The Strengths and Limits of the Theory of Retributive Punishment

Kurt Baier
University of Pittsburgh

Follow this and additional works at: https://digitalcommons.brockport.edu/phil_ex

Part of the Philosophy Commons

Repository Citation
Available at: https://digitalcommons.brockport.edu/phil_ex/vol8/iss1/3
Baier: The Strengths and Limits of the Theory of Retributive Punishment

KURT BAIER
Professor of Philosophy
University of Pittsburgh
THE STRENGTHS AND LIMITS OF THE RETRIBUTIVE THEORY OF PUNISHMENT

by

KURT BAIER

In textbooks on punishment one usually finds four major "theories" or "justifications" of punishment: (1) the retributive, (2) the deterrence, (3) the reform or rehabilitation, and (4) the incapacitation or social defense, theories. They are usually offered as rival theories of the proper (primary) purpose or function of punishment. And it is generally assumed that the general practice of punishing people and individual acts of punishment are morally justified if and only if, and to the extent that, they serve that purpose or perform that function.

During the fifties and sixties, the most popular theory both among social scientists and philosophers was, I think, the rehabilitation, and the least popular the retributive theory. During the last 5 years or so, however, with the growing conviction that there were no known effective measures or rehabilitation, and with the steadily rising figures in violent and other crimes, the popular and professional moods changed. More attention is now paid to the interests of actual and potential victims of crimes. Moreover social scientists have come up with evidence, rather dubious in my opinion, but widely publicized, that significant reductions of crimes could be achieved by increasing the certainty and the severity of punishment. Theorists divide over whether this reduction is better achieved by way of the incapacitation or the deterrence effect of imprisonment. And to some extent these aims of imprisonment involve different imprisonment strategies. Thus, the social scientists who believe in imprisonment for incapacitation believe that the important crimes, crimes against the 'person' (violence) and against property are committed by a fairly small proportion of the population who engage in a career of crime, between the ages of about 20—35 or so. Only incapacitation will stop them. Deterrentists on the other hand believe that a large number of criminals commit crime for the benefits or gains, whether material or psychic, which they hope for from criminal acts and that they can therefore be deterred by the knowledge that crime will not yield them these benefits. The problem group are the 1st-offenders, for more than half of them do not repeat crimes. Should 1st-offenders be dealt with lightly or severely? Should we let them off lightly, thereby failing to incapacitate the 50% who will become recidivists but keeping prisons freer for the hardened cases and avoiding the risks of turning 1st-offenders into hardened criminals by the association with other hardened criminals? Or should we punish them severely, with the converse results? Can we perhaps tell which 1st-offenders will become recidivists and if so, is it just as well as expedient to treat them differently?

While among social scientists the most popular theories now appear to be incapacitation and deterrence, among lawyers and philosophers, as well as, I guess, the general public, there seems to me a distinct anti-utilitarian and to some extent a pro-retributivist trend. I want to talk about this revival of and new interest in the retributive theory, and offer a few hunches about its attractions.

I shall, for clarity, distinguish between Pure and Impure Versions of Retributilism. The Pure Theory holds that the practice of punishment and any individual act of punishment is morally justified if and only if it conforms to all four of the principles of retributivism.
(a) All those convicted of a wrong-doing or crime deserve punishment;
(b) only those convicted of a wrong-doing or crime deserve punishment;
(c) the severity of the punishment should not be less than the gravity of the crime;
(d) the severity of the punishment should not be greater than the gravity of the crime.

On this view, all punishment and only punishment which conforms to these principles is deserved, therefore just, and therefore morally justified. The theory answers both (i) the question of what justifies punishing anyone, namely, that all wrong-doers deserve moral condemnation and deprivation of certain rights, and therefore should be punished, and (ii) the question of the extent to which they should be punished, namely, as much as they deserve, that is, as severely as their wrong-doing merits.

Impure theories are either what I shall call negative or positive versions. (i) Negative Retributivism accepts principles (b) and (d). It is offered by humane authors, such as H.L.A. Hart in opposition both to the Pure theory which adopts also (a) and (c), and to utilitarianism which ignores even (b) and (d). It rejects the Pure theory because it wants to allow utilitarian considerations to take the place of principles (a) and (c). It allows that it is right for some wrong-doers not to be punished at all or not as severely as they deserve, if thereby a lot of good can be achieved or evil avoided.

And it protects us against utilitarian governments who may wish to punish some who have not yet committed a crime but are predicted to commit them or to punish them more severely than they deserve (indeterminate sentence, parole discretion, etc.).

Positive Retributivism, by contrast, allows that for the good of society it may be necessary to violate principles (b) and (d), but insists that principles (a) and (c) must never be violated.

I shall concentrate on the Pure Theory. I want first to show how very plausible this theory is. It seems that we need only to examine carefully the nature of punishment in order to become convinced of its soundness.

What, then, is punishment? Punishment involves two main roles, recipient and imponent, and something, a deprivation, which the recipient suffers normally at the hands of the imponent. Finally, there is some condition (e.g., the commission of a wrong, a crime, or an offense) such that the deprivation suffered by the recipient amounts to punishment if and only if he suffers it on account of such a condition. It must also be mentioned that given cases of punishment may be flawless (deserved, just, effective, justified) or flawed, and that if we attempt to state the necessary and sufficient conditions of something being punishment, we must distinguish between those of flawless punishment and those of punishment whether flawed or flawless.

Let me now say a few words about each of these points to show how plausible it is to interpret them in a way that supports the retributive theory.

