# AN APPROACH TO THE PROBLEMS OF PUNISHMENT

S. I. BENN

I SHALL develop, in this article, certain distinctions suggested by recent contributions to the philosophical discussion of punishment, which help to clarify the issues involved. Having separated out what I consider the four central philosophical questions, I shall suggest an approach to them, which, while mainly utilitarian, takes due account, I believe, of the retributivist case where it is strongest, and meets the main retributivist objections.

I make three key distinctions:

(1) Between justifying punishment in general (i.e. as an institution), and justifying particular penal decisions as applications of it;
(2) Between what is implied in postulating guilt as a necessary, and as a sufficient, condition for punishment;
(3) Between postulating guilt in law and guilt in morals, as a condition for punishment.

I distinguish, further, four philosophical questions, to which a complete and coherent approach to punishment would have to provide answers:

What formal criteria must be satisfied in justifying:—

(1) Punishment in general, i.e. as an institution?
(2) Any particular operation of the institution?
(3) The degrees of punishment attached to different classes of offence?
(4) The particular penalty awarded to a given offender?

## *Preliminaries*

### A. *"Punishment" defined*

Prof. Flew[1] has suggested five criteria for the use of "punishment" in its primary sense, i.e. five conditions satisfied by a standard case to which the word would be applied:

(i) It must involve an "evil, an unpleasantness, to the victim";
(ii) It must be for an offence (actual or supposed);
(iii) It must be of an offender (actual or supposed);
(iv) It must be the work of personal agencies (i.e. not merely the natural consequences of an action);

[1] A. Flew: "The Justification of Punishment," in PHILOSOPHY, Vol. XXIX, 1954, pp. 291–307.



https://doi.org/10.1017/S0031819100055017 Published online by Cambridge University Press

(v) It must be imposed by authority (real or supposed), conferred by the system of rules (hereafter referred to as "law") against which the offence has been committed.

It is not a misuse to talk, for example, of "punishing the innocent", or of a boxer "punishing his opponent"; but since these usages, though related to the primary one, disregard one or more of the criteria ordinarily satisfied, they are extensions, or secondary usages. In considering the justification for punishment, I shall confine the word to the primary sense, unless I indicate otherwise.

B. *The distinction between justifying punishment in general, and justifying the particular application*

There would seem, on the face of it, to be a real difference between utilitarian and retributivist approaches to the justification of punishment, the former looking to its beneficent consequences, the latter exclusively to the wrongful act. It remains to be seen whether the gulf can be bridged. The first step is to distinguish between a rule, or an institution constituted by rules, and some particular application thereof. To ask what can justify punishment in general, is to ask why we should have the sort of rules that provide that those who contravene them should be made to suffer; and this is different from asking for a justification of a particular application of them, in punishing a given individual. Retributivist and utilitarian have tried to furnish answers to both questions, each in his own terms; the strength of the former's case rests on his answer to the second, of the latter's on his answer to the first. Their difficulties arise from attempting to make one answer do for both.

I. *What formal criteria must be satisfied in justifying punishment in general, as an institution?*

The retributivist refusal to look to consequences for justification makes it impossible to answer this question within his terms. Appeals to authority apart, we can provide ultimate justification for rules and institutions, only by showing that they yield advantages.[1]

[1] Admittedly, a rule might be justified *in the first place* by reference to one more general, under which it is subsumed as a particular application—e.g. "It is wrong to pick flowers from public gardens because it is wrong to steal—and this is a special case of stealing". But this would not be conclusive. It could be countered by making a distinction between private and public property, such that while the more general rule prohibits stealing the former, it does not extend to the latter. Whether the distinction can be accepted as relevant must depend on the reasons for the more general rule, understood in terms of its expected advantages, and on whether to allow the exception would tend to defeat them. Consider "Euthanasia is wrong because it is wrong to kill". It could be argued that the latter does not require the former; that a proper distinction can be made between killings generally, and those satisfying the conditions: i.

Consequently, what pass for retributivist justifications of punishment in general, can be shown to be either denials of the need to justify it, or mere reiterations of the principle to be justified, or disguised utilitarianism.

