1980

Towards a Theory of Punishment

Jonathan Bennett
*Syracuse University*

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/vol11/iss1/3
Bennett: Towards a Theory of Punishment

JONATHAN BENNETT
Professor of Philosophy
Syracuse University
Towards a Theory of Punishment

by

Jonathan Bennett

1. The accountability problem
2. Strawson’s solution
3. Justifying reactive attitudes
4. A problem about penal theory and practice
5. A case for a purely teleological penal system
6. Four counter-attacks met
7. Reactive attitudes and justice
8. Severity of penalties
9. Three constraints
10. Sympathy
11. Rights
12. Coda

1. The accountability problem

The fundamental problem about blame and responsibility has, in my opinion, been solved by P.F. Strawson’s paper ‘Freedom and Resentment’. I have discussed this at length elsewhere, and shall here merely offer a quick sketch, before proceeding to use this work of Strawson’s to illuminate the concept of punishment.

The basic problem about accountability is to explain and justify our way of distinguishing actions for which the agent is accountable from those for which he isn’t — e.g. between causes of an action which we do and those which we don’t regard as ‘compelling’ the action to be performed or ‘preventing’ its non-performance. We can say where this line is drawn, but the question is: why draw it just there?

I know of only three kinds of answer to this question.

The radical libertarian answer rests on a tendentious view about where the line falls, namely that an agent is accountable, and is not compelled to act, only if his action was not fully caused by antecedent events. I shall dismiss this without argument; but only because I have not time for the argument which it deserves.

Secondly, there are theories like that of Schlick, which go as follows. When someone acts in a manner which we welcome (regret), we do or should be concerned to increase (decrease) the likelihood of his doing so again. One way of doing this is by moral pressures — equipping the agent with encouraging (deterrent) beliefs of the form ‘If I perform an A, the upshot will be ...,’ so as to motivate him to perform (not perform) A actions on future comparable occasions. Moral pressures range from raised eyebrows to dire punishments, and from encouraging smiles to lavish rewards. Within this context, Schlickian theories offer their rationale for the line between accountable and non-accountable, by describing it as the line between cases where moral pressures do and ones where they don’t have some chance of being effective.

The trouble with Schlickian theories is that in providing only for a teleological, remedial, encouraging or deterrent approach to human conduct, they omit an important part of the story — the part involving genuine praise and genuine blame.
Towards a Theory of Punishment

It is sometimes said that we ought to give up blame and everything which conceptually depends upon it (e.g. genuine punishment) in favour of the forward-looking technique of deterrence through threatening moral pressures. But nobody urges us to give up praise, gratitude, etc. in favour of the clean-cut forward-looking use of encouraging moral pressures! Anyway, I shall assume that we all want to improve on Schlickian theories if we can.

2. Strawson’s solution

Strawson distinguishes reactive attitudes from a kind of attitude which he calls objective and which could also be called teleological. If you do something which I regret, my attitude is objective and teleological in so far as I think about why you did it and how I can reduce the likelihood of your doing it again: this kind of attitude could be adopted towards the catching of a cold, say, or the leaking of a roof. But if I am indignant or resentful or angry, or if my feelings are hurt, my attitude is reactive: it is a kind of attitude which an adult will have only towards a personal object — either an actual person or else something like theologized weather or an anthropomorphized dog. Similarly, if you do something which I welcome, I may in an objective frame of mind bend my efforts to encouraging you to do it again, or I may respond reactively with sheer gratitude.

Those remarks aim only at giving the general drift. They do not purport to define ‘reactive’, that being something I don’t know how to do.

Strawson presents moral praise and blame and their kin as special kinds of reactive attitudes. In a nutshell: blame is to be understood as a development of the more primitive response of resentment, and praise as an outgrowth of gratitude. To get clear about praise and blame and their near relatives, Strawson says, we must get clear about reactive attitudes generally. For a start, we can now pinpoint the defect in Schlickian theories, which is that they provide only for objective-teleological responses to actions and not for reactive ones.