(i) The core: Being punished consists in undergoing something normally unwanted. Receiving money could not therefore count as being punished, even if one hated it. Conversely, the fact that, let us say, imprisonment happens not to be unwanted by the person on whom it is inflicted, does not ipso facto disqualify it from being punishment. Of course, we may want to say that it is then flawed: perhaps unsuitable, perhaps ineffective, and so on.

Punishing someone consists in imposing on him something normally unwanted.
What is imposed must be intended as something unwanted: one cannot unintentionally or accidentally punish someone,\textsuperscript{15} but it would be unintentional punishment if what is inflicted were not intended as something normally unwanted. Punishment is a deliberate punitive measure and so must be intended as something deprivatory and so normally unwanted.

(ii) The relation between the two roles: Suffering something normally unwanted cannot be suffering punishment unless it is suffered at the hands of someone who inflicts it as something unwanted. The fact that someone deserves his suffering, does not by itself make that suffering punishment. Even if the suffering is "his own doing" and he deserves it—if as we say "he had it coming to him"—this is at best punishment in an extended or loose sense; as when a judge decides not to punish for larceny, on the grounds that "the culprit has been punished enough already,"—someone who as a consequence of an affair with his boss' daughter has lost his job, his wife, and his self-respect, and has been forced to steal to keep alive.\textsuperscript{16}

(iii) The social context: Punishment consists in something, deliberately depriving someone, which is prima facie wrong, but yet the word 'punishment' implies that it is something which is not prima facie wrong. Wherever the imposition of a deprivation is exhibited as punishment we take it to be prima facie justified, unless it is flawed (unjust) punishment. This means that those of the punished person's rights must be suspended which would otherwise be violated by the deprivation. In primitive societies, the suspension of rights is the only thing the society can manage to organize: the actual deprivation is left to individual initiative. When a person is outlawed, any other person may kill him and none of the duties of mutual aid apply to him. In more highly organized societies, the suspension of the individual's rights is only relative to the organs of punishment. The fact that a person has been condemned to death does not give anyone (except the executioner) the right to kill him.

This suspension of rights is always for a specific time; only in extreme cases is it permanent. The notion of expiation or atonement which is traditionally attached to the concept of punishment implies that during the time the punishment (the suspension of rights) lasts, the recipient is not a member in good standing, but when the punishment has been "served", he is received back into society as a member in good standing. This is true not only of legal punishment, but also of all other forms. The child under punishment is not a member in good standing in his family and, in so far as the society at large approves of the punishment of children by their parents, the child is not a member in good standing while he serves his punishment also in the larger social context.

(iv) Moral disapproval: Joel Feinberg\textsuperscript{17} has recently revived the idea\textsuperscript{18}, for some time ignored under the influence of Legal Positivism, that for the imposition of deprivation to be punishment it must express the moral disapproval or condemnation by the imponent of the conduct for which the deprivation is imposed. A system which attaches penalties to breaches of rules and then imposes those penalties on all and only on people whom it has found guilty of a breach of the rules, looks like a system of punishment but is not unless the imposition of penalties at the same time signifies moral condemnation by the imponent. If this is sound, as I think it is, then it shows that the idea of punishment is essentially at home in morality rather than in law, even if it is true that the law is among other things the enforcement of morality. This point is important, as we shall see more clearly below.

Both punishment and penalization must be distinguished from the social imposition of some burden, such as the heavy excise many states impose on smoking and drinking. These impositions differ from punishment and penalization in that they
THE STRENGTHS AND LIMITS OF THE RETRIBUTIVE THEORY OF PUNISHMENT

leave the individual morally free to choose between the course which carries certain pay offs as well as state-imposed burdens and the course which carries neither the pay offs nor the burdens. It would be wrong to think of this tax as a punishment or penalty for smoking or drinking (though someone opposed to smoking and drinking may well say that the penalty of, not for smoking, is lung cancer, of drinking, cirrhosis; this is a prudential warning against these practices, not a threat of reprisals) since most states count on people going on smoking and drinking as sources of state revenue. To speak of some imposition as a punishment or the imposition of a penalty is to imply that people are not (or generally are thought not to be) morally free to make such a choice, though of course they normally can choose what they are not morally free to do. Punishment expresses permissive moral condemnation of the conduct punished and is therefore inconsistent with a positive attitude of the imponent towards its recurrence or continuation, while taxation is perfectly compatible with it.

(v) “For” something: There is fairly general agreement that if someone’s suffering of a deprivation at the hands of another is to be punishment, then there must be something on account of which the deprivation is imposed, something it is for. Some have taken this to be merely a point about the meaning of the word ‘punishment’, others interpret it to mean that punishment cannot be just and therefore not justified, unless it is inflicted for something which is capable of providing such a justification.

a) The definitional point is settled comparatively easily. We can say that no deprivation can be punishment unless it is imposed on someone only after he has been properly found to satisfy the condition for which the system of punishment imposes it. This does not mean that the punitive system, say the parental one, must have spelled out in detail what these conditions are, but there must be some understanding on the part of the recipients of what these conditions are. The system is flawed to the extent to which this is not true. In a crude “common law” system, the only developed legal function is that of judge, i.e., of the official who “finds” (authoritatively determines) whether or not someone has broken the law. But this does not mean then because there is no legislator who has spelled out the rules, the people do not know, at least roughly, what these rules are.