Assertions of the type "it is fitting (or justice requires) that the guilty suffer" only reiterate the principle to be justified—for "it is fitting" means only that it ought to be the case, which is precisely the point at issue. Similarly, since justification must be in terms of something other than the thing in question, to say that punishment is a good in itself is to deny the need for justification. For those who feel the need, this is no answer at all. Given that punishment would not be justified for the breach of *any* rule, but only of legal rules, what is the peculiar virtue of law that makes it particularly fitting for breaches of just this type of rule? Even if we make punishment a definitional characteristic of "a legal system", so that "law" entails "punishment", we are still entitled to ask why we should have rule systems of precisely this sort.

Some retributivists argue that while punishment is a prima facie evil, and thus in need of justification, it is less objectionable than that the wicked should prosper. This is to subsume the rule "Crimes ought to be punished" under a more general rule: either "The wicked ought to be less well off than the virtuous" or "The wicked ought not to profit from their crimes". Now "wickedness" involves assessment of character; we do not punish men for their wickedness, but for particular breaches of law. There may be some ignoble but prudent characters who have never broken a law, and never been punished, and noble ones who have—our system of punishment is not necessarily the worse for that. We may have to answer for our characters on the Day of Judgment, but not at Quarter Sessions. The state is not an agent of cosmic justice; it punishes only such acts as are contrary to legal rules, conformity to which, even from unworthy motives like fear, is considered of public importance. And if we offer the narrower ground, that the wicked ought not to profit from their *crimes*, we are bound to justify the distinction between crimes and offences against morals in general. What is the special virtue of legal rules, that a

that the patient wants to be killed; ii. that the purpose is to put him out of pain; iii. that there is no hope for his recovery. Suppose the reason for the general prohibition is to ensure that the life of man shall not be "solitary, poor, nasty, brutish, and short"; then exceptions satisfying the above criteria might be admissible, on the grounds that not only would they not defeat the objectives of the rule, but that advantages would follow from distinguishing on the basis of these criteria, that would otherwise be missed. On the other hand, it might be said that it is *absolutely* wrong to kill—which is to deny the need for justification in terms of purpose or consequences, but is also to deny the need for *any* moral (as opposed to authoritative) justification. But in that case, how are we to decide whether "Thou shalt not kill" does, or does not, extend to a duty "officiously to keep alive"?

breach of them alone warrants punishment? It seems that the wicked are to be prevented from prospering only if their wickedness manifests itself in selected ways; but how is the selection made, unless in terms of its consequences? In any case, if we permit the subsumption of "Crime ought to be punished" under the more general "The wicked ought not to prosper", it would still be proper to seek justification for the latter. It would not help to say "Justice requires it", for this would only deny the right to ask for justification. I see no answer possible except that in a universe in which the wicked prospered, there would be no inducement to virtue. The subsumption, if allowed, would defer the utilitarian stage of justification; it would not render it superfluous.

A veiled utilitarianism underlies Hegel's treatment of punishment, as anulling a wrong. For if punishment could annul the wrong, it would be justified by the betterment of the victim of the crime or of society in general. Not indeed that the argument is a good one; for the only way to annul a wrong is by restitution or compensation, and neither of these is punishment. A man may be sent to prison for assault, and *also* be liable for damages. Similarly with the argument that punishment reaffirms the right. Why should a reaffirmation of right take precisely the form of punishment? Would not a formal declaration suffice? And even if the reaffirmation necessarily involved a need, right, or duty to punish, the justification would be utilitarian, for why should it be necessary to reaffirm the right, if not to uphold law for the general advantage?[1]

Others have treated punishment as a sort of reflex, a reaction of the social order to the crime following in the nature of things, like a hangover.[2] This is to confuse rules with scientific laws. The penal consequences of a breach of a rule follow only because men have decided to have rules of precisely this sort. Laws of nature, unlike rules, need no justification (except perhaps in theology) because they are independent of human choice. To treat punishment as a natural unwilled response to a breach of law is to deny the need for justification, not to justify.[3] Once we agree to have penal rules, any particular punish-

[1] Cf. Lord Justice Denning, in evidence to the Royal Commision on Capital Punishment: "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." Cmd. 8932, §53 (1953). But "denunciation" does not imply the deliberate imposition of suffering, which is the feature of punishment usually felt to need justification.

[2] Cf. Sir Ernest Barker, in *Principles of Social and Political Theory*, p. 182: "the mental rule of law which pays back a violation of itself by a violent return, much as the natural rules of health pay back a violation of themselves by a violent return."

[3] For J. D. Mabbott, too, punishment is a kind of automatic response, though in a different sense. "Punishment is a corollary not of law but of law-breaking.

ment might be justified (though not necessarily sufficiently justified) by reference to a rule. But this is to answer a different question from that at present under consideration.

