Luck and Retribution*

JONATHAN JACOBS

In the persistent debate about punishment, several different fundamental issues are involved. One is the question of the justification of punishment. Another is the question of the aim of punishment. Even if one justifies punishment in terms of its aims, there is still a point in recognizing that aim and justification need not be identical. A third issue concerns the conditions of liability to punishment. This is the question of who is to be punished, and what are excusing and mitigating conditions. While these issues are connected, it is important to distinguish them. Otherwise, the dialectic becomes confused, and there is unclarity with respect to the work that the various considerations are to do as reasons or as objections. The present discussion focuses on justification. In particular, it will be argued that there is a distinctive role for retributivist considerations in the justification of punishment, and that retribution is not properly an *aim* of punishment. The case for this will be made by first examining a vulnerability of any justification for punishment driven by considerations of its aim, and then looking at some positive reasons in favour of retributivism. A broad notion of luck will be relevant to both aspects of the argument, in a way that will be indicated in section II.

I

Retributivists sometimes speak in the idiom of annulling unfair advantage, restoring a just balance, denouncing the values or the

* I am very grateful to the Earhart Foundation for awarding me a Fellowship Research Grant to support my work during a sabbatical semester. I spent much of that semester as the John MacMurray Visiting Professor of Philosophy at the University of Edinburgh. The members of the department were both very welcoming and very helpful. Comments made by several members of the department, in particular, Dory Scaltsas, Vinit Haksar, and Tim Williamson, when I presented a version of this paper in the Department of Philosophy, were instructive and are much appreciated. Antony Duff, of the University of Stirling, made a number of helpful remarks and objections in conversation about this topic and this paper. Also, the Editor of the journal made a helpful suggestion which has been incorporated.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

behaviour of the offender, reprobation, and the like. These, while they may be appropriate as aims of punishment, are not quite (non-consequentialist) justifications of it. H. L. A. Hart writes of several retributivist theories, that they 'all either avoid the question of justification altogether or are in spite of their protestations disguised forms of Utilitarianism, ...'[1] and he may be largely correct in that observation. What often occurs is a confusion or conflation of retributivism with certain aims of punishment. If the aims are not directly concerned with social benefit, then it might be claimed that the justification is retributivist. But, as Hart notes, retributivist views tend to involve consequentialist considerations. Perhaps the conception of retributivism to be defended here is somewhat revisionary, but that should not disqualify it, especially if other, more familiar uses are problematic. The point is not to ignore the aims of punishment, and to make a case for 'pure' retributivism, but to identify the distinctive role of retributivist, non-consequentialist considerations in justification. A retributivist justification does not foreclose on the position that the aims of punishment may be multiple, and that consequentialist considerations should figure among them. That issue, though, is not just the same issue as the question of the justification of punitive sanction.

A fairly standard objection to retributivism is that it licenses punitive sanctions whether or not they do, or are even intended to do, any good. So, retributivism often meets with the charge of gratuitously bringing more suffering and injury into a world already abundantly supplied.[2] In a critique of retributivism, Anthony Kenny writes:

> Popular wisdom, which has many adages which seem to favour the retributive theory, has one which is conclusive against it: two wrongs don't make a right. We must not render evil for evil.[3]

He adds:

> Oddly enough, the root of the error contained in the retributive theory of punishment is the same as the root of the error con-

[1] H. L. A. Hart, *Punishment and Responsibility*, (Oxford University Press, 1995), p. 9.

[2] This is aggravated by the fact that disproportionate numbers of offenders are members of disadvantaged groups in society, and that punishment of them puts additional impediments in the way of those already struggling on account of severe and persistent inequalities of opportunities and resources.

[3] Anthony Kenny, *Freewill and Responsibility*, (New York: Routledge & Kegan Paul, 1978), p. 73.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

> tained in the purely remedial theory. Both theories attempt to give an account of crime and punishment as two episodes in a criminal's life, considered in isolation both from the authority imposing the punishment and the society in which the criminal lives. The retributive theory, starting from the premise that it is just to punish, reaches the conclusion that it is just to render evil for evil. The remedial theory, starting from the premise that it is unjust to render evil for evil, reaches the conclusion that it is unjust to punish.[4]

There is, though, a retributivist dimension to the justification of punishment that does not fall victim to the 'two wrongs' and 'isolation' objections that Kenny identifies. We will work our way towards it by fashioning responses to the questions, (a) what justifies sanction? and (b) what justifies punitive sanction? After all, there could be a system of laws or norms without institutionalized sanction (though, more on this below); and sanction can respond to violations in ways other than harming. It could respond with therapeutic or remedial measures. First, though, we shall consider some other important arguments against retributivism.

It might be supposed that a place for retributivism is secured by insisting on proportionality (between offence and sanction). This, however, will not make the case for retributivism on its own. Hart writes:

> the injunction 'treat like cases alike' with its corollary 'treat different cases differently' has indeed a place as a *prima facie* principle of fairness between offenders, but not as something which warrants going beyond the requirements of the forward-looking aims of deterrence, prevention and reform to find some apt expression of moral feeling. Fairness between different offenders expressed in terms of different punishments is not an end in itself, but a method of pursuing other aims which has indeed a moral claim on our attention; and we should give effect to it where it does not impede the pursuit of the main aims of punishment.[5]

The relationship of these two injunctions of fairness to the issue of proportionality is that in punishing more serious crimes more severely, we are aiming to do what we deem to be required to effectively check more serious evils, and to do so in a way that meets public demands for denunciation of serious crimes. There is fairness across two dimensions: (i) across the dimension of treating

