

Varieties of Retribution*

JOHN COTTINGHAM

One of the functions of philosophy is to combat ambiguity and muddle. Yet in discussions of the justification of punishment, philosophers persist in talking of 'retribution' and 'retributive theory' as if these labels stood for something relatively simple and straightforward. The fact is that the term 'retributive' as used in philosophy has become so imprecise and multivocal that it is doubtful whether it any longer serves a useful purpose. In this paper I shall attempt to separate out no less than nine distinct approaches to punishment— I shall call them 'theories', but the term will be qualified later— which are or have been labelled retributive. The first object of the exercise will be clarification for its own sake; but along the way I shall allow myself some brief observations on the logical adequacy of various strategies for justifying punishment.

(1) *Repayment theory*

Most dictionaries give the first meaning of 'retribution' as *repayment*. The etymology of the term is clear: *re + tribuo*, Latin, to *pay back*. I am inclined to suggest that this notion encapsulates the basic or fundamental sense of 'retribution'. Etymology, of course, is no arbiter in philosophy. But there is a perfectly ordinary (pre-philosophical) way of talking about punishment which exactly fits this original sense of 'retribution'.

Newspapers and public speakers often hold forth about the need for the criminal to *repay his debt* to society. And the argot of criminals (or at least the criminal of fiction) is full of expressions like 'I've paid the price':

I'm going straight, I am, straight as an arser,
I've paid the price and done me time;
I'm going straight, and I don't mean the straight and narrer,
'Cos I'm going straight back to crime!

On our standard theory of retribution, then, punishment is inflicted in order to make the offender pay for his offence. Exactly how or why suffering something unpleasant (e.g. 'doing time' for six months) should count as payment for an offence is left unexplained; but it has to be admitted that the notion is both ancient and widely held. In the Pauline doctrine of the atonement, the Passion (suffering) of Christ is supposed to 'pay for' the offences of mankind. (Note: this is merely an example of the belief that suffering pays for wrong; I do not mean to suggest that the repayment version of retributivism is really an atonement theory. There could not be an atonement theory of punishment; atonement is something voluntarily undertaken, punishment something exacted.)

(2) *Desert theory*

A recent textbook defines 'retributive punishment' as 'punishment meted out because it is deserved'.¹ Similarly, Ted Honderich in his book *Punishment* asserts quite bluntly that to give as a reason for the rightness of punishment that it is deserved by offenders is

* This is a typescript of which the definitive version was published by Blackwell Publishing in *The Philosophical Quarterly*, Vol. 29, No. 116. (July 1979), pp. 238-246.

¹ Golding, M. P., *Philosophy of Law* (Englewood Cliffs, 1976), p. 89.

retributivism.² Unlike our first theory, which is grounded in the original meaning of ‘retribution’, these accounts are initially puzzling. No doubt it may be true that many who call themselves retributivists have held that desert is a sufficient condition for just punishment; but to say, *tout court*, that this is what makes them retributivists leaves the rationale for the label compressed to the point of unintelligibility. Suppose I say: “I’m a retributivist: I believe that where punishment is deserved this is sufficient to justify it”. I think the initial reaction of that ubiquitous figure the ‘intelligent layman’ would be: “Well, go on, explain! Where does the retribution come in?” To put the point more precisely, if someone claims that a sufficient condition for the justification of university scholarships is that they are deserved, does this mean that he has a retributive theory of scholarship-awarding? Perhaps it does; but only in a rather old-fashioned sense of the term ‘retribution’, meaning any kind of reward or recompense (“never did a charitable act go away without the retribution of a blessing”³). If we are allowed to unpack and amplify the desert theory in this way, the claim becomes that punishment is justified because it is a deserved requital or reward for wrong-doing. Viewed as an exercise in justification, this account shares the curiously jejune quality of theory (1); and its detractors have made the complaint— not without force— that it reduces to the bald assertion that it is simply just that the offender should be punished.

Is there a logical connection between the desert account (2), and our basic account (1)? Thesis (2) is certainly compatible with— perhaps even entailed by— thesis (1). For it may be that to claim that punishment is a just retribution (repayment) implies that the offender deserves to pay. The converse entailment, however, does not seem to hold. For it seems that one could be a ‘retributivist’ in the unexpanded Honderichian sense (punishment is justified because deserved) without subscribing to the idea that punishment is some kind of repayment of a debt. This alone is enough to warrant keeping theories (1) and (2) distinct.