For Bosanquet, punishment was retributive in the sense that, ideally at least, it was the returning upon the offender of "his own will, implied in the maintenance of a system to which he is a party", in the form of pain. It tends to "a recognition of the end by the person punished"; it is "his right, of which he must not be defrauded".[1] Now while a criminal may not seek to destroy the entire social order, and may even agree in principle that law-breakers should be punished, his efforts to elude the police are evidence that he does not will his own punishment in any ordinary sense. He may be unreasonable and immoral in making exceptions in his own favour—but we cannot therefore construct a theory of punishment on a hypothetical will that would be his were he reasonable and moral, for then he might not be a criminal. To say that punishment is his "right" is to disregard one of the usual criteria for the use of that word, namely, that it is something which will be enforced only if its subject so chooses, the corollary being that it operates to his advantage. Only by pretending that punishment is self-imposed can we think of the criminal as exercising choice; and only by treating it as reformative can we regard it as to his advantage. By claiming that punishment tends "to a recognition of the end by the person punished", Bosanquet introduces such a reformative justification; but to that extent the argument is utilitarian.

To sum up: retributive justifications of punishment in general are unsatisfactory for the very reason that they refuse to look to the consequences of a rule, thereby denying a necessary part of the procedure for justifying it. To look to the consequences does not entail treating the criminal merely as a means to a social end, as critics have asserted; for in weighing advantages and disadvantages, the criminal, too, must "count for one". But equally, he must count "for no more than one". While we must not lose sight of his welfare altogether, we are not bound to treat him as our sole legitimate concern.

Bentham's case is that punishment is a technique of social control, justified so long as it prevents more mischief than it produces. At the point where damage to criminals outweighs the expected advantage to society, it loses that justification. It operates by reforming

Legislators do not *choose* to punish. They hope no punishment will be needed. The criminal makes the essential choice; he 'brings it on himself'." ("Punishment", in *Mind*, Vol. 48, 1939, p. 161. He reaffirms the position in "Freewill and Punishment", in *Contemporary British Philosophy*, 3rd Series, ed. H. D. Lewis, 1956, p. 303.) But legislators choose to make *penal* rules, and it is this choice that needs justification.

[1] *The Philosophical Theory of the State*, 4th edn. (1923), p. 211.

the criminal, by preventing a repetition of the offence, and by deterring others from imitating it. (These need not exhaust the possibilities of advantage—Bentham included the satisfaction of vengeance for the injured party.)

Not all theories dealing with the reform of criminals are theories of punishment. Prison reformers concerned with moral re-education offer theories of punishment only if they expect the suffering involved in loss of liberty, etc., itself to lead to reformation. Reformative treatment might cure criminal inclinations by relaxing the rigours of punishment; it might nevertheless defeat its purpose by reducing the deterrent effect for others. "Reformation" is in any case ambiguous. A man would be "a reformed character" only if he showed remorse for his past misdeeds, and determined not to repeat them, not through fear of further punishment, but simply because they were wrong. A criminal who decides that "crime does not pay" is merely deterred by his own experience which is as much "an example" to himself as to others.

Sentences of preventive detention, transportation, deportation, and the death penalty, may all be examples of punishment operating as a preventive. Punishment might be aimed at preventing repetitions of an offence by the criminal himself where there are good grounds (e.g. a long criminal record) for supposing him undeterrable.

The strongest utilitarian argument for punishment in general is that it serves to deter potential offenders by inflicting suffering on actual ones. On this view, punishment is not the main thing; the technique works by threat. Every act of punishment is to that extent an admission of failure; we punish only that the technique may retain a limited effectiveness for the future. Thus the problem of justifying punishment arises only because it is not completely effective; if it were, there would be no suffering to justify.

Retributivists do not deny that punishment may act in these ways, nor that it has these advantages. They maintain only that they are incidental; that a system of punishment constructed entirely on these principles would lead to monstrous injustices. These I consider below. It is evident, however, that while *some* sort of justification can be offered within the utilitarian framework, the retributivist is at best denying the need for justification, or offering utilitarianism in disguise. I conclude, therefore, that any justification for punishment in general must satisfy the formal condition that the consequences for everyone concerned of adopting the technique shall be preferable to the consequences of not doing so. If the main advantage arises from a lower incidence of crime (by way of reform, prevention, deterrence, or otherwise), this must be weighed against the penal suffering actually inflicted, and these together must be preferable to a higher incidence of crime, but with no additional suffering inflicted as

punishment. This is a frankly utilitarian conclusion. The strength of the retributivist position lies in its answer to the second question, to which I now turn.