Using the concept of a reactive attitude, Strawson offers a third rationale for the line between accountable and non-accountable. Where the radical libertarian rationale rests on bad metaphysics, and the Schlickian one is shallow, Strawson’s is at once coherent and adequate to the facts of the human condition.

It has two premises. (1) Reactive attitudes are essentially responses to the attitudes of other people. (2) Reactive attitudes properly belong to a context of involved, reciprocating, adult, interpersonal relationships. Because of (1), it is inappropriate to respond reactively to a good action which wasn’t really well-meant, or to a bad one which wasn’t really ill-meant — e.g. if the agent ‘didn’t realize ...’, or ‘couldn’t help ...’, or was acting under duress, etc. And because of (2), a reactive attitude is inappropriate if the agent is, at least temporarily, incapable of entering into fully adult interpersonal relationships, e.g. because he is one year old or hopelessly schizophrenic. These two lines of thought combine to generate a fresh account of the point and purpose of the accountable/non-accountable distinction: the realm of accountability is that in which reactive attitudes would not be downright inappropriate for one or other of the two reasons I have outlined. This puts the distinction just about where Schlickian theories put it; but it explains better than they do why it is placed just there.

3. Justifying reactive attitudes

Strawson says that reactive attitudes are so deep in our natures that there
Jonathan Bennett
couldn't be a serious question of giving them up. Still, if we could give them up, ought we to? Strawson says No: in so far as one can imagine a world where all our inter-personal responses were strictly objective and teleological, the prospect is horrifyingly unattractive.

Notice that the question 'Ought we to give up reactive attitudes?' is being treated as a practical one — more like 'Ought we to give up music?' than like 'Ought we to give up quantum theory?'. That is because reactive attitudes are matters of feeling rather than of judgment. If you do something which I regret, there is the factual question of whether a reactive attitude would be simply inappropriate, e.g. because you are insane or because your hand slipped; but if I know that it would not be inappropriate, my actually being indignant, or instead my coolly viewing the matter in an objective and remedial frame of mind, depends upon personal choice, the mood of the moment, some aspect of my personal style. That is all it can depend upon: it cannot arise from some further perception of the objective indignation-worthiness of your action.

So when we contemplate the picture offered us by Schlickian theories, and observe that there is a gap in them, what we are noticing is the absence of reactive attitudes. Do not think that the gap is to be filled by a further kind of fact about the moral quality of actions (their objective blameworthiness or praiseworthiness) and that reactive attitudes involve judgments about or perceptions of those facts. Rather, Strawson says, 'it is just these attitudes themselves which fill the gap.'

4. **A problem about penal theory and practice**

Granted that we should not relinquish reactive attitudes altogether, there are contexts in which they are so counter-productive that they ought to be excluded as a matter of policy: Strawson instances the attitudes of a psycho-analyst towards his patient. Perhaps another such context is the application of the penal system. Although it would be dreadful if all human inter-reactions were of the teleological sorts provided for by Schlickian theories, it might be a positively good thing if every question about the disposition of offenders against the law were settled by teleological considerations. I shall state that case for that view, as abstractly as possible. I shall say little about what the ends should be — whether they should include prevention, deterrence, reform, the forestalling of private revenge, etc.; and shall say nothing about how the welfare of the offenders, victims, etc. should be balanced against that of society, or about what it is for an individual or a society to fare well. My concern is only with the highly abstract thesis that the treatment of offenders should be controlled purely by thoughts of the desirability of upshots. The rival thesis says that the handling of offenders should be coloured by reactive attitudes — that it should involve not merely treatment/reform/deterrence/prevention and the like, but also genuine punishment, with this being thought of as connected with blame and indignation and reactive attitudes generally.

Of course the traditional rival to an ends-and-means approach to punishment is the view that it is all right to punish offenders because they deserve it. I don't think that that is wrong: I offer Strawson's approach not as a replacement for the notion of desert, but rather as a way of deepening and strengthening it, though no doubt also changing it a bit in the process. I shall say a little more about this shortly.