From this it plainly does not follow that a person suffers punishment only if he actually satisfies these conditions. What does follow is that it is not deserved punishment if he does not. If the imponent made a mistake in judging that the accused did satisfy the condition, then there was a miscarriage of justice. Nevertheless, the imponent meted out punishment and the recipient suffered it.

It also follows that if the accused is found not guilty, that is, if the judge’s or jury’s verdict is ‘Not guilty’, then an imposition of deprivation after that cannot be punishment. It is not even a miscarriage of justice, but naked, blatant, shameless, mistreatment or victimization or something of the sort.

But what about the imponent’s state of mind? Can the imposition be punishment if the imponent does not believe that the recipient is guilty (or believes that he is not) but pretends to believe it? I think it is punishment if the other conditions are satisfied. But it is punishment so seriously flawed that it amounts to victimization: victimization by deliberately wrongful punishment, though its being victimization is now concealed from the public. Thus, K.B. Armstrong in a much-quoted article which was one of the first in the Retributivist revival, says that this is a mistake because it would include cases such as those of an innocent man who had been found guilty by a court that had meticulously observed formal
court procedures, been sentenced to death and been executed, hence such impositions were cases not of punishment but of something else, perhaps victimization. Similarly, John Kleinig\textsuperscript{24} says “deliberately to impose unpleasant treatment on a person knowing full well that he has not committed any offense (even though he may have been ‘alleged’ to have committed an offense or may even have been formally ‘found guilty’ of an offense) is not to punish him, but to victimize, take advantage of, bully, persecute, tyrannize, or oppress him (and perhaps others) for various ends, and to cover this up by making it look as though he is being punished.” It seems that both of them must think that nothing can be both punishment and victimization, for otherwise the fact that these are cases of victimization would not show that they are not cases of punishment. But this is a mistake.

It may perhaps be thought that punishment and victimization must be incompatible, for this reason: if it is known that the imposition is victimization, then it cannot be punishment, hence no one who knows or believes that it is victimization can consistently believe that it is punishment, hence those who are ignorant of its being victimization merely believe, wrongly, that it is punishment until they discover that it is victimization. But the analogous case of a miscarriage of justice should make us suspicious of this argument. The fact that someone discovers that an imposition is a miscarriage of justice does not show that it was or is not punishment. Plainly the crucial question is who can know what and when.

The imposition may be punishment even if at the time the whole moral community (except the court) knows that it is a miscarriage of justice, but not if they know it is victimization. And it can be punishment in both cases if they find out only later: then it was and of course always will be a case of a miscarriage of justice or victimization as well as a case of (flawed) punishment.

That this can be so is explained by what should by now have become obvious, namely, that what makes the imposition punishment is that it is a “branding” of the recipient as one who satisfies the conditions of punishment (commission of an offense, crime, or wrong), on account of which he deserves the community’s moral condemnation and the deprivation of certain rights. This “stigma” can, of course, be put on wrongly, whether by mistake (miscarriage of justice) or deliberately (victimization). But this is beside the point: it is punishment if he has been branded. If everybody knows at the time that he is innocent and that the court knows this too, then, there is only the unsuccessful attempt at “branding” him. If the community discovers the fraud only later, then the attempt was successful. The unmasking of the fraud does not undo the punishment, it merely reveals it to have been wrong. If he is “branded” by a miscarriage of justice, then he has been punished even though everybody (except, of course, the court) knows that he was “branded” by mistake.

The “branding” is thus a rather complex one. Success is not simply a matter of convincing the moral community that the recipient really was guilty, for he can get branded even if they know he is innocent (the case of miscarriage of justice). Nor is it simply the uttering in the appropriate context, of the magic formula “we find you guilty” for he does not get “branded” when the whole community knows he is innocent (the case of victimization or travesty of justice). I think the pronunciation of the magic formula in the appropriate context normally suffices to “brand” the recipient but it can fail to do so, namely, in the abnormal case when the whole community knows (or for good reason believes) that the court proceedings have been a travesty of justice, that the formula “we find you guilty” was fraudulently applied.

There is, however, a further question, namely, whether what the imponent of
THE STRENGTHS AND LIMITS OF THE RETRIBUTIVE THEORY OF PUNISHMENT

the deprivation is doing to the recipient amounts to punishing him if the imponent himself knows or believes the recipient innocent. It is natural to think that if the imposition amounts to punishment, then it must amount to punishing. But this need not be so. One of the reasons why this is overlooked is a confusion between two things. The first is that no one can suffer punishment unless his deprivation is deliberately imposed by someone (as I said above, under (ii)), hence no one can suffer punishment unless he is punished by someone, i.e., deliberately has imposed on him by that person a deprivation amounting to punishment. The second is that under certain conditions, someone's deliberately imposing a deprivation on someone amounts to his punishing him. The two may be logically unconnected. It may be the case that even though Jones is being punished by Smith, Smith is not ipso facto punishing Jones, just as (perhaps) Jones may be lying to Smith (if in order to deceive him he tells him something he thinks untrue) without Smith's being lied to by Jones (if what Jones tells him is in fact true). Whether or not the two are logically connected, depends on whether a person (logically) can punish another when he believes him not guilty. Consider the case of Jones who believes that his favorite daughter has broken the window, but then inflicts (what he had said would be, or now says is) the penalty for breaking windows on his stepson whom he dislikes but, mistakenly, thinks innocent of the misdeed. We might then want to say that the stepson suffers deserved punishment at the hands of his stepfather (because Jones is deliberately imposing the deprivation on the pretense that his stepson has broken the window) but might nevertheless insist that Jones has not punished him for breaking the window (or anything else) since he does not believe him guilty. We might want to insist that when Jones says “I am punishing you for breaking the window” this is false as well as being a lie, but also that if he had said “You are being punished for breaking the window”, it would have been true, though even then Jones’ remark would be dishonest or even lying.