II. *What formal criteria must be satisfied in justifying any particular application of the technique of punishment?*

Critics of the utilitarian approach contend that a justification of punishment in terms of deterrence, prevention, and reform could be extended to justify (i) punishing the innocent, providing they were widely believed to be guilty (in the interest of deterrence); (ii) making a show of punishment, without actually inflicting it (again, deterrence, but this time on the cheap); (iii) punishment in anticipation of the offence (in the interests of prevention or reform). These criticisms, if just, would surely be conclusive. They are based, however, on a misconception of what the utilitarian theory is about. "Punishment" implies, in its primary sense, inflicting suffering only under specified conditions, of which one is that it must be for a breach of a rule. Now if we insist on this criterion for the word, "punishment of the innocent" is a logical impossibility, for by definition, suffering inflicted on the innocent, or in anticipation of a breach of the rule, cannot be "punishment". It is not a question of what is morally justified, but of what is logically possible. (An analogous relation between "guilt" and "pardon" accounts for the oddity of granting "a free pardon" to a convicted man, later found to be innocent.) When we speak of "punishing the innocent", we may mean: (i) "pretending to punish", in the sense of manufacturing evidence, or otherwise imputing guilt, while knowing a man to be innocent. This would be to treat him *as if* he were guilty, and involve the lying assertion that he was. It is objectionable, not only as a lie, but also because it involves treating an innocent person differently from others without justification, or for an irrelevant reason, the reason offered being falsely grounded.[1] (ii) We may mean, by "punish", simply "cause to suffer" i.e. guilt may not be imputed. This would be a secondary use of the word. In that case, it could not be said that, as a matter of logical necessity, it is either impossible or wrong to punish the innocent. To imprison members of a subversive party (e.g. under Defence Regulation 18B) treating them *in that respect* like criminals, though no offence is even charged, would not necessarily be immoral. Critics might describe it as "punishing the innocent", but they would be illegitimately borrowing implications of the primary sense to attack a type of action to which these did not apply. It is only necessarily improper to "punish the innocent" if we pretend they are guilty, i.e. if we accept all the primary usage criteria; in any looser sense,

[1] Cf. A. Quinton: "On Punishment," in *Analysis*, Vol. 14, reprinted in *Philosophy, Politics, and Society*, ed. P. Laslett, 1956.

there need be nothing wrong in any given case. For in exceptional conditions it may be legitimate to deprive people of their liberty as part of a control technique, without reference to an offence (e.g. the detention of lunatics or enemy aliens). Similar arguments apply in the case of the show of punishment. A utilitarian justification of punishment cannot be extended to cover lies, or the making of distinctions where there are no relevant differences; it would be impossible merely to pretend to punish *every* criminal—and unless a relevant criterion could be found, there could be no grounds for treating some differently from others.

The short answer to the critics of utilitarian theories of punishment, is that they are theories of *punishment*, not of *any* sort of technique involving suffering.

We may now turn to the retributivist position itself. F. H. Bradley asserted "the necessary connection of punishment and guilt. Punishment is punishment, only where it is deserved . . . if punishment is inflicted for any other reason whatever than because it is merited by wrong, it is a gross immorality, a crying injustice, an abominable crime, and not what it pretends to be."[1] Now, we must distinguish between legal and moral guilt. If the necessary connection asserted is between punishment and legal guilt, then this is a definition of "punishment" masquerading as a moral judgment. It would be more accurate to write "Punishment is 'punishment', only when it is deserved", for the sentence is then about the use of a word, not about the rightness of the act. "The infliction of suffering on a person is only properly described as punishment if that person is guilty. The retributivist thesis, therefore, is not a moral doctrine, but an account of the meaning of the word 'punishment'."[2]

But this is not the only form of retributive thesis. There are at least four possibilities:

(i) That guilt (i.e. a breach of law) is a *necessary* condition of punishment (this is the position just examined);

(ii) That guilt (i.e. a breach of a *moral* rule) is a *necessary* condition of punishment;

(iii) That guilt (*legal*) is a *sufficient* condition of punishment;

(iv) That guilt (*moral*) is a *sufficient* condition of punishment.