5. **A case for a purely teleological penal system**

On any theory, a penal system is bound to involve unpleasantness for the
Towards a Theory of Punishment

offender. Consider the effects of that upon his character. Even if we don't aim to reform him, presumably we don't contentedly acquiesce in the destruction or serious damaging of his character, making him insecure, self-hating, bitter, alienated, aimless and despairing. But any penal system which makes offenders suffer runs a permanent risk of damaging them, because of that very basic fact of human nature which Auden used in his short explanation of Hitler:

I and the public know
What all schoolchildren learn,
Those to whom evil is done
Do evil in return.  

This inescapable problem of minimizing damage to the offenders' characters, while not being self-defeatingly lenient, is too serious to be ignored; and it is too difficult to be soluble without the help of the most informed, careful and intelligent attention to ends-and-means that we can possibly muster — in short, without an unremittingly objective and teleological attitude.

Other damage which penal systems could do generates other arguments for objectivity and teleology in penal theory and practice. I offer that central one as a stand-in for the rest.

It might be replied that although of course penal systems should involve careful attention to means and ends, they should also express reactive attitudes. But where in the system are such attitudes to find expression? In the penal code itself, or in its application? Each alternative involves grave difficulties.

First, the penal code. General policies fixing penalties for kinds of offence might be shaped partly by our indignation over this or that sort of offence. But our indignations cannot make a difference unless they sometimes lead us to assign a more severe penalty than we would have assigned had we been guided purely by thoughts of ends and means: e.g. in fixing the penalty for armed robbery, we balance the need not to damage the offender's character against the needs of deterrence and physical prevention and the forestalling of private vengeance, etc., and then we tilt the balance towards greater severity because armed robbery makes our blood boil. This seems clearly to be morally intolerable. It is one thing to accept a degree of imprudence as the price of spontaneity in our daily lives; it is quite another to depart carefully from the course of prudence, wisdom, humane and careful firmness, etc. so that our cool policies may reflect the way we behave when we are neither cool nor politic.

Secondly, the application of the penal code. Reactive attitudes might enter the system by our leaving certain decisions open, to be settled on the spur of the moment according to the feelings of the judge, the jury, the prosecutor, the victim or his relatives, and so on. That would be a stirring reminder of that warm, erratic spontaneity which is so valuable in our personal lives; but it would greatly increase the risk of the penal system's doing gross damage, and — worse — it would violate the essence of law, which must be expressible in general principles which can be objectively applied. The suggested use of reactive attitudes would expose the law to those influences of passion and of personal involvement which civilized legal systems rightly strive to keep out.

6. **Four counter-attacks met**

It is sometimes said that a purely teleological approach to penology involves...
a demeaning or dehumanizing attitude towards offenders (the offender's alleged 'right of punishment' comes in here). But nobody would think that a psycho-analyst must dehumanize his patient when he treats him, with clinical objectivity, as a 'case', rather than responding to him in a fully personal and thus partly reactive manner. If indeed he subtracts from anyone's humanity, it is from his own. Nor does he imply that his patient is not a fit partner for a fully interacting relationship: it is just that there is a time and a place for everything, and psycho-analysis goes better — to put it mildly — if the analyst's disgusts and indignations are kept out of it. All this might apply equally to the sentencing judge's attitude to the convict. An opponent might say, threateningly, 'If to all offenders, then why not to all mankind?' — that threat is mentioned by Strawson, and though he doesn't endorse it he doesn't say either, as I now do, that there are powerful reasons for restrictions of attitude towards 'all offenders' which do not imply similar restrictions of attitude towards 'all mankind'.

It has been said that in treating offenders teleologically one exhibits a mean, impoverished, lowering view of what matters most — e.g. in putting 'law and order,' naively construed, before individual welfare. But one can pursue ends without pursuing those ends! A strictly teleological penal system could reflect a really generous view of human values and rights and potentialities.

