to; and there the injustice is done to the party to whom the duty is owed, who need not be the creditor.

We speak of a criminal paying his debt to society by having the state exact punishment. However, we do not suppose that states or other human agencies have a general duty to ensure that everyone under their protection is awarded his or her due. Perhaps a state has a duty to protect its citizens' rights and thus is justified in punishing violators of those rights in fulfillment of its duty. And, of course, once a criminal code is in place, officials will have duties of enforcement. But, I take it, while the debt owed to society or the ill deserts owed the criminal may justify the state in punishing, we do not suppose that those owings by themselves impose a duty on the state.

Again, a lack of justice need not impoverish anyone, for the good of justice is not itself possessed by anyone. Yet it can be enjoyed. Our sense of justice can be satisfied. A just man is pleased by the contemplation of justice done. But is a just and decent person cheered by the thought of sinners in hells or criminals in cells? If not, if we can derive no decent satisfaction from doling out the dolorous, then what respectable motive can we have for doing it?

Justification

The punisher intentionally inflicts an evil to do something good. Presumably, if his act is just, he thereby does something good. Yet, so it seems, if an act is to be just there must be some good done other than the good of justice itself. For, if we can find in the situation nothing in whose goodness we can justifiably believe

without assuming the alleged justice of the situation, how could we justify our believing that it is just? But what good of this sort can there be in inflicting an evil for the doing of evil?

Of course, a fallible punisher may, in any instance, be mistaken in believing it is good that the condemned

suffer the inflicted evil. But there is no a priori reason to think his belief cannot ever be true or to doubt that it could be. At least we know of no valid reason to that

*Justice is not itself and need
not involve a good for anyone.*

effect nor any reason sufficiently compelling to convince anyone not predisposed to the doubt. That belief can sound paradoxical, but it is not self-contradictory or logically incoherent.¹⁰

The air of paradox can be dispelled by considering the structure of anger (or hatred). Anger is an innate, aggressive response triggered by the perception of a (threatened) harm and directed toward its apparent source. We are endowed with this response pattern because it is adaptive; our having it promotes our survival. However, we do not (normally) choose to be angry as a means of self-protection. The internal aim of the anger itself is (threatened) harm to its object. A self-conscious being regards harm to the object of its anger as a good in itself, in the same way that he regards the fulfillment of any of his desires as a good. And though the anger is not motivated by a justification, if the anger is justified by the harm being wrongful, then the belief and desire involved in the anger are justified. Seen in this light the punisher's belief is hardly counter-intuitive. Such a story makes it understandable, sketches an explanation of the belief. Of course, the explanation is not a justification, since the last line of the story presupposes precisely the principle in question. Yet certainly the explanation does cast doubt on the assumption that the punisher's belief is a peculiarly dubious moral belief.

Nevertheless, we think punishment is specially in need of a justification. This is a curious fact about us. It is curious, first, because, obviously, the only thing in the world that can need or in any way benefit from a justification is a person. We project *our* needs for justification onto the world.

It is curious, too, to say that the punisher needs a justification, because, as we have seen, he has one and knows that he does, as he must if he is really a punisher. Punishers lack nothing essential for being punishers. If we doubt whether what they have is really good enough, then presumably *we* should have

some conception of what more is needed and what it is needed for.

What is especially curious is that our culture, and not just the theorists among us, feels a need for a justification of the punisher's belief. After all, it is the only one of our basic beliefs shared by every human society about which so many of us feel so personally anxious. More peculiarly still, our worries are not transcultural; indeed, no other people seem to have suffered the anxiety we do. We seem more anxious now whether we punish too severely or indiscriminately; in this we may be distinctive mainly in degree, but this is due largely to our ambivalence about punishing at all. Occasionally, prophets of old may have awakened some recognition of the presumptuousness and other perils of exacting retribution, but in our insistently consequentialist culture, people commonly are, in varying degrees of consciousness and consternation, unsure whether the very idea of retribution makes any sense. This modern worry is more republican and more radical.