(3) *Penalty theory*

Kant, universally regarded as a retributivist, asserts that the reason why punishment is imposed must always be because the individual on whom it is inflicted has committed a crime.⁴ A possible, though unlikely, interpretation of this remark is that the breaking of a law is, in itself, a sufficient condition for just punishment. This latter proposition was put forward, half a century ago, in a highly influential paper by John Mabbott. In connection with his own experiences in dealing with defaulters who had broken college rules, Mabbott wrote: ‘they had broken a rule and they knew it and I knew it; nothing more was necessary to make punishment proper’.⁵ Mabbott’s theory seems to construe punishment as a kind of automatic penalty, whose appropriateness is ensured simply by the knowing commission of an offence.

It is not clear that this theory, as it stands, will serve as an account of the justification of punishment (and the author later made substantial modifications in his position to deal with this point). However, Mabbott’s original unmodified theory deserves inclusion in our list for historical reasons. Many of those influenced by Mabbott took him to be putting forward not a moral thesis but a logical one. They took him to be claiming, as a necessary truth, that there is an essential ‘retrospective’ feature

² Honderich, T., *Punishment. The Supposed Justifications* (London, 1969), p. 9.

³ This seventeenth-century quotation comes from the *Oxford English Dictionary*, s.v. ‘retribution’.

⁴ *Rechtslehre* (1796), Part II, Section 49E.

⁵ Mabbott, J. D., ‘Punishment’, *Mind*, 48 (1939); reprinted in Acton, H. B. (ed.), *The Philosophy of Punishment* (London, 1969), p. 42.

to punishment; punishment, that is, must logically look backward to the past commission of an offence.⁶ Because Mabbott adopted the label ‘retributive’ to describe his theory (though he warned that he was not an ‘orthodox retributivist’), subsequent writers who wished to support this logical thesis tended to characterize their position as— at least partly— retributivist. To be aware of this background is a vital prerequisite for following the argument of section (4) below.

Whatever the truth of the purely logical claim that proper punishment must necessarily involve reference to a past offence, it seems both arbitrary and inappropriate to signal this claim by using the label ‘retributive’. There are many terms in the language of obligation whose correct understanding involves a necessary reference to a past act. ‘Promise’ is one of them. To understand why promises should be kept, it is, it seems to me, crucial to understand that— whatever the future benefits of keeping promises may be— the obligatoriness relates to a past act: the giving of a promise. But in saying this have I advanced a ‘retributive theory of promising’? Hardly! A retrospective or backward-looking element in a theory is never normally characterized as ‘retributive’; it is only in the literature on punishment that these two notions are muddled (so that, in some discussions, ‘retributive’ becomes virtually an antonym of ‘teleological’).

(4) *Minimalism*

‘Most contemporary retributivists’, asserts Martin Golding in his book *Philosophy of Law*, ‘maintain a minimalist position. This holds that no one should be punished unless he is guilty of a crime and culpable.’⁷ Similarly, Anthony Quinton claims ‘the fundamental thesis [of retributivists] must . . . be that only the guilty are to be punished’.⁸ Thesis (4), then, asserts that guilt is a necessary condition of just punishment.

As with our definition (2), some initial puzzlement is in order here. What on earth is supposed to be distinctly retributivist about thesis (4)? The most likely explanation for the classification of (4) as a kind of retributivism seems to be the historical one sketched in the previous section: even though Mabbott’s thesis was principally about the logically sufficient, as opposed to necessary, conditions for just punishment, his adoption of the label ‘retributivist’ paved the way for any theory asserting a logical connection between proper punishment and past offence to be classified as retributivist.

But whatever its historical genesis, the labelling of thesis (4) as a kind of retributive theory has sowed nothing but confusion. For it is clearly possible to hold thesis (4) as a principle of natural justice, or as a utilitarian secondary rule, without having any truck with the notion of retribution in the standard senses outlined under (1) and (2) above. A good example of the awkwardness that can arise in this connection is an otherwise exemplarily clear paper by H. L. A. Hart. Hart mounts a staunch and well constructed defence of thesis (4), but then proceeds to insist on labelling it the principle of ‘retribution-in-distribution’. This in spite of the fact that the route taken by Hart to defend (4) has, as he himself admits, nothing to do with retribution in its ordinary sense, but is concerned instead with principles of simple justice, and in particular the need to

⁶ Cf. Flew, A., ‘The Justification of Punishment’, *Philosophy*, 29 (1954), reprinted in Acton, *op. cit.*

⁷ *Op. cit.*, p. 85.