My third point, which probably doesn't need to be made, is that in keeping reactive attitudes entirely out of the penal system, we need not ignore any relevant facts about reactive attitudes. If people feel so indignant about burglary, say, that if burglars are not punished severely there is a risk of vigilante action, that may be a reason to step up the penalties for burglary. If rape-victims have feelings of shame or guilt which constitute part of the harm they have suffered, that adds to the gravity of the offence of rape, which may have implications about how rapists should be dealt with. But thought about these matters may belong entirely to the realm of means and ends, the realm of the objective teleological attitude.

Fourthly: it may be said that a purely teleological system would not really be penal, would not involve punishment properly so-called. That may well be right, for punishment seems essentially to look back to the deed rather than forward to the future. But what's in a name? The defender of a purely teleological disposition of offenders can agree that what he is recommending is indeed a 'remedial' rather than a 'penal' system.

7. Reactive attitudes and justice

So much for difficulties which can be met; now for one which cannot. It is the famous old difficulty about fairness or justice, which is something we think penal systems should aim at, and which cannot be fully captured by a strictly teleological (consequentialist, utilitarian) approach. Even if in very extraordinary circumstances it could be right to 'punish' someone who was known to be innocent, that is still prima facie a dreadful thing to do — something which couldn't be justified merely because it led to a somewhat better state of affairs than could otherwise have been obtained. Our thinking about penal systems, then, retains a non-teleological element, even if we do everything we decently can to eliminate the idea of punishment as retribution: it is the element which says that the permissibility of making someone suffer depends partly upon whether he has done wrong.

A determined utilitarian might say that this element should be weeded out also. 'It would in fact seldom be expedient to "punish" an innocent person,' he might say, 'and it is our muddled apprehension of this fact which leads us to think
Towards a Theory of Punishment

that it would be wrong to "punish" such a person even if it were expedient to do so. This position, properly understood, is less disgraceful than it seems on the surface; but no one would accept it if they could see a valid alternative.

The alternative, of course, is the non-teleological thesis that it is prima facie wrong to inflict suffering on the innocent, even when it is expedient to do so. The problem is to give a clear, deep reason for this without committing oneself to some unacceptable theory. Traditional theories about how offenders deserve punishment have characteristically had two bad features. One is that they don't just permit us to punish offenders but require us to do so. A 'desert' theorist may assert that the punishment of the guilty is permissible but not mandatory; but I have never seen a principled basis for this position, as distinct from its being offered ad hoc as a pair of doctrines which are tailored to fit our moral intuitions without being made to fit into any underlying theory. When the notion of desert is rooted in Strawson's doctrine about reactive attitudes, however, we get as a deeply grounded theorem the result that punishment of the guilty is permissible but not mandatory; which is just what most of us want.

The second bad feature of traditional theories about desert is that they can't help stumbling over the notion of an offender who doesn't deserve punishment because in offending he was not free or responsible or accountable; and it is a symptom of this difficulty that such theories tend to be associated with that radical libertarianism which I mentioned earlier, saying that an offender is not accountable or deserving of punishment unless his offence was not fully caused by prior states of affairs. This is a muddle, I believe, but not an easy one to escape from. One of the great virtues of Strawson's doctrine of reactive attitudes is that it shows us how to get out of this tangled trouble about determinism and accountability.

Those are my two main reasons for thinking that the notion of deservingness stands in need of help which is lavishly provided by Strawson. I haven't shown that I am right on either count: I have merely indicated what my conclusion is on this matter, and hinted at how I would argue for it if there were time to do so. Since there isn't time, I now turn my back on the concept of desert, and proceed purely in terms of Strawson's concept of reactive attitudes.

The most basic way in which our willingness that a man should suffer is connected with our belief that he has offended is through their roles in adverse reactive attitudes. If I resent some attitude of yours towards me, my resentment must involve some measure of willingness that some unpleasantness should befall you. An adverse reactive attitude essentially involves some disposition to hit back or to be pleased if God or Nature does it for one. To divorce judging someone to be an offender from willingness that he should suffer, therefore, we should have to strip ourselves of all adverse reactive attitudes; and that is unthinkable or unacceptable. So the connection between 'He offended' and 'It is all right to make him suffer' is forged by an aspect of human nature which is ineliminable and undisgraceful; and that aspect of our natures is what is expressed in our judgment that guilt but not innocence justifies the infliction of suffering.