Theorists now standardly identify our doubts about retribution as due to a certain circularity. Hugo Bedau has argued, for instance, that:

Either [the retributivist] appeals to something else—some good end—that is accomplished by the practice of punishment, in which case he is open to the criticism that he has a nonretributivist, consequentialist justification for the practice of punishment. Or his justification does not appeal to something else, in which case it is open to the criticism that it is circular and futile.¹¹

Bedau has called this a "familiar dilemma." But it is all too familiar due to the logic of justifying any moral principle. For, if a consequentialist justification must begin with some ultimate consequentialist principle,

*A lack of justice need not impoverish
anyone, for the good of justice is not
itself possessed by anyone.*

its first principle cannot be given a consequentialist justification which is not circular or futile. And if a nonconsequentialist principle must be justified by beginning with some ultimate nonconsequentialist principle, its first principle cannot be given a nonconsequentialist justification which is not circular or futile.

To suppose it worth posing the dilemma in consequentialist terms presupposes some question-begging consequentialism.¹² This formulation of the formal problem in moral theory is not as perspicuous as others.

Consequentialism is a thesis about the structure of justification of principles of rational choice and action.

The past calls to us through our present normative commitments.

Its distinctive position is that the principles of justice are justified, if at all, by their realization being a good, a good other than justice itself, a good whose evaluation as a good we can justify without presupposing anything like a principle of justice, a good we can justifiably believe is greater than the good we can bring about without the realization of the principle of justice. That is not an implausible idea, but it is not unproblematic.

Now, the question, "Why should I be just or moral?" can naturally arise in anyone's mind when he thinks he has some reason not to be. For the theorist the question exists by itself, asked or unasked, motivated or not, like any abstract, formal object, except that this one is of special interest to an ethical theorist. It is a formulation of the most general and fundamental question about moral motivation and justification. And it poses a familiar dilemma. Either a just man's reason for choosing to do (what, upon reasonable deliberation, he believes to be) justice is simply

that it is (what, upon reasonable deliberation, he believes to be) justice, or he has some other justification. The first alternative seems to beg the question, but it also seems to be the only possible answer, for it defines the essential condition of his being a just man. If he is motivated, instead or in addition, by some other reason, either he is not a just man or that reason is subordinated to and thus cannot be his justification for his dominant reason. Incidentally, this dilemma is not peculiar to practical reason; the structure of the dilemma applies to rational thought in general.

Now, when we are not condemning the particulars of the sentences our predecessors imposed, do we have reason to think they should have been as worried as we about the punisher's belief? A reasonable person does not feel a general need to justify each of his beliefs, especially if the belief is among his deepest and most widely shared convictions, because such convictions are ipso facto reasonably presumed to be true. Despite and because of our fallibility and finitude, it is unreasonable to doubt or bother to defend such beliefs until confronted with some reason to do so, some conflict with another reasonable belief. The fact that no attempted justification has been satisfactory cannot by itself be a reason for doubt, no matter how numerous and blatant the failures, for we have no clear conception of what should satisfy us as a justification of any belief as basic as the punisher's, nor any examples of a satisfactory justification of a comparably basic belief in or out of morality. No theoretical dilemma posed by the justification of the punisher's belief justifies our doubting that belief. More to the point, no such theoretical puzzle could *explain* our doubting that belief.

Retrospectivity

The modern mind was liberated by freeing itself from acknowledging any inherent authority, practical or epistemic, in custom, tradition, precedent, and the past. The fact that something had long or always been done or believed ceased to be sufficient proof that it ought to have been, and still less that it ought to continue to be. The liberal mind seems distinctively more systematically consequentialist, more disposed to suppose that only the future matters, for only in the from-

here-on-in can anything still be done to better or worsen our lot. Such a mind is tempted to say, "The past is over and done," as though a tautology entailed that past wrongs must lose their potency for a rational man.

Yet, however rational it may be to require that the maintainance of a norm be justified by its continuing to do more good than harm, it is myopic to regard *every* part of the past as utterly dead, over and done. Viola-

tions of and conformity to a norm must precede the responses to them. For a norm to remain in force, violations of it must continue to justify negative responses to them. Thus the past calls to us through our present normative commitments. We may have reason

to restrict our responses with statutes of limitations, principles of forgiveness and the like, but to prohibit or forgo all negative responses is to nullify the norm. We must hold ourselves to be held by the past if we are to be rational beings in the present or future.