Position (iii) is *not* logically necessary, for it does not follow from the definition of punishment; we *cannot* "punish" where there has been no breach, but we can, and often do, let off with a caution where there has. Other conditions besides guilt may have to be satisfied before punishment is wholly justified in a given case.

[1] *Ethical Studies*, 2nd edn. 1927, pp. 26–7.

[2] A. Quinton: op. cit., in *Analysis*, p. 137, in *Philosophy, Politics, and Society*, p. 86.

The introduction, in (ii) and (iv), of moral guilt puts a new complexion on retributive theory. A person who is morally guilty deserves blame, and the conditions for blameworthiness could be listed. But it is in no sense necessary that a person who is blameworthy should also be punishable. We may blame liars, but unless, e.g., they make false tax returns, or lie to a court of law, we should not feel bound to punish them. If the conditions of blameworthiness cannot be assimilated completely to the conditions for punishment, moral guilt cannot be a sufficient condition for punishment.

Position (ii) might be supported in two ways:

(*a*) A prima facie moral duty to obey law may yield, in the case of an immoral law, to a stronger duty; a breach of law would not then entail moral guilt, and we should question the justice of the punishment.[1]

(*b*) Certain conditions, like unavoidable ignorance or mistake of fact, lunacy, infancy, and irresistible duress, would exonerate from blame; offences committed under these conditions should not be punishable—and are not in fact punished, though the deterrent effects of the punishment would be no less in these cases than in others. Therefore, in a negative sense at least, the criteria of blameworthiness must be satisfied, if the necessary conditions for punishment are to be satisfied. Punishment is retribution for such moral lapses as the law recognizes.

The first argument (*a*) might be met in two ways. From the judge's standpoint, so long as he continued in office, it would be his duty to enforce the law, whatever his opinion of it.[2] For him, at least, the absence of moral guilt would not be a bar to punishment. Secondly, criticism of a rule is only indirectly criticism of the justice of a punishment inflicted for a breach of it. The utilitarian could argue that a law that is itself mischievous (in Bentham's sense of "mischief") cannot justify the further mischief of punishment; no good can come of it any way. This is not, therefore, a defence of a retributive theory of punishment, so much as a statement of conditions that a rule must satisfy if punishment is properly to attach to it.

The second argument (*b*) is inconclusive. If the technique of punishment operates primarily by deterrence, it can serve its purpose only in respect of deliberate acts. No act committed under any of the above conditions would be deliberate. If, therefore, offences of these

[1] Cf. C. W. K. Mundle: "Punishment and Desert," in *Philosophical Quarterly*, Vol. 4, 1954: "the retributive theory implies that punishment of a person by the state is morally justifiable if, and only if he has done something which is both a legal and moral offence, and only if the penalty is proportionate to the moral gravity of his offence", p. 227.

[2] This is roughly Mabbott's view (op. cit.). He is a rare example of a retributivist who dissociates punishment and moral guilt.

types are left unpunished, the threat in relation to other offences remains unimpaired, for the sane potential murderer gets no comfort from mercy extended to the homicidal maniac, and other homicidal maniacs will be unaffected either way. Consequently to punish in such cases would be a pointless mischief. In any case, because some of the conditions for blame and punishment coincide, it does not follow that the satisfaction of the former is a necessary condition for the satisfaction of the latter.[1] I shall return to this point later in relation to motive.

Of the four possible interpretations of the retributivist relation of guilt to punishment, it is the first only, whereby guilt in law is a necessary condition for punishment, that is completely persuasive; and this is precisely because it is a definition and not a justification. Consequently, it need not conflict with a utilitarian view.

For a utilitarian to require, for every case of punishment, that it be justified in terms of preventing more mischief than it causes, would be to miss the point of punishment as an institution. Indeed, any rule would be pointless if every decision still required to be justified in the light of its expected consequences. But this is particularly true of penal rules; for the effectiveness of punishment as a deterrent depends on its regular application, save under conditions sufficiently well understood for them not to constitute a source of uncertainty. Legal guilt once established, then, the initial utilitarian presumption against causing deliberate suffering has been overcome, and a case for a penalty has been made out. But it may still be defeated; for since guilt is not a sufficient condition, there may well be other relevant considerations (e.g. that this is a first offence). The following formal criterion may be postulated, however, which any such consideration must satisfy, namely, that to recognize it as a general ground for waiving the penalty would not involve an otherwise avoidable mischief to society *greater* than the mischief of punishing the offender.