⁸ Quinton, A., ‘On Punishment’, *Analysis* 14 (1964), reprinted in Acton, *op. cit.*, p. 68.

guarantee 'a method of social control which maximizes individual freedom'.⁹

A further unfortunate consequence of calling thesis (4) 'retributivist', or mislabelling it the principle of 'retribution-in-distribution', has been the widespread view (hinted at by Hart) that some kind of compromise is possible between utilitarian and retributive accounts of punishment. The supposed compromise runs: utilitarian considerations provide the answer to the general question 'why punish at all?', while 'retribution' provides the answer to the more specific question 'whom shall we punish?' (answer: 'only the guilty').¹⁰ To put the matter this way can be highly misleading, since, once again, if thesis (4) is held as a utilitarian secondary rule or as a liberty-maximizing principle, then of course no concession at all has necessarily been made to the strict notion of retribution— i.e., to the notion that punishment is due as repayment for a crime.

(5) *Satisfaction theory*

'A man is rightly punished because his punishment brings satisfaction to others.' This thesis, which I label the satisfaction theory, has been described as a view which 'has given force to retributivism'.¹¹ The idea behind the claim seems to be that there should be some kind of reciprocity between the sense of grievance felt by the victim of an offence and the satisfaction he gets from the suffering of the offender. What connection, if any, does this theory have with our basic sense of retribution? The answer, I think, hinges on which of two possible interpretations is given to the theory.

(5a) On the first interpretation, the claim put forward is that it is intrinsically desirable or appropriate that grievances of victims should be matched by suffering of offenders. There is a close link here with retribution as repayment. If child A hits child B causing him pain and a sense of grievance, child B will frequently be heard to say, "I'll make you pay for that!". The payment is felt to have been exacted once B has inflicted a similar hurt on A. (Such beliefs are by no means confined to children, but children tend to make them more explicit.) Unfortunately for this version of the satisfaction theory, it is far from clear how the indubitable psychological facts just cited are capable of providing a satisfactory moral justification for the practice of punishment (unless we fall into the error commonly attributed to Mill and argue that what is desired is therefore desirable).

(5b) A second, and more sophisticated, version of the satisfaction theory is put forward by Justice Steven: 'the criminal law regulates sanctions and provides a satisfaction for the passion of revenge.'¹² If the underlying idea here is that the penal system provides a substitute for private revenge, then it turns out that the focus of justification does not centre on the notion of retribution at all. Rather, we seem to be dealing with a utilitarian approach, where the penal system is justified as a mechanism for the prevention of vendettas, which furthers the goal of social stability-making society better ordered and more secure.

(6) *Fair play theory*

'Failure to punish is unfair to those who practise self-restraint and respect the rights of

⁹ Hart, H. L. A., *Punishment and Responsibility* (Oxford, 1968), p. 23.

¹⁰ For the slightly different and more subtle distinction between justifying the institution of punishment and justifying an individual act of punishment cf. Rawls, J., 'Two Concepts of Rules', *Phil. Review*, 44 (1966).

¹¹ Honderich, *op. cit.*, p. 31.

¹² Steven; J. F., *A General View of the Criminal Law of England* (London, 1890), p. 99.

others.¹³ '[Retributivism proposes that] a man's punishment is justified or obligatory because . . . unlike non-offenders he has gained satisfactions attendant on the commission of an offence.'¹⁴ These quotations are strongly reminiscent of John Rawls's 'justice-as-fairness' defence of political obligation. The citizen, according to Rawls, has a duty to obey laws of which he disapproves because it is unfair to his fellow citizens if he voluntarily accepts the benefits of the social system without being prepared to shoulder his share of its burdens.¹⁵ The parallel suggestion, in the sphere of punishment, seems to be that the state is justified in punishing the offender on behalf of the good because otherwise he would be gaining an unfair advantage: he would be profiting from his own selfish refusal to play fair and respect the rights of his fellow citizens.

Is it correct to describe the fair-play theory as 'retributive' in our basic sense? Only, I think, indirectly. The object of punishment, on theory (6), is the preservation of justice and the maintenance of fair play. The immediate focus of justification thus centres not on the offender but on the law-abiding citizen and the duties owed to him. It is not that punishing the offender is intrinsically appropriate, but that failure to punish would be unjust to others. It is, however, true that the means actually chosen for upholding fairness is to make the offender 'pay' for the unfair advantage he has obtained (other means of preserving fairness are theoretically available— e.g., special bonuses for the law-abiding); it thus turns out that there is in practice a strong, though indirect, connection between the fair play theory and retribution in the basic sense.