[4] Ibid., pp. 73–4.
[5] Hart, *Punishment and Responsibility*, p. 172.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

what is alike in a like manner, and (ii) across the dimension of responding with more severity to more grave offences. Neither of these essentially depends upon endorsement of retributivism.[6]

This is an example of a widely used argumentative strategy with respect to retributivism; namely, when there is apparent normative authority in what is claimed to be a retributivist consideration, domesticate that consideration to non-retributivist theory. Utilitarians, for example, argue on grounds of utility, that only offenders should be punished. (That *sounds* like retributivist 'just deserts' but the reasons given are utilitarian reasons.) Similarly, consequentialist justifications can accommodate the principle that more severe offences merit more severe penalties, and at the same time require that punishments be no more severe than is necessary to serve permissible social ends. This, it can be argued, is not a 'partial' retributivism; it is a textured consequentialism. It can appear that whatever contribution retributivist theory makes, it can be secured on non-retributivist grounds.

Nigel Walker pursues this strategy by arguing that as soon as we bring into the justification such matters as denunciation, or connecting the offender with correct values, or rectifying unfair advantage, we are no longer really articulating retributivism, but appealing to other kinds of considerations. Even with regard to punishment as conveying a message, as a kind of communication, he asks why such a message *must* be delivered, if the rationale for delivery has nothing to do with what good it might do.[7] Retributivism may seem to be 'good for nothing' both practically and theoretically.

It is fair to criticize retributivist justifications when they rely on consequentialist considerations. Retribution is *not* an aim of punishment, if by 'aim' we mean some state of affairs to be brought about, or intended to be brought about, by it. In fact, I confess to finding the notion of retribution as an aim of punishment to be rather odd. Whatever is *brought about* by punishment, it is not retribution. Retributivist considerations properly concern what punishment normatively is, or what is done *in* punishing. What practical ends punishment is engaged to, and the degree of its success in

[6] Hart writes:

> Much confusing shadow-fighting between utilitarians and their opponents may be avoided if it is recognized that it is perfectly consistent to assert *both* that the General Justifying aim of the practice of punishment is its beneficial consequences *and* that the pursuit of this General Aim should be qualified or restricted out of deference to principles of Distribution which require that punishment should be only of an offender for an offence. (*Punishment and Responsibility*, p. 9).

[7] Walker, *Why Punish*?, (Oxford University Press, 1991), p. 81.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

realizing them, are not directly the concern of retributivist justification. It is sometimes argued that denunciation or condemnation of the criminal is a retributive aim of punishment. However, to whatever extent moral condemnation is or should be an *aim* of punishment, there are other (though not distantly related) grounds for recognizing retributivist considerations as having a place in the justification of it.[8]

II

What does luck have to do with the debate over justification? By 'luck' I do not mean chance or randomness, but the variety of ways in which people are not in control of many of the circumstances of their lives, and also the results and consequences of their actions.[9] One area in which luck is in play in a large way is with regard to the effects and efficacy of punishment.

It is clear that punishment affects different people in different ways. Some who are punished are thereby motivated to genuinely reform themselves. Some just decide that it is wiser to not risk being punished again, so they stop committing offences. Some feel shamed by conviction and punishment, others become embittered or defiant or alienated. Some people may be deterred by knowing what are the sanctions and knowing that they are meted out. Yet, there are no doubt others who do not know the sanctions, or do not care what they are. Others may know that conviction is uncertain

[8] See, for example, James Fitzjames Stephen, *Liberty, Equality, Fraternity*, for a version of this view, and see Hart's responses to it in *Punishment and Responsibility*, (pp. 169–73). Other defenders of denunciation and the expressive view are Joel Feinberg ('The Expressive Function of Punishment' reprinted in *A Reader on Punishment*, A. Duff and D. Garland (eds), (Oxford University Press, 1994), and Hyman Gross (*A Theory of Criminal Justice*, (Oxford University Press, 1979).

[9] The issue of moral luck has been extensively discussed in recent years. See, for example, Thomas Nagel, 'Moral Luck' in his *Mortal Questions*, (Cambridge University Press, 1979), and also Bernard Williams, 'Moral Luck', in the collection of his papers, *Moral Luck*, (Cambridge University Press, 1981). In the present discussion I do not enter in any detailed way into the issues of moral psychology that are at the centre of many treatments of moral luck. I am considering luck in a wider, looser way, as concerning the many ways in which the conditions and consequences of our acts are not under the control of our intentions and exertions. It is very easy for human beings to exaggerate their autonomy and to think that they have control, when in fact, they are only considering a very narrow 'slice' of the usually untidy and complex realities.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

and sentencing is erratic, so they take their chances. The list could grow to great length.

There are several important results, the following among them. (1) There is considerable uncertainty about just what consequences are brought about by specific practices and policies of punishment. (2) It is unclear to what extent offenders are corrected, or are connected or reconnected to correct values. (3) It is unclear to what extent feelings of public safety are enhanced by different penal policies. (4) It is unclear how much criminal activity is deterred by fear of being caught and convicted. (5) It is unclear to what extent crime is effectively denounced, or to what extent victims of offences are, or feel that they are, vindicated by the sanctioning of offenders. (6) It is unclear which offenders would respond more positively to a punitive or a remedial sanction. And so forth.