Finally, it should be noted that in the case of legal punishment, where there is a division of labor between prosecutor, jury, judge, and actual administrator of the deprivation (executioner, prison warden, etc.), there is no longer a single punisher and so the simple question of what the punisher must believe if his imposition is to count as his punishing the recipient is no longer applicable. We can no longer find a straightforward answer to the question, ‘Who punished him?’ The punishing is a cooperative enterprise involving the playing of many parts, any of which may be performed faultily without necessarily preventing the recipient from getting punished, i.e., getting the stigma of guilt attached to him.

For these reasons I am more confident of the truth of (i): that the deliberate imposition of a deprivation when the recipient has been properly found guilty normally amounts to punishment of him, than I am of the truth of (ii): that no imposition can be punishing someone unless the imponent believes the recipient guilty or (iii): that (ii) is incompatible with (i).

b) The second, substantive interpretation of what punishment must be for, the one advanced by the Retributive Theory (R.T.), claims it to be something capable of providing a justification for what is prima facie wrong, namely, imposing a deprivation on someone. This explains why, despite the fact that punishment consists in a deprivation which is prima facie wrong, the fact that this deprivation is for that thing, punishment is prima facie right. “If we see a mother beating her child, it is usually sufficient for her to answer the question ‘Why are you doing that?’ (a question asked to ascertain whether what she is doing is right or wrong) with ‘I’m punishing him’.”25 Teleologists, such as deterrestists, incapacitationists, or re-
habilitationists, by contrast take punishment to be justified if and only if it takes a form in which it does in fact adequately promote their preferred end. But supporters of R.T. can muster a number of telling arguments against all such teleological theories of the proper function of punishment. In the first place, if what justified the imposition we call punishment were that it was an imposition by which, as in compulsory vaccination, the recipient or as in quarantine, the people around him, can alone or best be protected from harm, then whether and to what extent a given punishment is justified would be discovered by what turns out to be its consequences: if execution for petty thieving is more effective as a deterrent or incapacitant, then it would necessarily be more justified than less effective punitive measures. Secondly, it would then be sensible to look for more effective alternatives of achieving these ends than punitive measures, say, paying them large sums of money, so that they would not be tempted to commit their crimes. But it simply does not make sense to look for non-deprivatory, non-punitive forms of punishment. Unlike quarantine, punishment is not an evil that is necessary only as long as we do not know better ways of achieving certain ends, such as deterrence, etc.

Supporters of R.T. infer from this, not unreasonably, that the justification for punishment, i.e., for imposing deprivations as punishment, must lie in the very thing that punishment is for, not in something to be achieved by the deprivation. For only then can we know whom to punish and how much, and whether the punishment is justified, quite irrespective of how it turns out.

The question then arises what sort of thing it is that punishment is for. The three main candidates are offenses, crimes, wrongs. Under the influence of legal positivists, such as John Austin, who want to sever law from morality, even some retributivists have said that punishment is for offenses, breaches of social rules. But this is not a plausible view because the fact that someone has committed an offense, i.e., a breach of a social rule, is not itself a justification for imposing a deprivation. What makes it plausible to say that it is, is the underlying assumption that societies prohibit conduct which, for adequate prior reasons, they ought not to engage in.

The second candidate, crimes (or even misdemeanors), meets this point. However, it does so only if it is not defined, the way it often has been under the influence of Legal Positivism, as whatever is forbidden under threat of the criminal sanction. For when so defined, this candidate, crimes, is open to the same objection as the previous one, offenses. What makes crimes more plausible candidates is that by ‘a crime’ we ordinarily mean something that is wrong as well as prohibited under threat of the criminal sanction. To call someone a criminal is to condemn him morally, which implies that one thinks his behavior morally wrong. Of course, typically, conduct is prohibited because it is believed wrong and when it is prohibited, it normally also is generally believed wrong. But this is not always so. Many now believe that certain types of conduct are wrongly held to be crimes, are not really crimes, and therefore should be “decriminalized”, e.g., the so-called victimless crimes.

But even when interpreted in this morally loaded way, the candidate remains open to a telling objection. Not all punishment is for crimes: parents, teachers, bishops punish those under their authority but not for crimes. R.T. is thus naturally driven to the third candidate, moral wrongs. And in this, the commission of a moral wrong, R.T. finds a genuine justification for the deliberate imposition of a deprivation on the wrong-doer. If punishment is such an imposition on someone for a moral wrong he has committed, then it is justified because that is what is due to him, his desert, “because the vicious deserve to suffer".
THE STRENGTHS AND LIMITS OF THE RETRIBUTIVE THEORY OF PUNISHMENT

Thus, on R.T., justifiable, because just punishment is some wrong-doer getting the deprivation he deserves.