Responsibility

A popular explanation of our qualms regarding retribution is that modern science has driven us to be determinists and thereby denied us the presuppositions of blame. Yet, as a matter of logic, it matters little to our culpability whether we go crooked but for the grace of God or of good genes. (The causal stories we tell may be more detailed, longer, and less fictional, but their nexus with the will is no closer.) And, as a matter of history, our era is notable as much for its insistence on the reality of individual responsibility and its faith in knowledge-as-freedom as for its valuation of autonomy. If we are the ones who started supposing that everyone's life is a throw of the dice, then why is ours also the culture which is uniquely circumspect, stringent, and insistent about distinguishing the culpable from the exonerable?

Our retributivist worries about responsibility run the reverse of what the bogey of determinism implies. For the liberal, the more reason we have to punish, the less reason we have to punish. Liberal antipaternalism denies us the use of punishment (or any coercive force) as a means of character reform for the reprobate's sake—as distinct from the community's—precisely to the degree that such punishment is most proper; that is, to the degree the culprit is a fully responsible agent. Conversely, the less justice there is in punishment, the more justified we feel. No punishment seems purer than those with which loving parents scrupulously impart some self-mastery to their impulse-ridden children. When a reluctantly punishing papa truthfully

says, "This hurts me more than you," the pains he refers to are not pangs of conscience, nor should they be.

A more plausible account of the origin of our worries about retribution would focus on our altered attitudes about anger. The density and discipline of urban-bureaucratic-industrial-technological life are specially strained by anger acted out in its midst. We learn to be thoroughly frightened of anger, particularly our own.

*Liberal antipaternalism denies us the use
of punishment as a means of character
reform for the reprobate's sake
precisely to the degree that such
punishment is most proper.*

We are trained to control, suppress, and repress anger. We come to condemn anger, not only its violent expression, but anger itself in all its forms (unless, perhaps, it is detoxified by being institutionalized). We deprive ourselves of anger's best objects, deny the devil his due, and deny any theory of man that recognizes unmotivated malevolence. As a result, our compassion, uncompromised, makes us quicker to excuse and even condone.

Vengeance

We wish the good of punishment to be like other goods—something we can enjoy, and whose enjoyment we can enjoy, feel good about enjoying, and feel

ourselves to be good in enjoying. Elsewhere, we relish righteousness triumphant. But for us, unlike all but our rather recent forefathers, a pall hangs round the

pleasure of punishment.

We still allow a joyous shout for our successes in sleuthing, chasing, catching, and convicting the nasty and nefarious. Perhaps even more than those before us, we cheer and are cheered by the drama and the sport when justice and truth itself win out, injustice is overcome, confronted, captured, and convicted. But for us, once the culprit is at our mercy and our moral certainty, and the time for condemnation has come, once the game is over, so too is the fun.

The pleasure is gone from punishment. But how could there have been any in punishment itself? Punishment seems tasteless. There is nothing to savor. Guilt and indignation have no flavor. Resentment is bitter. Revenge is sweet. So a common worry is whether punishment is vengeance or whether it must be motivated by vengefulness.

Punishment cannot be animated without any animus, and vengeance does provide a specially satisfying explanation of punishing. Revenge inherently satisfies a self-interested motive. It fulfills the demands of the pride of the avenged (and of anyone who proudly identifies with his), and provides the pleasures of glorification. Its internal aim is the vindication of honor, so the act and its motive are relatively readily intelligible. Vengeance is a display of power that disconfirms and invalidates an imputation against the avenged's honor, an imputation that the avenged event appears to make and provide evidence for. That event seems to say that the avenged person is so worthless that he need not be reckoned with, so impotent and inconsequential that he can be offended with impunity. That imputation is disconfirmed by vengeance, by showing that the avenged's worth is a reality recognized by the world, a reality evidenced by the world's being so constituted as to vouchsafe that worth.

More importantly, no other normative, self-interested motive shares the distinctive aim of punishment, the suffering of its object.¹³ And, assuredly, authorities are naturally prone to a pridefulness that competes with the pridefulness of their subordinates. The pride of an authority (and its agents, and its people) in its power may be stung by sheer disobedience (even by waywardness, let alone rebelliousness), and punishment is the perfect means of avenging the insult.

Yet, even though punishment can, perforce, be a means to that end, that implies that punishment cannot

be vengeance, identical with it or a subspecies of it. The two differ categorically in their conceptual and psychological structures. Vengeance looks not at the guilt of the offender but at the shame of the avenged. It is not essentially impersonal.