(7) *Placation theory*

In a famous passage in the *Rechtslehre*, Kant says that 'even if a civil society were to dissolve itself by common agreement . . . the last murderer remaining in prison must first be executed so that . . . the blood guilt thereof will not be fixed on the people.'¹⁶ This quotation is often put forward as a paradigm case of retributivist thinking. One commentator cites it triumphantly as showing how, for Kant, 'forward-looking considerations . . . are quite irrelevant to whether punishment ought to be inflicted.'¹⁷ This comment seems to involve the following extraordinary chain of reasoning: "Kant is a retributivist; retributivism is backward-looking (non-teleological); this passage is written by Kant; therefore it contains no reference to forward-looking considerations." Apart from the muddle about 'backward-looking' (see section (3) above), the most striking mistake here is the failure to see the highly teleological flavour of Kant's argument. For Kant is clearly worried about the future: blood guilt will be fixed on the people. The *telos* or goal of punishment is to prevent this consequence by dealing with the murderer.

The logic of justification implied here by Kant seems heavily tied up with the Old Testament notions of sacrifice and placation. Murder involves blood guilt for which an angry God will take vengeance. If the blood of - the guilty is spilled (or perhaps some alternative device— e.g., a scapegoat— is employed) the irate deity will be placated. This theory— the placation theory— is unmistakably consequentialist: it looks forward

¹³ Golding, *op. cit.*, p. 92.

¹⁴ Honderich, *op. cit.*, p. 18.

¹⁵ 'Legal Obligation and the Duty of Fair Play', in Hook, S. (ed.), *Law and Philosophy* (New York, 1964).

¹⁶ 'Selbst wenn sich die bürgerliche Gesellschaft mit aller Glieder Einstimmung auflöstete . . . müsste der letzte im Gefängnis befindliche Mörder vorher hingerichtet werden, damit . . . die Blutschuld nicht auf dem Volke hafte', *loc. cit.*

¹⁷ Golding, *op. cit.*, p. 90.

to the desired result of appeasing the wrath of God. The line of argument involved is unlikely to have much appeal to present-day thinkers concerned to provide a secular justification for punishment. But before leaving the theory it is worth noting that it does have some connection with our 'basic' sense of retributivism. Under (I), I noted that it was obscure how suffering is supposed to 'pay for' wrong done. However, once one grants the existence of a God who desires blood for blood, and who will wreak his vengeance "even to the fourth and fifth generation", then the suffering of the offender (or of a sacrificial victim) makes good sense: punishment avoids the disastrous consequences that would otherwise ensue were the deity's anger to be left unplaced.

(8) Annulment theory

We are obliged to punish because to do so is 'to annul the crime which otherwise would be [held] valid, and to restore the right'; 'the objective treatment of righting the wrong . . . is the primary and fundamental attitude in considering crime.'¹⁸ This celebrated, not to say notorious, theory of Hegel has— in my view rightly— had a pretty bad press from English-speaking philosophers; but they all seem to agree in calling it 'retributive'. However, the label only appears apt given a confusion between criminal retribution and civil restitution. If I smash one of your windows, you may sue me for damages. Once I have been made to pay for replacing the broken pane, the wrong is, literally, put right: things are exactly as they were before I did the damage.¹⁹ This is precisely what Hegel seems to be saying about criminal punishment: it is a means of restitution, it 'restores the right'.

It is, then, a confusion to assimilate the aim of the annulment or restitution theory, as it stands, to that of retribution. However, a kind of combination of theory (8) and our basic theory (1) might be cobbled together along the following lines. I restore the right by paying for your window (restitution); suffering a prison sentence 'pays for' the offence (repayment theory); thus, just as financial payment restores the right in the civil case, so penal 'payment' restores the right in the criminal case. If an (expanded) Hegelian theory were put forward along these lines, it would be appropriate to label it 'retributive' in character (in our basic sense). To say this is, of course, not to defend the logic of justification involved, which to say the least leaves much unexplained. The principal difficulty is that, in the broken window case, things after the restitution are exactly as they were before the wrong was done; how is this supposed to be the case after an execution, or a prison sentence for robbery? (It is nevertheless sometimes said that doing a 'stretch' in prison 'wipes the slate clean'— though this is hardly the view of judges who invariably sentence second offenders to severer punishments.)