Now if this is a correct account of the nature of punishment, then it does seem plausible to say that punishment is just and therefore justified if and only if it satisfies the requirements of R.T. For if punishment necessarily involves the imposition of deprivations, then two matters require justification: the imposition of any deprivation at all, and the imposition of deprivations of a certain magnitude on a particular person. R.T. provides the first justification by allowing punishment only in cases of moral wrong-doing. It then relies on the plausible, though unexamined intuition, that membership in good standing in a moral community is rightly made conditional on refraining from wrong-doing, and that the deprivation imposed in punishment is the price exacted from the culprit for receiving him back into the community. This is the basis of the deep conviction many people have about the propriety of demanding that criminals be made to expiate or atone for their crimes by suffering, that only punishment can “wipe out” the crime, and so on. R.T. is also capable of explaining and justifying our current moral conviction that punishing, although it consists in doing what is prima facie wrong, namely, imposing a deprivation on someone, nevertheless, because of what it is imposed for is prima facie right.

R.T. provides the second required justification by insisting that punishment is (prima facie) justified if and only if it is just, that is, accords with the four principles of retribution.

R.T. thus rests on an implicit contractarian conception of the moral order. We can bring this out as follows. On R.T., punishment is the activity of bringing wrong-doers to justice, that is, meting out to them what they deserve, namely, deprivation. They deserve this because all members of a moral community are required, as the price of the advantages conferred by membership, to refrain from doing wrong. Those who refuse to pay this price are excluded from membership in good standing, for a period depending on the gravity of their moral wrong-doing, in the worst cases, permanently.

II

I want now to reverse my argument and weaken the hold of R.T. by showing that its central claims rest on an assumption seldom (if ever) discussed which, when brought into the open, is far from plainly true and indeed is no longer even particularly plausible under the conditions of contemporary industrial societies.

As we have seen, the central contents of R.T. amount to the following:

1. Every community has a right or duty to exact, as the price of continued membership in good standing, a certain amount of suffering or deprivation from “culprits”—I use this term as neutral between offender, criminal, or wrong-doer.

2. The severity of this deprivation ought to correspond to the gravity of the “misdeed” (offense, crime, wrong), measured in terms of the harm done and the degree of culpability of the culprit.

It is plain, however, that these two claims are plausible only against the background of a further tacit assumption:

3. The only justifiable response to (misdeeds) is punishment, i.e., bringing the culprit to justice, i.e., giving him his deserts, i.e., branding him as a culprit, expressing the community’s moral condemnation, and depriving him of his good standing in the moral community for an adequate period of time.
As long as crimes (or other misdeeds) are linked to punishment as their only appropriate sequel, the only appropriate question to ask will therefore seem to be, ‘What justifies punishment?’ But when answering that question, we are led, almost inescapably, by the very meaning of ‘punishment’, to embrace R.T., rather than any of its supposed rivals mentioned on p. 1. For since the proper meaning of ‘punishment’ is, as we have seen, the branding of someone as a culprit, his deprivation and his moral condemnation for his misdeed, its proper primary function must be to bring the criminal to justice, i.e., to give him his desert, and so the justification of punishment in each and every case must be that it is a just deprivation, one he deserves, that is, one in accordance with the four principles of retributive justice. The other possible functions of a social response to crime, deterrence, rehabilitation, incapacitation may therefore be incorporated in a penal policy only as secondary functions, that is, to the extent that penal strategies designed to carry out these policies do not come into conflict with the proper primary function of punishment, namely, bringing the criminal to justice. The concentration on punishment as the only possible social response to crime keeps before our eyes primarily the fate of the person to be brought to justice. Therefore the question uppermost in our minds when we ask whether the social response to a crime was justified is whether it was just punishment, that is, whether it was a just branding, condemnation and deprivation of the person to whom the punishment is administered, i.e., whether it was just to him, and (perhaps) to his victim. What is not before our eyes is whether the punishment which brings justice to these persons also is just to others involved, above all his and others’ potential victims.

During periods when the crime rate is low, and so any representative (randomly chosen) citizen has a rather low expectation of becoming a victim of crime, peoples’ concern with crime reduction will tend not to be great. They may then be satisfied with social responses to crime which do justice to the actual victim and the criminal. But in a period of rising crime, the representative citizen will become more concerned to have social measures designed to reduce crime and increase his security. (People living in safe areas may, of course, continue to remain unconcerned about potential victims.) If this aim is pursued within the confines of the principles of R. T., the currently acceptable forms of punishment (e.g., imprisonment or fines, but not flogging, hard labor, bread and water, mutilation, and so on) may become very costly, perhaps more costly than the community is willing to pay for. In such a case, the total deprivations “ideally” to be meted out are reduced by lowered “clearance rates”, plea bargaining, suspended sentences, paroling, and the like. As a consequence, a decreasing proportion of criminals are brought to justice and the crime rate continues to rise. Developments such as these suggest that bringing the criminal to justice is and perhaps ought to be our primary concern only as long as other concerns, such as the cost and the criminogenic effect of doing so remain acceptable. Despite all this, it would appear that R.T. continues to have a strong grip on many people. The demand to have criminals brought to justice persists despite the unwillingness to pay the costs. Philosophers, by casting their opposition to R.T. in the form of “rival theories (justification) of punishment” have contributed to the continued fixation on retributive justice.

It seems to me therefore that what philosophers should ask is not the traditional question, ‘What is the justification of punishment?’, which as we have seen, very plausibly leads to R.T. For this answer is then wrongly taken to imply (by way of the implicit assumption that the only proper social response to crime is punishment) that the sole primary function of any social response to crime must be bringing the criminal to (retributive) justice. What we should ask is, rather, “What is (are) the
THE STRENGTHS AND LIMITS OF THE
RETRIBUTIVE THEORY OF PUNISHMENT

proper function(s) of social measures in response to crime? When the question is
put in this way, it does not (improperly) close the possibility that the ultimate
primary function of a social response to crime is crime control or crime reduction.
Of course, this large topic cannot be treated on this occasion.