One of the criticisms levelled against utilitarianism is that by relating the justification of punishment to its expected consequences, rather than to the crime itself, it would justify penalties divorced from the relative seriousness of crimes, permitting severe penalties for trivial offences, if that were the only way to reduce their number.

[1] A man who had broken a law (say, an import regulation), of the existence of which he was ignorant (but avoidably so), would be liable to punishment. It would be to counsel perfection to say that everyone has a moral duty to know of *every* law that might affect him. I should say, in this case, that the offender had been imprudent, but not immoral, in not ascertaining his legal position. I should impute no moral guilt either for his ignorance or for his breach of the rule; but I should not feel, on that account, that he was an injured innocent entitled to complain that he had been wrongly punished.

A serious but easily detected crime might warrant lesser penalties than a minor but secret one. This conclusion being intolerable, the retributivist contends that to escape it we must seek the measure of the penalty in the crime itself, according to the degree of wickedness involved in committing it.

Again, I distinguish the justification of rules from the justification of particular applications. To ask "How much punishment is appropriate to a given offence?" is ambiguous: it may refer either to the punishment allotted by a rule to a *class* of acts, or to a particular award for a given act, within that class. The distinction is pointed by the practice of laying down only maximum (and sometimes minimum) penalties in the rule, leaving particular determinations to judicial discretion.

### III. *What formal criteria must be satisfied in justifying the degrees of punishment attached to different classes of offence?*

"The only case" (said Kant) "in which the offender cannot complain that he is being treated unjustly is if his crime recoils upon himself and he suffers what he has inflicted on another, if not in a literal sense, at any rate according to the spirit of the law." "It is only *the right of requital* (*jus talionis*) which can fix definitely the quality and the quantity of the punishment." This is the most extreme retributive position; its essential weakness is present, however, in more moderate attempts to seek the determinants of punishment exclusively in the offence itself.

If retaliatory punishment is not to be effected "in a literal sense" (which might well be intolerably cruel, and in some cases physically impossible), but rather "according to the spirit of the law", it involves a sort of arithmetical equation of suffering as impracticable as the hedonistic calculus. Suffering of one sort cannot be *equated* with another, though it may be possible to prefer one to another (or to be indifferent as between one and another). I can certainly say that I would rather see A suffer in one way, than B in another, or that there is really nothing to choose between the two. But this is quite different from saying that A ought to be made to suffer in exactly the same degree as B, whom he has injured; for this involves not a preference enunciated by some third person, but a quasi-quantitative comparison of the sufferings of two different people, treated as objective facts. And there is no way of making this comparison, even though the external features of their suffering may be identical. It is even more evidently impossible when the suffering of one is occasioned by, say, blackmail, and of the other by imprisonment.[1]

[1] Hegel virtually admits the impossibility of answering this question rationally (*Philosophy of Right*, § 101) but insists nevertheless that there must

The difficulty remains in the compromise between a utilitarian and retaliatory position attempted by W. D. Ross. While admitting that the legislator must consider the deterrent ends of punishment in assessing penalties, he maintains that the injury inflicted by the criminal sets an upper limit to the injury that can legitimately be inflicted on him. "For he has lost his prima facie rights to life, liberty, or property, only in so far as these rested on an explicit or implicit undertaking to respect the corresponding rights in others, and in so far as he has failed to respect those rights."[1] But how are we to make this equation between the rights invaded and consequently sacrificed, and the amount of suffering so justified—unless there is already available a scale, or rule, fixing the relation? But then how is the scale to be justified?

J. D. Mabbott admits there can be no direct relation between offence and penalty, but seeks, by comparing one crime with another, to make an estimate of the penalties *relatively* appropriate. "We can grade crimes in a rough scale and penalties in a rough scale, and keep our heaviest penalties for what are socially the most serious wrongs regardless of whether these penalties . . . are exactly what deterrence would require."[2] But what are we to understand by "socially the most serious wrongs"? On the one hand, they might be those that shock us most deeply—we could then construct a shock scale, and punish accordingly. There are some shocking acts, however, that we should not want to punish at all (e.g. some sexual offences against morality); at the same time, we should be hard put to it to know what penalties to attach to new offences against, say, currency control regulations, where the initial shock reaction is either negligible, because the rule is unsupported by a specific rule of conventional morality, or where it is of a standard mild variety accompanying any offence against the law as such, irrespective of its particular quality. On the other hand, "the most serious wrongs" may be simply those we are least ready to tolerate. That, however, would be to introduce utilitarian considerations into our criteria of "seriousness". For to say that we are not prepared to tolerate an offence is to say that we should feel justified in imposing heavy penalties to deter people from committing it. But in making deterrent considerations secondary to the degree of "seriousness", Mr. Mabbott implicitly excludes this interpretation.