That is a rather peculiar justification of the judgment in question: it doesn't derive it from accepted moral principles, but rather shows it to be an upshot — not through muddle or error — of a valuable aspect of human nature. Without proving any theorem about guilt and suffering, it provides evidence that something would have gone wrong if we rejected, on principle, policies which connect guilt with suffering; and that does, after a fashion, justify such policies.

The foregoing attempt to base the concept of penal justice on reactive attitudes is Strawson's, I believe. He says that 'to speak in terms of social utility alone
Jonathan Bennett

is to leave out something vital in our conception of [penal] practices, and he indicates that this ‘something’ is justice. He does not explain in detail how reactive attitudes can make good this lacuna, and merely implies that the gap won’t be filled unless our ‘preparedness to acquiesce in [the] infliction of suffering on the offender...is all of a piece with’ — or is on a ‘continuum’ with — reactive attitudes generally. Those phrases, ‘all of a piece with’ and ‘on a continuum,’ vague as they are, are exactly right if the line of thought is the one I have been presenting.

8. Severity of penalties

Even if reactive attitudes underlie the view that someone’s being an offender is relevant to the permissibility of making him suffer, it might still be that every detailed question about the penalizing of offenders should be rigorously governed by teleological considerations. Reactive attitudes could underlie our willingness to have a penal code of any kind, without affecting what specific code we adopt. So the case I presented in § 5 has not yet been seriously harmed.

But there is a further problem to be faced. The thought that the innocent ought not to be made to suffer is linked with the thought that how much an offender should be made to suffer is limited by the mildness of his offence. It might be expedient to hang speeding drivers, for example, but we all find such a penalty morally excessive for such an offence. This is a further non-teleological element in our thinking which needs to be justified; explained, put in its place.

My concern is only with judgments of the form ‘P is too severe a penalty for offence O.’ As for ‘P is too lenient,’ I believe that the lower limits on penalties for kinds of offences should be set by considerations of utility alone. This is a common enough moral position, which is held for example by those who cannot swallow the orthodox Christian thesis that Man’s sin required the paying of a price. It is, furthermore, a view which is considerably reinforced by the Strawsonian position I have been developing for that position bases the notion of ‘all right to punish’ on that of ‘adverse reactive attitudes appropriate’ or ‘...not inappropriate.’ To provide a basis for ‘not all right not to punish’ it would have to make room for ‘adverse reactive attitude mandatory;’ but I submit that reactive attitudes are never mandatory, because in any individual case there is the legitimate option of retreating from reactive attitudes into a more detached objectivity.

Let us now consider judgments about undue severity of penalties. These are not mentioned in ‘Freedom and Resentment,’ and seem to be ignored throughout most of the philosophical literature on punishment. But they are worth investigation; and it will turn out that reactive attitudes are again involved, though this time they are not the whole story.

There is no special problem about our judgments that one offence is graver than another, or one penalty more severe than another. These rest on judgments about which situations are worse for their inhabitants than which others, and these lie near the heart of anyone’s value-system. So we have a fairly unproblematic notion of a more or less grave offence, and a more or less severe penalty.

What about the view that the less grave an offence is, the less severe should be the maximum penalty for it? That, I suggest, is underpinned by a fact about our everyday reactive attitudes. Even where resentment is not inappropriate, its intensity may be judged to be excessive: not necessarily harmful or inexpedient, but simply disproportionate to its object. And similarly with gratitude and the rest. If our penal thinking generally is ‘all of a piece with’ or on a continuum with our everyday reactive attitudes, our judgment that severity of penalty should be con-
Towards a Theory of Punishment

strained by gravity of offence could be on a continuum with our everyday sense that a given reactive attitude may be too intense for the instance of malignity, indifference, goodwill or whatever which aroused it. And to the extent that this aspect of our ordinary reactive attitudes is a necessary and undischgraceful one, this is again a justification — of a slightly peculiar sort — for the judgment that a penalty may be unduly severe for a given offence.