These matters are often carefully studied, and there is a great deal of data of various kinds. But no stable, clear picture emerges of just what is accomplished by punishment, and there is scant basis for the confident assertion of the relevant sorts of counterfactuals. For example, of which people, or of what proportion of a certain group, can we say with confidence, such things as 'If they had stayed in prison three years rather than eighteen months, the number of property crimes and personal assaults committed by them during their first two years out of prison would have been X% lower'? Or, 'If the sentence had been three years instead of one year, the difference in the deterrent effect would have been X?' What else would have to be different? What other differences would those differences make?

Moreover, there are further-reaching effects of punishment which also affect family members and associates of the offender. In one case an offender's sentence of three years may be a shaming experience for him and a survivable hardship for the family. In another case, it may prove much more destructive, with irreversible effects. The ghetto-dwelling, drug-dealing teenage son of a single mother on public assistance may regard doing time as a rite of passage, while the middle-class university student who gets caught at it the first time, regards it as a disaster, and so too does his family.[10] In fact, it may actually be more of a disaster for the former, though he does not recognize it as such. (That is part of the disaster.) The causal reach of punishment extends in ways that are often not fore-

[10] See M. Tonry, 'Proportionality, Parsimony, and Interchangeability of Punishments,' in *A Reader on Punishment*, Antony Duff and David Garland (eds), (Oxford University Press, 1994), pp. 136–60. That is a fuller discussion of this issue of how facts about offenders make a difference to what a punishment actually counts for.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

seen, and the causal efficacy of punishment is very difficult to predict.[11] This is one general reason for a measure of scepticism with regard to consequentialist justifications of punishment. We know that it does many things, but we are not sure just what things, and we are not sure how it does them.

The plain fact that others beside offenders are affected and are often harmed by sanctions has been used as an objection to retributivism. Retributivism can be criticized for isolating the relation between sanction and offender, and ignoring the effects (often considerable harms) on others. Walker writes:

> Those who object to utilitarianism because it seems to condone the punishment of the innocent ... might ask whether the harm done every day to prisoners' dependants in the name of retributive sentencing does not amount to the same thing.[12]

He says 'The retributive sentencer cannot escape this dilemma.'[13] That may be so, and retributivism has to acknowledge the reality of what Walker calls 'obiter punishment'.

Our insistence that we remain clear about the difference between justification and aim will not make the difficulties disappear; they are still real. Still, retributivism is no worse off than other justifications with respect to obiter punishment. It cannot, in any case, be avoided. When we look at who is punished for what, why that person is liable to sanction, and what is the aim of the sanctioning, we look at each of these things under quite narrowly circumscribed descriptions, the ones relevant to the legal dimensions of the situation. It is never the case that what is going on, or what results from what is going on, is limited to what is captured and expressed by those descriptions. There will always be additional descriptions, and many of them will describe unwanted and unintended effects. That fact is made no more bearable if the justification for punishment is non-retributivist.

Obiter punishment and incidental harms to the sanctioned offender (such as stigma, loss of job opportunities, and so forth) are aspects of the sort of luck we have been talking about. Luck of this sort is not eliminable from the facts of life. The justice of this or that action or practice cannot ensure that it is *totally* just in the sense that there is no injustice in what eventuates. But retributivist justification is at least minimally vulnerable to being interpreted as a bet that our luck will be good if we punish. Luck, in the sense of cir-

[11] See Walker, *Why Punish?* ch. 13, for discussion of the effects of punishment on family and associates of offenders.

[12] Walker, *Why Punish?*, p. 107.

[13] Ibid., p. 107.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

cumstances and outcomes that are not under our control or directly responsive to our intentions and purposes, makes for a looseness between any consequentialist justification and the actual states of affairs brought about by the practice. The connection between the practice of punishment and the aims of punishment is less like the connection between mandatory immunization, and reduction in the incidence of disease, and more analogous to the relation between sex education for adolescents, and their adolescent and adult sexual behaviour, or between inflation and unemployment.

The upshot is that we do not know to what extent punishment accomplishes the various ends that have been set for it, (which are often multiple and not particularly complementary, and also lack the relevant sorts of institutional infrastructure and support). In that regard, claims about what punishment is aimed at, *when offered as part of a justification of punishment*, are vulnerable on epistemic grounds, independently of any moral reservations we might have about consequentialist justifications. And that is a serious matter for a consequentialist justification.

Retributivism is not vulnerable in that way. The argument is not that retributivism 'wins' because of the other candidates' vulnerability on this score. The main point so far is that luck is especially problematic for consequentialist justifications. We turn now to some more positive considerations in favour of retributivism.

III

It is quite right to point out, as Kenny did, that it is a mistake to see crime and punishment as two episodes, isolated from the larger institutional and social setting. The most basic reason why there are criminal justice systems is to keep people from committing certain kinds of acts. (This is a matter distinct from whether the aim of punishment is or ought to be crime reduction.) We can agree with Kenny when he says:

> The primary purpose of punishment is to prevent citizens from becoming criminals; hence punishment will never be understood if it is regarded purely as a transaction between society and those who have already become criminals. When a man is punished, the purpose of punishment has, in his case, already been thwarted.[14]

It is socially beneficial to have such a system and for its authority to be recognized and effective. With respect to explaining the existence

[14] Kenny, *Freewill and Responsibility*, p 76.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

of institutions of criminal justice and the general social aim of punishment, there is no question about that.We are not trying to answer the question 'why are there criminal laws and sanctions for violating them?' Retributivist considerations enter the picture when we are trying to get clear about what ethically justifies punitive sanction as a feature of a criminal justice system.