The retributivists' difficulties arise from seeking the measure of the penalty in the crime, without first assuming a scale or a rule relating the two. Given the scale, any given penalty would require justification

be a right answer (§ 214) to which we must try empirically to approximate. But by what test shall we judge whether our shots at justice are approaching or receding from the target?

[1] *The Right and the Good*, 1930, pp. 62–3. [2] Op. cit., p. 162.

in terms of it; but the scale itself, like any rule, must in the end be justified in utilitarian terms. It remains to be seen whether this necessarily opens the way to severe penalties for trivial offences.

For the utilitarian, arguing in deterrent terms, it is the threat rather than the punishment itself which is primary. Could we rely on the threat being completely effective, there could be no objection to the death penalty for every offence, since *ex hypothesi* it would never be inflicted. Unhappily, we must reckon to inflict some penalties, for there will always be some offenders, no matter what the threatened punishment. We must suppose, then, for every class of crime, a scale of possible penalties, to each of which corresponds a probable number of offences, and therefore of occasions for punishment, the number probably diminishing as the severity increases. Ultimately, however, we should almost certainly arrive at a hard core of undeterrables. We should then choose, for each class of offence, that penalty at which the marginal increment of mischief inflicted on offenders would be just preferable to the extra mischief from which the community is protected by this increment of punishment. To inflict any heavier penalty would do more harm than it would prevent. (This is Bentham's principle of "frugality".)[1]

This involves not a quasi-quantitative comparison of suffering by the community and the offender, but only a preference. We might say something like this: To increase the penalty for parking offences to life imprisonment would reduce congestion on the roads; nevertheless the inconvenience of a large number of offences would not be serious enough to justify disregarding in so great a measure, the prima facie case for liberty, even of a very few offenders. With blackmail, or murder, the possibility of averting further instances defeats to a far greater extent the claims of the offender. One parking offence more or less is not of great moment; one murder more or less is.

In retaliatory theory we are asked to estimate the damage done by the crime, and to inflict just that amount (or no more than that amount) on the criminal; here we are required only to choose between one combination of circumstances and another. The choice may not always be easy; but it is not impossible, or even unusual. For we are well accustomed to choosing between things incapable of quantitative comparison; what is impossible is to assess what one man has suffered from blackmail, and then to impose its equivalent on the blackmailer in terms of a prison sentence. The difference is between a prescription and a description. To say, as I do above, that the right penalty is that at which the marginal increment of mischief inflicted is just preferable to the mischief thereby avoided, is to invite the critic to choose (or prescribe) one state of affairs rather than another. But to say that the penalty should equal (or should not exceed) the suffering of the

[1] *Introduction to the Principles of Morals and Legislation*, Chap. XV, §§ 11–12.

victim of the crime is to invite him to prescribe a course dependent not on his own preferences, but on a factual comparison of incomparables, on an equation of objective conditions.

The utilitarian case as I have now put it is not open to the objection that it would justify serious penalties for trivial offences. For to call an offence "trivial" is to say that we care less if this one is committed than if others are, i.e. we should be unwilling to inflict so much suffering to prevent it, as to prevent others. "Relatively serious crimes" are those relatively less tolerable, i.e. we prefer to inflict severer penalties rather than to suffer additional offences. If this is so, "Trivial crimes do not deserve severe penalties" is analytic; consequently a utilitarian justification could not be extended to cover a contrary principle.[1]

Some penalties we are unwilling to inflict whatever their deterrent force. We are less ready to torture offenders than to suffer their offences. And there are people who would rather risk murders than inflict the death penalty, even supposing it to be "the unique deterrent". To kill, they say, is absolutely wrong. Now this may mean only that no circumstances are imaginable in which its probable consequences would make it right, i.e. in which the mischief done would not outweigh the mischief prevented—not that it could *never* be right, only that in any imaginable conditions it would not be. This would not exclude justification by consequences, and is therefore compatible with the view of punishment I am advancing. On the other hand, if the absolutist denies altogether the relevance of consequences, he is making an ultimate judgment for which, in the nature of the case, justification can be neither sought nor offered, and which is therefore undiscussible.