One thing, though, which can and ought to be examined here is the central R.T.
claim that the “culprit” deserves to suffer, i.e., that it would be morally fitting or
appropriate if he suffered, and that this claim is based on a deep-seated and widely-
held intuition. I want to bring to light what is sound and what unsound in this
claim. I want also to offer a few speculations about why so many people find even
the unsound part of this claim plausible.

Let us begin with an example given by Kleinig. Thirty years after World War
II, a Nazi who has fled and managed to carve out for himself an idyllic existence,
is unmasked and brought to trial. Kleinig claims that he deserves to suffer even
though no good purpose is served by the suffering. According to Kleinig—and I
agree—that someone deserves to suffer implies that if he suffers he is not justified
in complaining: no wrong has been done to him, but it may still be wrong for
certain persons to inflict the undeserved suffering on him. Kleinig thinks it possible
that it “would be improper for a state to inflict it”, but he does not appear to
think it possible that it might not be right for anyone to inflict it, not just for a
state, and not just because no one has the necessary authority to do so. For it may
be that, even though criminals deserve to suffer, under certain conditions the infliction
of suffering on those who deserve it is socially undesirable, because of the con-
sequences it has for others.

To return to our context: when the crime rate is low, the retributivist need not
consider the criminogenic dimension of the institution by which the deprivation is
administered to criminals. But if during a period of rising crime rates there is evidence
that these retributive institutions have significant criminogenic effects, particularly when the community spends little money on their administration, then in order for the continuation of the practice to be justified the comparatively harm-
less proposition that the criminal deserves to suffer must be strengthened into some-
thing like ‘let (retributive) justice be done even if the heavens fall’. When people
adhere to such “ultimate” principles based on their intuitions, in the face of evidence
that adhering to them contributes to disasters, it becomes justifiable to ex-
pose the psychological springs which might make such principles so irresistibly
attractive to them. In the remainder of the paper I want to spell out two hunches
about these psychological springs. My first hunch is that R.T. is most attractive to
people of a certain quasi-Hobbesian psychological type and that retribution would
perhaps be justifiable also as the primary function of any social measure for coping
with crime (not just as the primary function of punishment) if all or most criminals
were of that psychological type. My second hunch is that the claim that the crim-in-
al deserves to suffer, in the strong sense, which implies that he ought to be made
to suffer, can be derived by a quasi-Hobbesian line of argument from certain quasi-
Hobbesian psychological assumptions about human nature.

My first hunch is about the psychology of those who find R.T. attractive. To
formulate my hunch, I need to introduce a classification of possible attitudes to-
wards crime and the criminal law. I first distinguish three main groups:

1. The unconditional conformists, i.e., those who conform to the law whatever
society does about law-breakers. Society need not do anything about this
group (if indeed it has any members) since they will never break the law anyway.
2. The unconditional non-conformists, i.e., those who will break the law
whatever society does about crime as long as they are physically capable of doing so. As I have already mentioned, it is now widely believed that the bulk of crimes against the person and property is committed by members of this group during their active career of crime. About them society can do nothing to reduce their criminal output except to incapacitate them, either by keeping them in prison, executing them, or mutilating them, e.g., cutting off their hands (as I believe is still done in certain Arab states) or castrating them, as is done in some Western countries.

(3) The conditional conformists, i.e., those who will conform when society adopts certain measures against crime, but will not conform when it adopts others.

(3a) The calculating offenders, i.e., those who break the law whenever doing so promises them a net gain, whether pecuniary or psychic. This group of calculating rational egoistic criminals (organized crime, white collar crime) will not commit crimes if the cost of committing crimes is raised so as to make criminality unprofitable. This can be done by raising the certainty and severity of punishment. This group of criminals provides the strongest support for the general deterrence theory, i.e., the theory which holds that the intermediate primary function of any crime reduction measure is to deter potential criminals.

(3b) The desperate offenders, i.e., those who wish to conform if conformity is compatible with the solution of their major daily problems (e.g., getting food, drugs, and whatever else they need to keep going), but who cannot solve these problems except by breaking the law. This group of criminals provides the strongest support for the rehabilitation theory, i.e., the theory which says that the intermediate primary function of any crime reduction measure is to reform or rehabilitate those who have committed crimes.

(3c) The resentful offenders, i.e., those who wish to conform, but are determined not to conform if others don't, and who believe that others will conform if criminals are punished, i.e., taught the lesson that crime does not pay, or are given a taste of their own medicine.

It seems clear that theorists who think of the typical criminal as belonging to (3a) will be attracted to the deterrence theory, those who think of him as belonging to (3b) to the rehabilitation theory, and those who think of him as belonging to (3c) to the retributive theory. What's more, people tend to think of their own general motivational patterns as characteristic of human nature. They will therefore tend to think of criminals as being like themselves (though they will not necessarily put it to themselves in these words). Hence they will tend to favor the theories best suited to deal with criminals of the psychological type to which they themselves belong.

I conclude this first point with some speculations about what would be a rational response to crime. If a rising crime rate makes it unwise to ignore the question of how to keep crime in check, and if at the same time it is impossible or unjust to deal differentially with psychologically different criminals, that is, if the society has to choose a single type of response to crime, then the rational way to choose would seem to be to assess the relative size of these groups, and pick the method which has the best overall effect on the crime rate.