(9) Denunciation theory

'The ultimate justification of any punishment is not that it is a deterrent but that it is the emphatic denunciation by the community of a crime.'²⁰ This celebrated dictum of Lord Denning is frequently quoted by commentators in the course of discussions of retributivism,²¹ and his theory is therefore included in our list. But it is doubtful

¹⁸ ' . . . das aufheben des Verbrechens, das sonst gelten würde. und . . . die Wiederherstellung des Rechts'; 'die objective Betrachtung der Gerechtigkeit . . . der erste und substantielle Gesichtspunkt bei dem Verbrechen ist': Hegel, G. F., *Grundlinien der Philosophie des Rechts* (1833), Part I, Section 99.

¹⁹ I do not mean to suggest that all successful suits for restitution have this tidy result.

²⁰ *Report of the Royal Commission on Capital Punishment* (London, 1963), Section 63.

²¹ E.g., Honderich, *op. cit.*, p. 23; Golding, *op. cit.*, p. 87.

whether clarity is served by labelling Denning's thesis 'retributivist'. No doubt exacting retribution could be an emphatic way of denouncing a crime. But one could presumably exact retribution quietly— even secretly— without denouncing. And, conversely, it is evidently logically possible to denounce a crime without exacting retribution (indeed without punishing at all). One suspects that the main reason for calling theory (9) 'retributivist' is the tendency, already noted and condemned above, to call any non-consequentialist account of punishment 'retributivist'.

Concluding remarks

I shall not attempt to summarize the characteristics of the various accounts sketched; and I hope that their respective relationship (or lack of it) to the basic sense of 'retribution' will have been apparent section by section. But the following general points are perhaps worth noting in conclusion.

First, the promised caveat about 'theory'. Philosophers are apt to be lavish in their employment of this high-falutin label; it is not uncommon to find a collection of analytic observations on the concept of, say, pleasure, called a 'theory of pleasure'. If we look at our basic (repayment) sense of 'retribution' it is evident that what we really have is not so much a theory as a metaphor; a metaphor which, I claim, is central to the basic signification of 'retribution', but which, as noted, cuts remarkably little ice as a justificatory device.

Some of the other accounts described can, as I have tried to show, be exhibited in a way which reveals a good deal in common with the central metaphor of repayment: these are (2), (5a), (7), and (8)— desert, simple grievance-satisfaction, annulment and placation. And it is interesting that— with the exception of the placation account— they all turn out to share the inherent thinness of the repayment account. The placation account is more solidly structured, and perhaps alone of those just mentioned merits the accolade 'theory'; but the structure on which it rests incorporates controversial assumptions which are available only to theists (and which nowadays are in most cases probably uncongenial even to them).

Although the accounts just mentioned are logically distinct (commitment to any one need not entail commitment to any of the others) it perhaps does no great harm to regard them as planets revolving around a single sun. But when we come to accounts (3) and (4), Mabbottian penalty theory and minimalism, we need to make a complete break. If the Mabbottian theory is to be construed as a logical thesis about a necessarily backward-looking element in the concept of punishment, then nothing but confusion is served by dubbing holders of this view 'retributivists', since they need have no allegiance to the thesis that repayment and allied notions can be used to justify punishment. If they must have a label, let it be 'retrospectivists'. A similar danger lurks if we suppose that advocates of minimalism (i.e., of Hart's misleadingly dubbed principle of 'retribution-in-distribution') need feel any attraction to the repayment account and its brothers. Of the remaining accounts, (5b) and (9), vendetta-prevention and denunciation, can be disposed of quite briefly. If my remarks on these have been anywhere near right, then the label 'retributivist' is, in these cases, straightforwardly misplaced. The vendetta prevention account, when properly understood, has a utilitarian rationale; and the denunciation view counts as retributive only if retribution is, muddle-headedly, allowed to mean no more than 'non-teleological'.

This leaves the fair play theory (6), which seems to me to have the best chance of providing a non-utilitarian rationale for the practice of punishment. But if this approach is to be characterized as 'retributivist', enormous caution is needed. For, as argued

above, the theory draws its attraction from being derived from very general principles of justice, and seems to be retributive only in an indirect sense.

It has not been the purpose of this paper to propose that the term 'retributive' should henceforth be restricted to the basic (repayment) sense or to senses closely connected with it. Such a proposal would in any case be doomed to failure. It is almost certainly too much to hope that even the most suspect members of the above list will cease to be characterized as retributivist. But something will have been achieved if some writers in future see a mental red light flashing when they are tempted to use the phrase 'the retributive theory of punishment'.

*Department of Philosophy
University of Reading*