Unlike punishment, vengeance is not essentially communicational: The self of shame is not committed

Vengeance looks not at the guilt of the offender but at the shame of the avenged. It is not essentially impersonal.

to a world of communal norms. Vengeance does have the structure of a speech act, but in the way an oath does—that is, as a pledge of honor by which one binds oneself by one's word, a personal performance that need not be witnessed by anyone. It suffices for vengeance to be done that the avenger succeed in his intent to inflict the requisite suffering. No one, not even the avenger, need believe he succeeded, and no one but he need believe he even tried.

The engines of natural anger and hatred are available to propel the avenger. But no assault on the self can stir it to punish without the depersonalization contained in conceiving the assault to be a wrong. The modern mind demands that its thought and actions be ruled by reason, an impersonal power blind to the particularities of self. Here the urges of natural anger and hatred must be detached from the self; their gratification is channeled and checked, and the pleasures of inflicting suffering are distanced and subject to doubt. The structure of vengeance often precludes it from being a motive for punishment, particularly in the modern political and personal world, where the agencies of state and self aspire to impersonality.

Modern rulers must relate to citizens through law, not through the personal relations that kings and patriarchs claimed.¹⁴ Then too, secular civil authority no longer presumes to be the great Doler of Deserts, Nemesis, or anything of the sort, however noble the role may be.

Punishment, law, and guilt must be conceptualized in aural terms—the voice of authority, the voice of a sentencing judge, a scolding parent, a wrathful God, a

guilty conscience. Vengeance responds to shame, a response of fearing to be seen. Vengeance responds to and wipes out a stain, a blot, a visual discordance. The only criminal punishments we still witness are executions, and there the audience is select—the general public is provided with only a verbal account.

A culture aspiring to impersonal control fosters feelings of guilt, not shame, and breeds a need for self-punishment and nurtures its indulgence; but the very concept of self-vengeance is incomprehensible. ("Avenge oneself" is pointedly not parallel to "punish oneself.") Moreover, the self-centeredness of vengefulness deafens it to pleas of pity, makes it immune to the sympathies that can restore the bonds of community. To punish we need inflict only *some* suffering. But vengefulness enjoys no satisfaction until it has served desert; to avenge, the score must be evened, no less than an equivalent evil inflicted. An avenger can show mercy only by limiting the undeserved evil he has the power to impose. And at the other end, while a punishment may be excessive, greater than deserved, if the punisher himself believes that the suffering he inflicts exceeds what is due, then that overplus of suffering is not a punishment, but something quite different. Punishment per se sets an internal limit, an upper bound but no lower one (though one may be inserted by desert). Vengeance sets a lower bound, but no upper one. If you beat me at tennis, I may try to win the rematch, and failing that

I may try to beat you in golf, and failing that I may torch your car. Such behavior gives revenge a bad name.

By now it is well-nigh analytic that vindictiveness is vicious. In our language *vindictiveness* refers to a vice, an extreme, the Aristotelian mean of which is the virtue which seeks vindication for the right reasons, in the right way, at the right time, and so on. The name of that virtue is *honor*. Vengeance is now less respectable. Honor is now less esteemed; its status shifts when the public morality gives pride of place to an impersonal *respect* for persons as its first principle. The biblical God denied his followers vengeance, but did not damn vengeance itself. He arrogated to himself the propriety of pride and the pleasures of protecting it. Earthly rulers have done the same. Modern nations deny the validity of violence in defense of honor, except, of course, their own honor. The results of all this are mixed, not uniformly delightful. Our lives are more secure, more limited by law, more impersonal, and more pervaded by a sense of impotence. The valuation of courage has changed, and with it so has our language: The remarkable role of *dare* as a modal auxiliary verb is becoming archaic. At the same time, the cost of increased ambivalence about punishment may rise. Perhaps now it is natural and common for people to see vengeance in punishment in the way we commonly see in a person's face the face of another we wish to see.

Love and Respect

The problem of reconciling condemnatory hostility with love has bothered Christian moralists for a millennium.¹⁵ The puzzle is not *whether* we can be angry with someone we love, but *how* it is possible. The condition is as common as self-deception or weakness of will, and no easier to make coherent. The puzzle will seem sheer mirage if we misperceive love or anger—if, for example, we imagine these feelings to be episodic sensations stirring us at separate times and places. That, however, explains the possibility at the cost of precluding the real conflict in our wishing to harm someone whose welfare we wish to protect and promote.