Retributivism makes a point about what punishment is, as a rational undertaking, not about punishment's efficacy as a causal process. Or, we might put the point this way: the retributivist is making a point about the rational sense of imposing penalties as a response to offences. We will see below that there are considerable costs to the coherence of our ethical beliefs and commitments if we abandon retributivism. We will first say a bit about the role of sanction in a normative system, (briefly deferring the question of *punitive* sanction) and then argue that punitive sanction, retributivist sanction, is appropriate.

While the extent to which the aims of punishment are realized is contingent, the relation of sanction to an effective system of normative authority is not contingent. This is the first point, viz., in the absence of sanction, a system of normative authority would be merely notional, and the extent to which it was actual would be entirely a matter of luck. (Primarily, luck in respect to how strongly individual agents' internal sanctions coincided with the norms of the system.) In the absence of sanction, enactment of the norms would be, so to speak, discretionary. It would be as if the norms were put forward merely as recommendations. In the absence of sanction, there would be no institutional, morally regulated, public system through which violations of norms could be addressed in a systematic and norm-regulated manner. This may sound like a Hobbesian point, but the focus here is not on 'the Sword' but on the fact that a system of norms without sanction is incomplete as a system of norms.

The answer to the question, 'what *difference* is made by the imposition of sanction?' *is*, in many respects, a matter of luck. It is a causal matter, subject to all variety of causal complexities and contingencies and it is subject to the sorts of epistemic uncertainties indicated above. But *that* a normative order permits or requires sanction is not a contingent matter. It is an essential mode of actualization of that order. It is one dimension of the reality of the norms. Still, an objector may persist in asking why sanction needs to be understood *retributively.* It could be, for example, educative, therapeutic, or correctional. What is the moral rationale which motivates *punitive* sanction?

We might note that a focus on luck may actually seem to provide

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

materials for the case *against* retributivism. There is luck with regard to what sorts of natural tendencies and abilities one has. This makes a difference to one's character, and as a result, to one's policies of choice and action. What we are like, and what moves us and concerns us is very much a matter of natural endowment and the influences of others. There is luck with regard to who we are surrounded by, and what we are encouraged in and discouraged from. There is luck with respect to what circumstances we face, and whether our courage and integrity, for example, are severely tested. And so on. Taken jointly, these can motivate scepticism about the sorts of agency and responsibility that are standardly assumed by retributivism. If indeed human action is just one strand in a more global causal nexus, then should we not reinterpret both the justification and the practice of punishment in causal (and consequentialist) terms?

We need not attempt to settle the free will and determinism debate in order to answer that question. The extent to which luck is involved in human action does not undermine the significant distinctions that can be made between voluntary and involuntary acts, and between conditions for liability to punishment and excusing and exempting conditions. It is plain that sometimes causal factors operate in ways that warrant attributions of diminished responsibility. It is also plain that most human action is not so affected. What count as coercion, compulsion, provocation, defects of sanity and the like, are by no means obvious or unproblematic. Still, we do appropriately regard ourselves and others as responsible agents, unless there are special reasons not to. There is no special burden on retributivism to metaphysically underwrite that fact; there is no special burden on retributivism to make a successful case for libertarianism or contra-causal agency. We can acknowledge the significance of each person's 'story' without the individual's agency being assimilated to it in a way that dissolves the grounds for reward and penalty.

If rationality and internal sanction were adequate for virtue, then punitive sanction would be unnecessary. In addition, we would not have occasion for resentment, because there would be no actions that merited resentful opposition; people would regularly act lawfully. We are not like that. If one wishes to exploit this fact and argue that it shows that retributivism itself rests on an especially fundamental contingency, namely, the fact that we have the sort of constitution that we have, then the point is taken. Maybe there is luck or contingency at the 'bottom,' but that does not mean that there are only contingent connections between the main elements of a system of norms for human beings.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

It must be admitted that punitive sanctions cannot literally undo the past, annul the crime, or restore a just balance between offender and victim. In addition, it may be that the offender is the sort of person who is unimpressed by, or even uninterested in the justice of his conviction and sentence.[15] In any case, when we are seeking the *justification* of punishment, we are not automatically required to think of this in terms of its efficacy or its intended purposes as a means to something. Those matters do not supply the only possible rationale for including punitive sanction as a dimension of a system of norms.

The suggestion with respect to retributivist justification is this: punitive sanction can be rationally invited just as resentment or gratitude can be rationally invited. *They* are invited in a sense in which it would not be correct to say that envy and spite, for example, are invited. We can often understand why someone is envious or spiteful, but these are states for which there is not an ethically fitting rationale as there is for the other states mentioned. The acts and facts that cause them do not merit them in the way that other attitudes and responses are merited. The actual occurrence of both the ethically fit and the less ethically fit responses depends substantially on the dispositions and sensibilities of the agents who have the various responses. So, it could be that nearly everyone felt only the lust for vengeance, rather than recognizing the appropriateness of ethically sound retribution. That the responses to criminal wrongdoing should be institutionally regulated in a manner designed to establish and preserve fitness, is part of what law (ideally) is about. The responses are depersonalized. Legally enacted punitive sanction, like resentment and gratitude, is (or, it can and should be) a