But that implies nothing about what penalties are excessive for what offences. And we shan’t get judgments about that from our ordinary reactive attitudes: they normally carry the thought of ‘unpleasantness for the offender’ in far too thin and abstract a manner to put content into a penal code. It seems, then, that our specific judgments about undue severity — e.g. that hanging is a morally excessive penalty for speeding on the road — must have more than just our propensity for reactive attitudes to support them.

Before saying in §10 what else I think is involved, I note three conditions which apparently must be satisfied by any acceptable theory about what underlies our specific judgments about the undue severity of sentences.

9. Three constraints

There is no non-moral relationship R such that our judgments about undue severity of penalties derive from a single principle of the form ‘Every penalty should have relation R to the corresponding offence.’ I thus reject the attempt to use anything like ‘The punishment should fit the crime’ as a basis for our undue-severity judgments, though it may be all right as a casual summing up of them. The most promising such attempt is the one which, assuming that there are no victimless offences, takes it that gravity of offence and severity of penalty both involve amount of harm or suffering or disutility — in one case to victims and in the other to the offender. The idea is that that common element might generate a co-ordinating relation R between gravity and severity. But what, specifically, could R be? It would have to be something like ‘not hurting the offender more than his offence risked hurting its victims,’ so that the governing principle of our undue-severity judgments would be that no penalty should in that way outrun the offence for which it is being inflicted. But it is not hard to see that in countless cases this principle would yield no answer (e.g. someone who burns down a National Forest), in others it would set the ceiling much too low (e.g. someone who steals a million dollars from a wealthy bank), and in yet others it would set it far too high (e.g. someone who tortures a child to death).

So (a) an acceptable theory about our undue-severity judgments will not derive them from any single principle of that form; it will represent the moral situation as more complex than that.

Our undue-severity judgments apparently do not derive from any underlying moral principles: for we apparently cannot argue much about them, as we can about moral judgments which have more general ones behind them. I may try to convince you that penalty P is excessive for offence O by arguing that it is not after all expedient, or by making you more aware of what P is actually like for someone undergoing it; but those moves bring me to the end of my resources, which is not the case in a moral disagreement which has more moral theory underlying it.

So (b) an acceptable theory will not trace our undue-severity judgments back to any substantial body of moral principles.

But I do not infer that they are a mere jumble, resting on nothing of a principled or systematic kind. Evidence that they are not is provided by something which
has been happening throughout the western world over the past four centuries. The maximum penalties for various kinds of offence have been fairly steadily decreasing in severity; and while that may reflect changing opinions about what is expedient (connected, for instance, with discoveries about how environment can harm character), it also reflects changes in what is found to be morally permissible. The relative steadiness of this change suggests that such judgments have a basis which saves them from being a mere jumble.

And (c) any acceptable theory about that basis must respect, and if possible explain, the historical change to which I have referred.

The above three constraints are all satisfied by the hypothesis I shall offer regarding the source of our undue-severity judgments.

10. **Sympathy**

Let us start with one aspect of the historical change which I have mentioned. Although we contemporaries sometimes disagree about the severity of penalties, none of us thinks that any conceivable offence could make it permissible for the offender to be hung, drawn and quartered, or to be burned at the stake. What produces this difference between ourselves and the people of the England of Elizabeth I?

Well, we in a sense cannot bear the thought that someone is at this moment being deliberately burned to death, whereas an Elizabethan could with equanimity bear the thought, and even the sight and sound and smell, of the burning. That is because we are less callous than the Elizabethans were: more prone than them to sympathy, fellow-feeling, pity, etc., when consciously aware of the suffering of others. There are reasons for that in turn, but I shall not pursue them.