The intellectual puzzle is real and realized as a practical problem, a felt conflict evidenced by the pervasive and profound resistance we have to acknowledging to others—or to ourselves—the reality of our anger or resentment toward someone we love, or someone we think we ought to love or be grateful to or sympathetic with. The conflict is commonly felt without being identified, because the reality we resist acknowledging is the pleasure we would or do take in seeing the suffering of the other, and because the nature of the pleasure and of the seeing takes so many forms. We enjoy recognized fantasized events we would be traumatized to see realized. And we enjoy fantasies and realities with-

out identifying the object and cause of our enjoyment.

Wisdom may say that we should love our enemies but not that every soul can succeed at this. Some of the wise say that nothing less than the grace of God is needed to accomplish this. Wisdom does not demand that we let no one be our enemy, for it is not always up to us whether someone sets his will and force against what we love. It might be up to us whether we feel enmity for our enemies, for we may have the choice of refusing to recognize the reality of their enmity. We may acknowledge that someone is hell-bent on injuring our beloved while lovingly dismissing such antics as rash, brash, misdirected: He knows not what he does. But how do we assume that stance toward those we respect as equal, autonomous members of our community and commonwealth?

Our situation is exacerbated because, unlike our pre-liberal Judaeo-Christian predecessors, we profess, and occasionally possess, a commitment to a principle of inalienable respect for persons in addition to, and perhaps in priority to, a principle of inalienable love for our neighbors. None but the smugly self-righteous supposes it is obvious that penning someone in a cage for two or twenty years is consistent with (let alone betokens) the operation of *either* principle, yet we now feel a need to reconcile it with *both*, and each presents its own problems.

*Love and punishment conflict
because the internal
aims of love and anger
are essentially opposed.*

Plausible stories may be told of how nonretributivist rules of justice might be consistent with and even expressive of, if not motivated and explained by, some kind of universal love or respect of all human beings. It may be in the interests of every human being, qua human being, and/or in accord with the will of every human being, qua rational agent, that we govern ourselves by certain proscriptions. And if they are genuinely normative, they must justify some kind of negative response to their violation. The conundrums come

when we try to reconcile that love and respect with the aggressively hostile character of the responses to wrongdoing respected by morality and justice.

The conundrum is that the very character of the primary rules as expressions of love or respect explains all too well the character of the aggressively hostile response. Can you love someone without being disposed to feel hatred toward those who threaten your beloved by willfully violating the rules protecting him? Can you respect someone's will as a rational human being if you do not respect, regard as justified, and endorse the outrage he feels toward those who violate the rules authorized by his rational will and thereby violate his rights?

However, to show *why* a response of angry hostility toward violators may be *necessitated* by the nature of norms expressive of universal love or respect does not help to explain *how* that response is *possible*, reconcilable with the demands of love and respect. Our dilemma is not dissolved by our being disposed to bear the same hostility to someone who does the violator an injustice. What needs explaining is how our intentionally inflicting or even wishing suffering on someone due to his guilt can be expressive of love or respect for *him* and not just of love and respect for his victims. After all, condemnation and punishment hardly presuppose love or respect for the wrongdoer: Consider the punishment of pets and slaves.

We may desire reciprocation from, or the perfection of, someone we love or respect, and thus an extra increment of anger and hate for a wrongdoer may be motivated by a prior or persisting love or respect for him. That extra fury only magnifies the intensity of the conflict and does not clarify the compatibility of love or respect with the quantity and quality of condemnatory hostility directed at a wrongdoer independent of our loving or respecting him.

Yet only psychological naivete causes anyone consternation about punishing children. We may argue over the best manner and means, the proper severity and occasions, but we do not doubt that, if punishing a child is truly good for the child, if it is necessary for developing the self-control requisite for independent survival or the sense of justice requisite for a respectable life, then punishment is permissible, indeed a parental duty. When we sense that punishment is compatible with and can be motivated by love, we sense no moral dilemma. And we do experience that possibility being realized in some parental punishments. We

may, if we look, see the conceptual puzzle about the possibility, but when we see the reality, the paradoxes are like Zeno's, incapable of making us question the real possibility.