[15] See my 'Taking Ethical Disability Seriously,' in *Ratio*, Vol. XI, No. 2, September 1998, pp. 141–58. I develop the view that an agent may voluntarily come to have a character such that it is practically impossible for him to engage with ethical considerations in a sound manner. He may be disabled for virtue, but none the less responsible. Thus, there may be offenders who do not accept or even really see the moral justification for the punishment they suffer; but that does not make the punishment unjust. There may also be offenders who are genuinely remorseful, and who fully understand the wrongfulness of their act, and who genuinely resolve not to act in such a way (or any seriously wrong way) again. Such an agent, we might think, does not 'need' punitive sanction. But my point is that the justification of sanction is not a matter of what efficacy it is likely to have. The remorseful and repentant agent does not need punishment to connect him with correct values. But he *did* the thing. That's the main factor determinative of whether punitive sanction is appropriate.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

rational response to agents and actions with certain features.[16] It is when it is motivated and shaped by morally sound sentiments reflected in the law.[17]

[16] After having written this section, and almost all of this paper, I read Antony Duff's 'Expression, Penance and Reform,' which is a chapter of his *Trials and Punishments*, (Cambridge University Press, 1986) and is reprinted in *Punishment and Rehabilitation*, third edition, Jeffrie G. Murphy (ed.), (Belmont California: Wadsworth Publishing Company, 1995), pp. 169–209. What I have said about gratitude and admiration as a guide to understanding resentment, sounds, in some respects, similar to much that Duff says about gratitude and admiration in section 2 of 'Expression, Penance and Reform'. His emphasis, though, is different, chiefly on account of the expressive role that he assigns to moral sentiments, and the way in which he interprets the communicative role of punishment and its relation to penance.

[17] Adam Smith, in *The Theory of Moral Sentiments*, has a good deal to say about resentment. Two features of his treatment of the topic that are particularly relevant here are expressed in the following two quotations:

> … the violation of justice is injury: it does real and positive hurt to some particular persons, from motives which are naturally disapproved of. It is, therefore, the proper object of resentment, and of punishment, which is the natural consequence of resentment. (Adam Smith, *The Theory of Moral Sentiments*, D. D. Raphael and A. L. MacFie (eds), (Indianapolis: Liberty Fund, 1984), p. 79 [II. ii. 1. 4]) And:
>
> When the guilty is about to suffer that just retaliation, which the natural indignation of mankind tells them is due to his crimes; when the insolence of his injustice is broken and humbled by the terror of his approaching punishment; when he ceases to be an object of fear, with the generous and humane he begins to be an object of pity. The thought of what he is about to suffer extinguishes their resentment for the sufferings of others to which he has given occasion. They are disposed to pardon and forgive him, and to save him from that punishment, which in all their cool hours they had considered as the retribution due to such crimes. Here, therefore, they have occasion to call to their assistance the consideration of the general interest of society. They counterbalance the impulse of this weak and partial humanity by the dictates of a humanity that is more generous and comprehensive. They reflect that mercy to the guilty is cruelty to the innocent, and oppose to the emotions of compassion which they feel for a particular person, a more enlarged compassion which they feel for mankind. (Ibid., pp. 88–9.)

The argument in this paper does not depend upon an endorsement of Smith's moral theory, but what he says about moral sentiment has important points of contact with the present view. There *is* a proper object of resentment, and there *is* a proper enactment of it. I think he is quite right to note that resentment often changes into pity when the offender has been caught, been convicted, and faces punishment. However, I think he is also correct to insist that when resentment is enacted in an impartial manner, as

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

It is a mistake to see these sentiments as 'mere' passions which we must try to manage in effective and salutary ways, passions which we should harness and 'put to work' in a way that is beneficial rather than injurious. (This is another symptom of the instrumentalist model of rationality.) Speaking more generally of the ethical context, of course it is a failing, in certain circumstances, not to feel and to show gratitude or to admire others. The failure is an ethical flaw, not simply a matter of the absence of certain passions. The flaw often involves a *refusal* to be grateful, or a refusal to admire. This is often more than just the absence of certain feelings; there may be unwillingness to acknowledge certain considerations as reasons of certain types. Or, if there is not a deliberate effort at refusal, then the unwillingness may be indicative of dispositions that are ethically unsound in a quite deep way, concerning one's valuations and modes of engagement with others.

In the concluding chapter of *Why Punish*?, Walker recapitulates his arguments against retributivism, and writes, 'This does not mean that retributivists are completely routed: only that they must retreat to ground where they are less open to attack—their right to have retributive feelings.'[18] It is though, tendentious to characterize retributivism as a matter of sensibility in *that* way. The implication is that there is a measure of unavoidability about this feeling, so, it has a place in the theory of punishment because it has to be accommodated. This suggests that we are in the position of the host who says to the guest who *will not* leave, 'Oh, all right, I guess you can stay.'

It is better to see retributive sentiments for what they are, viz., responses and attitudes which can themselves be more rational or less rational on account of what sort of comprehension of things they figure in, and how they figure in it. There is, as Aristotle observes, a way of being angry 'to the right extent, at the right time, with the right motive, and in the right way…' (1109a 27–8). Resentment need not be vindictive, and it can be punitive without being vengeful or ethically dubious. It can itself be a proper feature of sound judgment rather than a passion that interferes with, or

[18] Walker, *Why Punish*?, p. 139.

through sanction, supported by due process, we see that it is not vengeful, and that objective considerations are in favour of it rather than against it. This is one reason why conviction without sentencing is not enough. If we were to say, 'Well, look; he's been caught and tried and found guilty. It would be wrong or hardhearted to punish him' we have confounded the legitimacy of the considerations that gave us reason to charge and convict the offender in the first place.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

overwhelms rational judgment. Aquinas, for example, distinguishes between anger antecedent to a judgment of reason and anger consequent to a judgment of reason.[19] In the latter case, the passions 'do not impede the judgment of reason which has already preceded, but rather help to execute the judgment of reason more promptly, and in this way the passions are useful to virtue.'[20] And Reid writes:

> But in most of the operations of mind in which judgment or belief is combined with feeling, the feeling is the consequence of the judgment, and is regulated by it. Thus, an account of the good conduct of a friend at a distance gives me a very agreeable feeling, and a contrary account would give me a very uneasy feeling; but these feelings depend entirely upon my belief of the report.[21]

Aquinas and Reid have quite different philosophies of mind and moral psychologies. The point of referring to them is just to note that the claim that moral sentiments and passions can be well-placed or not in ways that are rationally assessible, is a familiar and plausible one in philosophical theorizing; it is also familiar and plausible claim of the plain person's philosophically untutored understanding.

A retributivist justification of punishment is not a way of 'dealing with' retributive feelings, (feelings which may indeed be intense and undisciplined). Nor is it a way of substituting passions for reason as determinative of moral judgments. The point is not that if one feels resentment, that, in its own right is *evidence* that one's opposition to a type of action is ethically sound. The point is rather that appropriate resentment is partially *constitutive* of ethically sound opposition to certain kinds of behaviour, and that this can be reflected in the law.[22] For it not to be reflected through the law via punitive sanction would be a distortion and misrepresentation of a legitimate and significant moral response. The notion that the relevant sentiments could be transformed or redeployed in the service

[19] See Aquinas, *De Malo*, 12, 1.

[20] Ibid., *De Malo*, 12, 1.

[21] Thomas Reid, *Essays on the Active Powers of Man*, in Thomas Reid, *Inquiry and Essays*, Ronald E. Beanblossom and Keith Lehrer (eds), (Indianapolis: Hackett Publishing Company, Inc., 1983), p. 365.

[22] There are of course, many types of behaviour that merit being resented, but which are not, and should not, be criminal behaviours meriting punitive sanction. Not everything that (properly) invites resentment is or should be a *criminal* offence. Thank heaven for that. The importance of acknowledging this is a point I owe to Vinit Haksar. The point of noting the distinction between the evidential role of the moral sentiments and their constitutive role was suggested by remarks by Tim Williamson.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

of fundamentally meliorative purposes is symptomatic of a denial of their cognitive substance and rational legitimacy, and also indicative of an exaggerated and misplaced optimism about the plasticity of moral sentiments.

To fail to address criminal violations with punitive sanction would be a denial of the normative authority of ethically sound sentiments.[23] This would be no more sustainable or legitimate than a denial of the suitability or enactment of gratitude, or admiration, when they are merited. There is no guarantee that in fact the sorts of resentment that most people feel *are* apt or well-placed. They are often mean-spirited, indulgent, and motivationally toxic. Moreover, in the writing of laws and in sentencing practice there surely can be a disturbingly wide gap between actual practice and what is required by sound moral sentiments. This account of the justification of punitive sanction does not guarantee its own good faith employment. But it is a justification that is not dependent upon good luck with respect to the effects of this or that penal practice or the purposes of any specific policy.

We can admit that punitive sanction often does not motivate rehabilitation or penance and reform. Also, though it is a way of acknowledging the voluntary agency of the offender, we should not overestimate the extent to which those sanctioned will welcome sanction as a mode of respect, though some may well regard being punished as an occasion for reformative reflection and exertion. These facts do not undermine retributivist justifying considerations, and they do not outweigh the fact that abandonment of retributivism would require a significant dimension of ethical rationality to be disengaged from practice or somehow redeployed.

In addition, we are, for example, accustomed to using considerations of sympathy and compassion (not simply pathos or melodramatic manipulations of emotions) as grounds for calibrating judgments and sentences. This is not the interference of something outside the cool rationality of reason or the law. The law, in its refinements concerning liability to punishment, degrees of culpability, and discretion in sentencing, can reflect morally serious discriminations, many of which register the rationality of moral sentiments.[23]

[23] In 'Hatred: a qualified defence,' Jeffrie Murphy writes:

> Since I regard retributive hatred as in principle the natural fitting, and proper response to certain instances of wrong-doing, I do not regard the passion itself as either immoral or irrational. (It is not the moral equivalent of a phobia.) I do, however, believe that it is generally both irrational and immoral to be *led*, to use Spinoza's phrase, by this dangerous and often blind passion. [In Jeffrie G. Murphy and Jean Hampton, *Forgiveness and Mercy*, (New York: Cambridge University Press, 1988), p. 108.]

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

IV

We will take a brief look now at two approaches that may seem to share a good deal with the present approach, but which are in fact different in important respects.[24]

In 'The retributive idea,' one of her chapters in *Forgiveness and Mercy*, Jean Hampton writes:

> So I am proposing that retributive punishment is the defeat of the wrongdoer at the hands of the victim (either directly or indirectly through an agent of the victim's, *e.g.*, the state) that symbolizes the correct relative value of wrongdoer and victim. It is a symbol that is conceptually required to reaffirm a victim's equal worth in the face of a challenge to it. Thus the punishment has a telos, but the telos is not so much to produce good as it is to establish goodness.[25]