So much for the judgment that \( P \) could never be justified by the magnitude of any offence. As for the judgment that \( P \) is morally excessive for offence \( O \) in particular: I suggest that all such judgments arise from the inter-action of (1) our disposition towards sympathy, pity and the like, and (2) our propensity for reactive attitudes. My sympathy makes me *prima facie* unwilling that someone should suffer \( P \); but my belief that he has committed \( O \) generates in me a level of resentment, indignation, etc. which dampens down my sympathy to a point where I am willing after all that \( P \) should be inflicted. The worse the offence, the greater the indignation, and so the greater the amount of sympathy that can be dampened down by it; but only within certain limits, for there are sympathy-levels which cannot be overcome by any reactive feelings, such as the pity I feel towards any actual or imagined sufferer of death at the stake. (That is not to deny that it might be overcome by something else, e.g. if the alternative to burning one person to death were, certainly, ten people’s being burned to death.)

The gradual lowering of the penalty-ceilings over the past four centuries, on this account, could be due to a gradual increase in sympathy, the lessening of callousness; and the latter change has indeed occurred and been manifested in other areas as well as penology — e.g. in views about slavery, about indigence, about school discipline, and so on. There may also have been a change in the propensity for adverse reactive attitudes, but we don’t have to postulate one: the change in moral judgments could be due to a change in one of the two interacting elements while the other held steady.

(In all of this I am agreeing with Strawson that sympathy, pity, etc. are not themselves reactive attitudes. One can pity a baby, a horse, a mentally ill person, perhaps a sparrow, in respect of which resentment, gratitude, anger, jealousy, etc.)
Towards a Theory of Punishment

would be glaringly inappropriate.)

I do not contend that those who devise a penal code set penalty-limits on the basis of the outcome of a struggle within them between spurts of indignation on the one hand and wellings of sympathy on the other. On the contrary, they may dispassionately judge that O is a very grave offence, and dispassionately conclude from that that P is the appropriate maximum penalty for it. But I contend that when they make these judgments what they are doing is 'all of a piece with' (on a continuum with), and best understood in terms of, what happens when sympathy is tempered by indignation.

On this account, our undue-severity judgments express a kind of limit on what we will tolerate in the way of suffering for an offender, these limits being direct upshots of natural human sympathies rather than theorems derived from broader moral principles. The absence of backing in moral theory explains why we can do so little to defend or argue about specific undue-severity judgments; while their rooting in fairly steady aspects of human nature saves them from being a mere chaos, and helps to explain how they can have undergone change over the centuries. That satisfies two of my constraints, and the other is also clearly satisfied — my account does not invoke any super-principle saying that the punishment should 'fit' the crime.

11. Rights

A certain rival theory about undue-severity judgments should be presented and refuted. It says that our undue-severity judgments are based upon judgments about rights. 'Everyone has certain rights, some of which are lost when one offends against the law, the amount of loss being proportional to the gravity of the offence. So someone who has committed offence O may be subjected to penalty P only if O is grave enough to have deprived the offender of his right not to undergo P.'

This fits the historical fact I have mentioned, for the western world has taken a steadily more generous view of human rights. Whether the 'rights' approach also satisfies the other two constraints is not worth discussing, however, for the approach fails anyway.

Some questions about the handling of offenders are discussable in the language of 'rights': does his crime deprive him of his right to physical freedom? his right to sexual companionship? his right to vote? his right to complain? But if 'rights' underlie all our judgments about undue severity of penalties, then innocent people must have, for instance, a right not to be gaol for a month and a right not to be gaol for a year, so that in a particular case an offender may be judged to have forfeited the former right but not the latter. And there are countless other examples: the suggested approach would require an endlessly complex and elaborate set of beliefs about negative rights — the right not to be treated thus-and-so — which everyone has until they start being peeled off him by his offences against the law. Our apparatus of moral thinking and reasoning does not include, even unconsciously or dispositionally, any such structure as that.