One of the most troublesome questions here will be the following. One may for instance want to adopt a mixed method, say, combining a policy of deterrence with a policy of negative retributivism. The problem, then, is this. Suppose insisting on depriving the criminal would interfere with his rehabilitation or that setting an upper bound to the severity of punishment according to the gravity of the crime would interfere with the deterrent or incapacitative effect, would it be morally permissible to set aside this requirement? H.L.A. Hart would think the former but
not the latter, while Bradley, under the influence of Darwinism, thought the latter but not the former. The matter is complicated by the probability that the social order is unjust and that some criminals are driven into crime by such injustices. At the same time, one's concern for the protection of the rights of criminals should not blind one to the force of the argument that a negative retributivist strategy of punishing (limiting the severity of the penalty by the gravity of the crime) which has very poor results in terms of deterrence is an injustice to potential innocent victims including the criminal himself on future occasions, when he may be a victim.

My second hunch concerns the reason why the claim that criminals deserve to suffer seems so obvious to so many. My hunch is that this claim in its strong sense, meaning that criminals ought to be made to suffer, can be made plausible by a type of Hobbesian argument if we assume that limited good will is a feature of all or most men. The steps of such an argument might be set out as follows:12

(a) Everybody has more to gain than to lose from being a member of a society, with clearly spelled out rules of conduct.
(b) It is in everyone's interest if the social order prohibits only conduct which inflicts significant harm on others, for only in that case does the society provide the maximal desirable security and the maximal desirable freedom.
(c) Everybody has adequate reason to want to be a member in good rather than a member in bad standing, i.e., with all of a member's rights fully protected rather than some suspended, since he is necessarily better off in the former than in the latter case.
(d) But everyone has adequate reason to want there to be adequate sanctions in support of the requirements of the social order, for (i) anyone is better off if he is not the victim of a crime, yet (ii) anyone can on occasion benefit himself by committing a crime and may, therefore, (for all that others may be able to ascertain) be tempted to commit one, but (iii) the existence of adequate sanctions would disincline such potential criminals from committing a crime.
(e) Sanctions are adequate if and only if they correspond in severity to the gravity of the violation, for only such sanctions teach the criminal the required lesson of what it is to suffer the harm inflicted by his crime. This lesson consists in what is often called “giving the criminal a taste of his own medicine”. It is what underlies the lex talionis. Of course, it need not be the same medicine, “the same taste” is enough.
(f) Therefore everyone has adequate reason to want criminals made to suffer by the application of adequate sanctions.

If this line of reasoning is sound, then under the conditions envisaged criminals deserve and ought to be made to suffer by the application of social sanctions fashioned in accordance with the principles of retribution. And those who have already committed one have done so contrary to reason. Hence they cannot in reason deny that there is adequate reason for anyone including themselves to want there to be such adequate sanctions, or that criminals deserve to suffer.

It seems to me than an argument such as this one confronts the retributivist with a dilemma. If the argument is sound or if it can be made sound by some modification, then we do not need a basic retributivist intuition, but can make the central claim of retributivism hinge on the deterrentist premiss (above (d,iii) and (e)) : that sanctions of a certain magnitude do in fact deter. Or else, if the argument is not sound, if, for instance, such sanctions do not in fact deter or disincline the potential criminals, if the central retributivist claim is based only on intuition, then the central retributivist owes us an answer to the question why, if non-retributive re-
sponses to crime (such as deterrence and incapacitation) are more effective ways of bringing down the crime rate, this is morally irrelevant; why the claims of the potential innocent victims need not concern the moralist.

FOOTNOTES


2A function (that someone or something can have) is an activity which someone or something is supposed to perform. We can distinguish between resultant and assigned functions. If something has a resultant function (e.g., the heart’s, to pump the blood) then normally it must be performing it. This is not true of assigned functions (e.g., the secretary’s, to type letters). Resultant functions, therefore, may be discoverable by observation, whereas assigned functions need not be. Of a thing’s assigned function we can always ask whether that is its proper function, the function that should be assigned to it, the function it ought to perform. If the CIA is assigned the function of spying on persons suspected of opposing government policy, this may not be (one of) its proper function(s).

A thing may have more than one function. We can lexically order them, and then speak of them as primary, secondary, --n-ary functions. We can also distinguish between a thing’s ultimate and its intermediate function(s). If crime reduction is the ultimate function of a certain social measure, then deterrence may be its sole primary, rehabilitation its sole secondary intermediate function, or they may both be primary, and so on.


FOOTNOTES


7See e.g., Shinnar & Shinnar, op. cit., in Note 6. The authors estimate (a) that more than 80% of solved crimes are committed by a comparatively small fraction of criminals, a group of hardened recidivists, and that a comparable percentage of unsolved crimes is committed by the same recidivists and (b) that, given a mugger’s present chances of going to prison, sending every convicted mugger and robber to prison for five years would reduce this type of violent crime by a factor of five (p. 605). And if we raised the present chance of 1.2% of a mugger’s getting convicted, to 20%, the value now (1975) prevailing in England, then a net prison stay of three years would do the job. (pp. 592-599).

8See e.g., Ehrlich, “The Deterrent Effect on Criminal Law Enforcement”, p.261.


10For a discussion of the classical retributive theory, see e.g., Edmund Pincoffs, The Rationale of Legal Punishment. Humanities Press, 1966. For a survey of contemporary views, see e.g., Nigel Walker, Sentencing in a Rational Society, Basic Books, 1969/71, Ch. 1.


12Cf., e.g., von Hirsch, op. cit.