She says a few pages later, 'But the retributive motive for inflicting suffering is to annul or counter the appearance of the wrongdoer's superiority and thus affirm the victim's real value.'[26] The retributivist, she says, 'is using a competitive victory to deny an incorrect claim about relative worth.'[27] And, 'the demand for retribution is the demand for the defeat of the wrongdoer, where his humbling is seen as a way of symbolizing the value of the victim whom he transgressed.'[28] And, 'an adequate theory of retribution would show how

[24] One large issue raised by these matters, but also falling outside the scope of the present discussion, is the relation between justice and mercy. Suppose justice permits or requires punitive sanction. Would it then be wrong for a sentencer to show mercy to an offender? It is important to distinguish between suspending the requirements of justice, and calibrating them to the particular features of situations. The law is inevitably something of a blunt instrument: it cannot in itself address or foresee all of the ethically relevant fine-grain of particular situations. Still, its countenance of mitigating factors, its attempts at discriminating degrees of severity and culpability, and the discretion sometimes given to sentencers, are among the ways in which the calibration is achieved. Again, I will not enter here into a full-scale discussion of the issue, but just note that many calibrations that might be (mistakenly) identified as suspending the requirements of justice and 'showing mercy' are in fact ways of answering to the requirements of justice.

[25] Jean Hampton, 'The retributive idea,' in *Forgiveness and Mercy*, Jeffrie G. Murphy and Jean Hampton (eds), (New York: Cambridge University Press, 1988), pp. 125–26.

[26] Ibid., p. 130.

[27] Ibid., p. 138.

[28] Ibid., p. 157.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

it was a response to what was immoral about the wrongdoer's action. This is precisely what the present theory tries to do.'[29]

On the one hand, a strength of her analysis is that it indicates a moral justification for punishment which articulates the good of punishment without making its good depend upon expectations about what states of affairs will be brought about by it. It does not rationalize punishment on instrumental grounds. On the other hand, what is problematic about the view is that punishment may be the playing-out of a moral drama in which the *dramatis personae* do not understand their roles. Perhaps the wrongdoer does not believe that his message has been annulled. Perhaps the victim doesn't think so either. Whether this happens is a matter of luck. The victim's value can be said to be vindicated in the sense that there is a rejection (which is not exactly the same thing as a defeat) of the offender's values, even if this goes unacknowledged by the offender, and there is an assertion of the value of the victim through the sanctioning force of the law which protects it. Still, I believe that there is a sense in which, on this interpretation, retributive punishment runs the risk of too often being *mainly* symbolic.

Another view in which the symbolic function is crucial is presented by Joel Feinberg in 'The Expressive Function of Punishment'. He argues that the distinction between punishment and mere penalties is marked by 'the essentially reprobative function of the former,'[30] and that punishment (hard treatment, in contrast to mere penalty) has as a central function to symbolize condemnation. In considering the objection that the symbolizing function could perhaps be fulfilled without hard treatment, he concludes that the 'question is surely open'[31] and that the problem of justifying punishment 'may really be that of justifying our particular symbols of infamy.'[32] That is the key place where his analysis departs from the present one. Granted, Feinberg does not say that the problem of justifying punishment just *is* the problem of 'justifying our particular symbols of infamy.' But if symbolizing condemnation or resentment were central features and functions of punitive sanction, then the question of eliminating its punitive dimension would indeed be open.

We have argued that punitive sanction is a morally apt enactment or actualization of a judgment. Could the judgment be enacted

[29] Ibid., p. 130.

[30] Joel Feinberg, 'The Expressive Function of Punishment,' in *A Reader on Punishment*, A. Duff and D. Garland (eds), (Oxford University Press, 1994), p. 83.

[31] Ibid., p. 87.

[32] Ibid., p. 87.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

without 'its usual physical media—incarceration and corporal mistreatment'?[33] This, I think, is doubtful. (Also, as Walker points out, the main condemnatory function could be performed by conviction, not sentencing.[34]) It is difficult to conceive of any substitute that would not be either merely symbolic, or an ethically dubious strategy of humiliation or some attempt at translating the appropriate moral sentiments into a quite different, and probably distorted form. An objector might prefer the merely symbolic substitute, since it could be claimed on its behalf that at least it did not inflict harm (or as much harm, supposing that at least to some extent it shamed or stigmatized offenders) as punitive sanction. But what would be the *point* of the symbolism that would make it more than a communicative gesture? How, if at all, would it be guaranteed that the communicated message is received? To what extent should it matter that it is? Sanction is not primarily undertaken as a dramatic expression of attitudes in ritualized manner. It may have that feature by the way, (and that may be welcome) but it is ethically rationalized by other considerations. This, anyway, is intended to be the key point of difference between the present view and the views of Hampton and Feinberg.

Punitive sanction may affect offenders by expressing or communicating condemnation. For some offenders, that may be constructive, in so far as it motivates reconsideration by them of their actions and their characters. It is easy however, to overestimate the extent to which condemnation is constructively expressed. The rationale for punitive sanction should not be dependent upon that. It has a rationale independent of that, in terms of the way in which sanction can actualize ethically merited resentment in a dispassionate, institutionally regulated manner. (Merited, or as I put it, 'invited' resentment is not just the same thing as elicited hostility. The distinction between being merited and being caused should be respected here.)