However, when we punish our peers, a conceptual puzzle pinches a moral nerve. If we see any love at all for the convicts in our criminal justice system, it appears only in the form of constraints upon the severity of treatment; and actually the limitations we now place on the acceptable manner and means of criminal punishment seem generally as well or better explained by some form of respect rather than by something like benevolence. In any case, to explain why we do not do worse to a convict is not to explain how worsening his lot at all is compatible with love or respect. We see ways to ameliorate the conflict, but no way to motivate a punishment for our peers while respecting them as responsible moral agents.

Love and punishment conflict because the internal aims of love and anger are essentially opposed. Respect is not an emotion; it has no aims or desires. Rather than motivating the will, it constrains it. Though respect is not antithetical to anger or hatred, it does constrain maleficence. Yet punishment troubles our respect for persons even more than our love for them.¹⁶

It is no accident that enslavement has been a prime form of civil punishment throughout human history. Punishment is like enslavement, even when the sentence is not a servitude of forced labor in the service of the authority and even when a prisoner sits in a cell serving or doing nothing but time, his own time. Like enslavement, civil punishment imposes a condition of total subordination, a profound social and political inequality radically unlike anything acceptable in a liberal democracy. Mutual love does not require equality. Mutual respect does.

Our pre-Enlightenment predecessors had little difficulty legitimizing punishment and slavery, since any degree of socio-political inequality is—in theory, if not in practice—compatible with the transcendental equality bestowed on us by our being equally undeservedly beloved by God. That love which we were enjoined to imitate had no principled qualms about subjugating or injuring the beloved as long as this could work to benefit the beloved, and—in theory, if not in practice—any bondage or burdens in this world could be a boon to the bonded or burdened.

A parent can love his children and seek to secure

their interests without respecting his children—that is, without acknowledging that they have the cognitive and volitional powers which justify respecting their judgments and decisions and refraining from interfering except when they threaten others. Anger can coexist with mutual love partly because lovers can be reconciled through a mutual understanding reached by one party humbling himself before the other, submitting to condemnation, begging forgiveness, recognizing and accepting an inequality. This is not an inequality in value: A punishing parent might willingly sacrifice himself for the sake of his child. It is an inequality in the powers and authority that can demand respect.

There is nothing new in the notion of an inalienable transcendental worth of persons which is independent of a person's moral character and of his cognitive and volitional capacities. This worth can be possessed by infants, idiots, and the insane, and even perhaps by a zygote. The politically and culturally revolutionary conception of respect for persons that blossomed in

*Mutual love does not require equality.
Mutual respect does.*

the Enlightenment accords a kind of sovereignty to every "normal" adult, everyone whose cognitive and volitional capacities are developed and functioning beyond some (problematically defined) minimal level.

To many minds, subjecting a legally competent agent to a regimen of enforced psychotherapy is hard or even impossible to reconcile with a proper respect for persons. To some, the propriety of coerced character reform largely depends on the specific procedures employed.¹⁷ However, understandably, to many people such worries seem a bit precious given the inevitably radical lack of respect inherent in imprisonment. And surely the devotion of many more people to the nobility of autonomy would be badly taxed if they could maintain a faith in the efficacy of the therapy and its cost-effectiveness in serving the civic goal of crime reduction.

Of course, if the community is to be the primary

beneficiary of punishment, then some strong showing needs to be made for sacrificing the guilty for the greater good. But more worrisome still is that the more we regard our convicts with respect and not merely with sympathy for their penal plight, the more we must take seriously their point of view, and the more we must be troubled by the unsurprising parallel between social deviance and social disaffection and disenfranchisement. Unlike our predecessors who suffered at most pangs of pity for the objects of their com-

munal wrath, we who profess respect for those in our social cellars are not so certain of the justice of punishing persons for breaking the rules of a "game" they are compelled to play and doomed to lose. And behind that worry lies a doubt too unsettling for public debate: If the concepts and principles of retribution are inherent in the concepts and principles of justice, and if they conflict with our concept and principle of respect for persons, then what should we conclude about the latter?

NOTES

This paper is composed of some condensed sections of a work in progress under the same title. I am indebted to my editors, the Baiers, and the SELFers.