14John Kleinig, in a book Punishment and Desert. Martinus Nijhoff, 1973, from which I have learnt and am borrowing much, argues that many of the disputes about the nature of punishment are due to the attempt to define ‘punishment’ by stating the necessary and sufficient conditions of something being punishment. He argues that it is necessary (or better) to lower our demands on definition and be satisfied with stating what is “typically or characteristically involved in punishment” (loc. cit., p. 37). He then goes on to explain that the reason why some disagree on whether it makes sense to speak of ‘punishing the innocent’ is, at least in part, that people have tried to define ‘punishment’ by listing “a set of necessary and suffi-
cient conditions for its use'', including 'guilt', 'offender', etc. If they were satisfied with saying ''that guilt is typically or characteristically involved when speaking of punishment, then there would be no problem. But this does not seem to me a solution of the problems of defining 'punishment' or even of dealing with the question of whether it makes sense to speak of 'punishing the innocent'. For (i) Kleinig himself distinguishes between qualified, e.g., unjust, wrongful, ineffective, etc., and unqualified punishment, and then argues that guilt is a necessary condition of something being punishment when it is ''unqualified'' but not when it is ''qualified'' punishment. But from this it would seem to follow that 'guilt' is not a necessary condition of something being punishment since it is not such a condition of something being punishment whether qualified or not. Hence it makes perfectly good sense to speak of punishing the innocent, though not therefore (necessarily) of justly punishing the innocent. But more importantly (ii), defining 'punishment' by merely saying what is typically or characteristically true of it does not solve the problem of what to say where a person is subjected to deprivation when he is known to be innocent. Is this a case of qualified punishment or not of punishment at all? Kleinig does not say, except that ''the situation is of course different when it is believed that the person is innocent''. (loc. cit., p. 37). It seems to me that on his approach to definition there can be no answer to this question. But then he seems to me wrong, or at least overstating his case when he says (loc. cit., p. 16) that with his approach to definition we are not precluded ''from doing all the things we hope for from a definition ... we will be able to give sufficient clarity to the term without being reduced to vagueness or unwarranted stipulation''. I see no reason for abandoning the search for necessary and sufficient conditions.

15See John Kleinig, p. 18.

16See John Kleinig, op. cit., p. 18.


20Cf., e.g., Edmund Pincoffs, op. cit., Ch. 1, or John Kleinig, op. cit., pp. 25-34.

21See e.g., John Kleinig, op. cit., p. 34.

22I have made this point before, in ''Is Punishment Retributive?'' Analysis, vol. 16, no. 2, Dec., 1955, also reprinted in H.B. Acton, ed., The Philosophy of Punishment, St. Martin's Press, 1969, and Gertrude Ezorsky, ed., Punishment, State University of New York Press, 1972, but the claim has often been disputed.

23''The Retributivist Hits Back'', Mind, October, 1961, reprinted in Acton, op. cit., p. 34.
FOOTNOTES

24 See John Kleinig, op. cit., pp. 66f.

25 Ibid., p. 42.


28 Cf., e.g., Norval Morris & Gordon Hawkins, op. cit., Ch. 1.

29 Cf., e.g., John Kleinig, op. cit., p. 66.


31 The point here is that the meaning of ‘punishment’ comprises its proper primary function, and that since it also implies that performing that function is prima facie morally right, that primary function is also morally acceptable, as e.g., the proper function of a hit-man is not. This is easily overlooked since, like ‘hammer’ and unlike ‘distributor’, ‘punishment’ does not wear its proper function on its lin­guistic sleeve.

32 ‘Clearance rate’ is the name for the proportion of those among the reported crimes which have been “solved”, that is, where a suspect has been arrested. They are a measure of the probability of arrest, not of conviction or imprisonment.

33 I here ignore the question, which an adequate treatment would have to in­vestigate, to what extent the increasing crime rate is due to a growing perception, by sections of the population, of the social and legal order as unjust, and to what extent that perception is sound. If this is a significant aspect of the situation, then that raises very great additional difficulties for the application of R.T. to actual cases of law-breaking.


35 Ibid.


37 This classification is based on a suggestion by Jackson Tobey, op. cit.


FOOTNOTES

40 The type I call “resentful offenders” are people of limited good will, not necessarily ruthless egoists. They may for instance be concerned primarily with the promotion of their own good, secondarily that of those they care about (though some of them may love a few as much as or more than themselves) while towards the bulk of mankind they feel more or less indifferent, but have nevertheless an attitude of limited (conditional) good will. That is to say, while not going out of their way to promote the good of these others, or prevent harm to them or even to refrain from doing what will affect them detrimentally, they are yet willing to cooperate with them, make and keep contracts and in other ways “play the game”. And they do this even where they could do better for themselves by breaking their contracts or violating the rules, but only on condition that others do likewise. For a discussion of the pattern of preferences involved in these psychological types, cf., e.g., A.K. Sen, “Choice, Orderings and Morality” in Practical Reason, ed., Stephan Koerner, Yale University Press, 1974, pp. 54-67, my critique of Sen, “Rationality and Morality”, and Sen’s “Rationality and Morality: A Reply”, both forthcoming in Erkenntnis.


42 I do not contend that everybody—or even anybody—who has the central retributivist intuition is clearly aware of these steps. For my hunch to be sound it is sufficient that when such reasoning is presented to such people, many of them should regard it as making their own thinking clearer to themselves and that they should find it cogent. For these reasons, it is not essential that the argument contain precisely these steps or even that the argument really be sound.