It might be objected that we need only to attribute to innocent people a single right, namely the right not to be made to suffer by the law; and then further judgments will be made about how much of that right is lost when one commits a certain offence. But then the alleged basis for our judgments about undue severity of penalties turns out to be a mere rewording of them. We are to base judgments of the form 'Penalty P is unduly severe for offence O' on ones of the form 'Someone who commits offence O does not lose enough of his right not to be punished to
make it legitimate to inflict P upon him.' Things have become more prolix, to be sure, but nothing has been done to exhibit system or structure in this part of our moral thinking: it could be that the putative judgments about 'how much of his right not to be punished he has lost because of his offence' are themselves a perfect jumble.

12. Coda

How far has the purely teleological picture presented in § 5 suffered from the incursions of the demands of justice? Those demands involve (a) reactive attitudes, or something on a continuum with them, to justify having a penal system at all, and (b) human sympathies, or something on a continuum with them, to set upper limits to the severity of penalties. Of these two elements, (a) merely restricts the penal code to actual offenders. It is (b) which could affect the content of the penal code, by keeping specific penalty-limits lower than they might have been set if utility were the only guide.

I contend that those are the only incursions of non-utilitarian elements which we should allow. So I am maintaining that in penal theory and practice considerations of utility should hold sway within limits set by the requirement that the upper (not the lower) limits on penalties should be just. That is a familiar enough doctrine, but I have tried to defend and deepen it through a general, Strawsonian theory of punishment.

Some people think that more harm than good is done by overtly emphasizing deterrence, reform, etc. in implementing a penal system, and that it would be better if ends were never mentioned, and offences were routinely penalized by penalty P to the accompaniment of judicial rhetoric about 'paying your debt to society.' If they are right, that is a reason for keeping quiet about utility in the courtroom and the prison; but so far from being a reason not to handle penological problems teleologically, it is itself a utilitarian reason.

One large problem-area remains untouched. Although I introduced Strawson's line of thought as a theory of the accountable/non-accountable line, I have not brought it to bear on the vexed question of criminal accountability or responsibility.

A purely remedial-teleological system needs only a Schlickian notion of responsibility, deeming offenders to be responsible in proportion as they are susceptible to deterrence by moral pressures. But a genuinely penal system, if this is understood in Strawson's way (and what other way is there?), needs the concept of responsibility sketched at the end of §2 above. That is, it will want to know whether the act really did manifest the unacceptable attitude or intention which it seemed to, and whether the agent is capable of entering into fully-fledged adult interpersonal relationships.

The choice between Schlick's account of the line and Strawson's might affect what happens in courts of law. For although over a large extent of everyday life we draw a single accountable/non-accountable line which is accurately located both by Schlick's and by Strawson's descriptions, the descriptions may sometimes part company, so that for a given action the agent is accountable according to Schlick but not to Strawson or vice versa.

The explicit adoption of either account would surely produce changes in the present indeterminate and muddled handling of the concept of responsibility in the courts. And although in public affairs clarity is not always desirable, I suspect that at least some uses of the concept of criminal responsibility would be improved if
Towards a Theory of Punishment

we — our courts, our civilization — became clearer about it. For example, many people tend to think that if a crime is utterly vile, cruel, depraved, the criminal must be insane and thus not a fit subject for punishment at all; yet none of us has a sufficiently clear, determinate, and valid concept of insanity to entitle us to such an inference. Perhaps we don't need one: perhaps we are right to think that punishment is inappropriate for the worst criminals, but not because they are insane. Strawsonian reasons might be given for that. But a full discussion of how reactive attitudes relate to criminal responsibility must be reserved for another paper.

FOOTNOTES


3 op. cit., p. 23.


5 op. cit., p. 21.

6 Ibid., p. 22.

7 Ibid., p. 22.

8 Peter Remnant has pointed out to me that if speeding were a capital offence, the death-rate from traffic accidents and from executions of speeders might well be lower than the present death-rate from traffic accidents alone.

9 I have been greatly helped by discussions with Gillian Bennett, S.C. Coval, Peter Remnant, and Robert J. Rowan.