Having argued that the distinction between aim and justification should be respected, we might now be criticized for having left nothing to be the material of justification. If the material is not the expected benefit, and not the (hoped-for) effective communication of condemnation, and not the righting of a balance of values put askew by the offender, what is there left with which to ethically rationalize punitive sanction? Note first that none of these aims or effects are rejected as ethically impermissible or unwelcome. Still, the justificatory case for punitive sanction is a matter of it being the

[33] Ibid., p. 87.
[34] See Nigel Walker, *Why Punish?*, p. 24.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

only practice which directly enacts appropriate sentiments of resentful opposition in an undiverted and faithful manner.[35] In asking the question of justification, we are not only asking, 'what is it that permits this practice, given that it imposes harms?' We are also asking, 'what legitimate requirements for an ethically sound system of norms are fulfilled by this practice?' The imposition of punitive sanction is something participants in a normative system should undertake, not mainly because of what it might bring about, but because of what are the conditions for plausibly actualizing a system of norms.

It may 'establish goodness' (Hampton's formulation) and it may defeat offenders and vindicate victims. Retribution can do these things *because* it is something more than just an orderly way of discharging anger, outrage, and vindictiveness. It may be that as well, but it is not primarily an outlet; it is a public, objective actualization of resentment of wrongdoing, and an affirmation of the norms violated. It need not be any more metaphysical or mysterious than that.

V

Retribution is not 'good for nothing,' unless what is meant by that is that there is a moral rationale for it that is not dependent upon expected consequences. In addition, the objection that a justification of this type is irrelevant to actual practice is off the mark. It is an account of the moral rationale of punitive sanction, an account of the ground on which punishment stands as a feature of the moral landscape. It is not a sketch of sentencing policy.[36]

[35] Here again, there is a point of contact with Duff's view. He writes: '... for the kind of reform at which punishment aims can be achieved only through the kind of suffering which punishment aims to impose on, or to induce in, the criminal.' 'Expression, Penance and Reform,' in *Punishment and Rehabilitation*, p. 170. The agreement is over the unique fitness of punitive sanction to the ethical features of the issue. We do, however, disagree over the point of punitive sanction.

[36] I am aware that this account will attract the objection that it is too 'philosophical,' too remote from the actual concerns of jurists and legislators and prisoners and prison officials. It is though, a misdirected objection. The point of the present discussion is to diagnose the moral rationale of retributivism, not to sketch the outlines of a penal policy. Questions about what should be the sanctions for specific crimes, what should be the relevance of character and previous record in sentencing, what criteria we should use in trying to ascertain the effectiveness of our policies, and so forth, are of the first importance. But attempts to answer them should be undertaken on a landscape on which the crucial normative landmarks are clearly marked, so that we know why we are trying to answer those questions. And that task is a philosophical project.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

Defenders of retributivism often find it difficult to argue for it without bringing in consequentialist considerations. Justification and aim get intertwined and confused. Yet, not to do this seems to leave the case for retributivism resting on no more than intuitions about just deserts or repaying harm with harm, and so forth.[37] The understanding that agents can be, as voluntary agents, responsible for wrong acts, is part of the overall fabric of conceptions through which we participate in a common moral and social world. Actions are not just events that happen, some of them (and their consequences) pleasing or desirable, and some of them (and their consequences) not. Actions (typically and paradigmatically) are voluntary undertakings that enact the agent's values, in at least the 'thin' sense that the agent did the thing thinking that it was a worthwhile, or reason-supported or desirable thing to do. When we ask 'how could you *do* such a thing?' we often mean, 'how could you see doing *that* in a favourable light or as worthwhile?' Wrongful acts, unlike the unwelcome involuntary acts of those who are helplessly coerced or those who are insane or otherwise disordered, reflect agents' judgments of what is choiceworthy. Looked at across this dimension, retribution is justified as a feature of a system of norms that applies by being norms for rational, voluntary agents. It is not a strategy for improving offenders or society. It is part of the rationality of norms, for beings such as we are.[38]

Given all that there is to be said against the deliberate infliction

[37] See Jean Hampton's 'The retributive idea' in *Forgiveness and Mercy*. She writes of retributivists' attempts to supply a reason for why punishment is the appropriate sanction:

> Advocates of retribution typically answer that there is no *telos*, or goal, behind the suffering retributivists want: it is simply desired because it is supposed to be 'fitting' or 'suitable' or 'right.' And in response to queries about *why* suffering is suitable for wrongdoers, retributivists typically insist that this question has no answer. It is supposed to be bedrock intuition that, at the very least, those who are not guilty ought not to suffer pain, and more positively, that those who are guilty deserve to suffer in proportion to the pain they have caused. (pp. 112–13)

[38] The argument is not offered as a vindication of current practice. There is a great deal that is confused and morally suspect about current penal practice. It is not clear what are its aims, but it is clear that it fails. Moreover, there are many opportunities for discretion to be exercised, and it is not exercised in ways that are consistent and fair. Still, there is a point of the first importance in trying to get clear about what are the fundamental justificatory considerations for punitive sanction. That will supply a basis on which debates about sentencing, prosecutorial discretion, and other issues can be conducted in a more coherent manner.

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press

of harm and the deprivation of liberties and the other costs that punitive sanction imposes, the case for retributivism can appear to be based upon an unadmirable sentiment dressed-up in moral trappings. However, attempts to eliminate retributivist considerations or to explain them away in non-retributivist terms will leave an unreduced and significant core of legitimate retributivist thought. The rejection or exclusion of retributivist considerations would mean that instead of considering a criminal offense in terms of what it is, and what response it merits, we should look at it as an occasion to bring about some possible good. That is a misdirection of normative attention, which is both theoretically unsound and pragmatically fated to fail, because the sentiments which are to be reformed, properly resist reformation.

Colgate University

https://doi.org/10.1017/S0031819199000662 Published online by Cambridge University Press