1 If this way of formulating norms seems troublesomely unintuitive, a rationale and framework are provided in my *The Significance of Sense* (Ithaca, N.Y.: Cornell University Press, 1972), Chapters 3 and 4. Note, if a singular command, such as "Platoon halt!", is (regarded as) normative, it presupposes some norm of the form: If *s* is commanded by *A*, then *s*. See the section entitled "Authority" below.

2 See Aurel Kolnai, "Forgiveness," in his *Ethics, Value and Reality* (Indianapolis: Hackett Publishing, 1978), pp. 211-224.

3 It is an evil for *the condemned*, something in itself contrary to *his* good. The condemned need not believe it is an evil for him, but the condemner must rightly believe it is. Further, though it must be *in itself* contrary to his good, it may have compensating, beneficial consequences for him or others, and the condemner may intend that too. However, the condemner cannot be intending the evil *only* as a means to such benefits.

4 The issues of this and succeeding paragraphs are discussed more fully in my "Capital Punishment," presented at Somerset College, Fall 1977, and my "The Psychologic of Punishment," presented at The Society for Ethical and Legal Philosophy and at the CMU—U. of Pittsburgh Philosophy Discussion Group, Spring, 1978.

5 See *The Significance of Sense*, pp. 70-72, 119-120, 161-163.

6 See H. P. Grice, "Meaning," *Philosophical Review* 66 (1957): 377-388; and Grice, "Utterer's Meaning and Intentions," *Philosophical Review* 78 (1969): 147-177.

7 *Discipline*, as I use the term, is directed at affecting character. Often, what is called rehabilitation includes vocational training and the like. However, while teaching useful skills may be prudent penal policy, analytically it is best conceived as akin to deterrence, providing incentives for compliance. So too, when people speak of teaching someone a lesson punitively, they may have in mind either discipline or deterrence. *Deterrence*, as I use the term, is not aimed at character change, but rather motivates through existing desires and aversions.

8 See Joel Feinberg, "The Expressive Function of Punishment," in his *Doing and Deserving* (Princeton: Princeton Uni-

versity Press, 1970), pp. 95-115.

9 Justice, here, is an attribute of states of affairs, and derivatively of acts and laws and institutions conformable to or causing them. The root notion of the just, like that of the right, refers to a property of states of affairs representable by a proposition: It is just (or right) that he suffers. (See *The Significance of Sense*, esp. Ch. 5.) Justice is also a virtue of persons, a trait of character, (roughly) an effective disposition to do what is just because it is just. The personal virtue is something in virtue of which the person is a good person. Whether it be in some further way a good *for* the person is another matter.

10 A punisher need not believe that two wrongs make a right, for he believes he does nothing wrong. He rightly believes that to inflict an evil is not ipso facto to do evil, something wrong. And he rightly believes that if the inflicted evil is deserved, inflicting it intentionally does not make the inflicting wrong.

11 Hugo Adam Bedau, "Retribution and the Theory of Punishment," *The Journal of Philosophy* 75 (1978): 616. For another statement of this well-known argument, see Stanley I. Benn & R. S. Peters, *Social Principles and the Democratic State* (London: Allen & Unwin, 1959), pp. 175-176.

12 See Ted Honderich, *Punishment* (London: Penguin Books, 1971), pp. 40-41.

13 Certainly, you might enjoy witnessing or causing someone's suffering without caring whether it is a punishment. And raw anger or hatred or refined sadism can propel a punishment, but, like vengeance, only if abetted by a pridefulness that permits the punisher to feel justified in employing punishment as a vehicle for venting such urges.

14 See Kurt H. Wolff, ed., *The Sociology of Georg Simmel*, (New York: Free Press, 1950), pp. 250ff; and Max Weber, *The Theory of Social and Economic Organizations*, ed. by Talcott Parsons (New York: Free Press, 1947), Part III.

15 For example, see Augustine, *The City of God*, I, 9; XIV, 6, 15, 19; XV, 25; XIX, 13, 16; XXI, 18. For Aquinas, see *Summa Theologica*, II, part 2-2, question 108.

16 On the contrast between love and respect, see my review of Alan Donagan's *The Theory of Morality* in *Nous* 17 (1983).

17 Cf. Herbert Morris, "Persons and Punishment," *The Monist* 52 (1968): 475-501.