I conclude, from this discussion, that any justification for the nature and degree of punishment attached to a given class of offence must satisfy the following formal criterion: that the marginal increment of mischief inflicted should be preferable to the mischief avoided by fixing that penalty rather than one slightly lower. Assuming that the advantages of punishment derive mainly from upholding rules, this means that the conformity secured, weighed against the suffering inflicted, should be preferable to a lower level of conformity, weighed against the suffering inflicted by imposing a lesser penalty. (This entails neither that a very few offenders suffering

[1] We could say "Some trivial crimes deserve serious penalties" if we wished to imply that some crimes are a good deal more serious than they are generally held to be. But the sentence would be better punctuated: "Some 'trivial' crimes . . . ", for they are "trivial" in the view of others, not of the speaker. Consider, in this connection, the difference of opinion between pedestrians' and motorists' associations on the gravity of driving offences—and on the penalties appropriate. A pedestrian might not think a prison sentence too severe a penalty for speeding—but neither is it, for him, a trivial offence.

heavy penalties must be preferred to a larger number of offenders suffering lighter penalties, nor the converse; preferences are not settled by multiplication.)

IV. *What formal criteria must be satisfied in justifying the particular penalty awarded to a given offender?*

Two men guilty of what is technically the same offence (i.e. who have broken the same rule) are not necessarily punished alike. This could be justified only by reference to relevant criteria, other than simple guilt, by which their cases are distinguished. Provocation, temptation, duress, and a clean record may all make a difference. But these are also relevant to the determination of blame. From these considerations arise two possible objections to the view I am advancing:

(*a*) Is it consistent with utilitarianism that in determining the sentence, we should look to the particular conditions of the crime, rather than to the consequences of the penalty? Should we not look forward to the exemplary advantages of the maximum penalty, rather than backward to extenuating circumstances?

(*b*) Since we do look backwards, and assess the penalty in the light of criteria also relevant to an assessment of blameworthiness, can we not say that men deserve punishment only in the measure that they deserve blame?

As to (*a*); a rule once accepted, there is no need to justify every application in terms of its consequences; it is necessary to justify in utilitarian terms only the criteria of extenuation, not every application of them. Now precisely because an offence has been committed under exceptional circumstances (e.g. severe temptation, provocation, duress), leniency would not seriously weaken the threat, since offenders would expect similar leniency only in similar circumstances, which are such, in any case, that a man would be unlikely to consider rationally the penal consequences of his act. Given that, the full measure of the penalty would be unjustifiable.[1]

As to (*b*); while some criteria tend to mitigate both blame and punishment, the latter need not depend on the degree of the former. The question of motive is crucial. We generally regard a man as less blameworthy if he breaks a rule "from the highest motives", rather than selfishly or maliciously. A traitor from conscientious conviction may be blamed for wrongheadedness, but, if we respect his integrity,

[1] Grading sentences according to the number of previous convictions might be justified by the failure, *ex hypothesi*, of lesser penalties on earlier occasions, to act as deterrents. Possible imitators with similar records may possibly require a similarly severe deterrent example. For most of the rest of us, with little criminal experience, lighter penalties awarded to less hardened offenders are sufficient deterrents. A case can therefore be made for reserving the severest penalties for the class of criminals least easily deterred.

we blame him less than a merely mercenary one. But honest motives will not always mitigate punishment. It may be vital for the effectiveness of government that conscientious recalcitrants (e.g. potential fifth columnists acting from political conviction) be deterred from action. But since strong moral convictions are often less amenable to threats than other motives, they could scarcely be admitted in such cases in extenuation of punishment. On the other hand, if the mischief of the penalty needed for a high degree of conformity exceeds its advantages, it may be reasonable to give up punishing conscientious offenders altogether, provided they can be discerned from the fakes.[1] We no longer punish conscientious objectors to military service, having found by experience that they are rarely amenable to threats, that they are unsatisfactory soldiers if coerced, and that, given a rigorous test of conscientiousness, their numbers are not likely to be so great as to impair the community purpose.

The considerable overlapping of the factors tending to mitigate blame and punishment nevertheless demands explanation. Morality and law are alike rule systems for controlling behaviour, and what blame is to one, punishment is to the other. Since they are closely analogous as techniques for discouraging undesirable conduct, by making its consequences in different ways disagreeable, the principles for awarding them largely coincide. But it does not follow that because we usually also blame the man we punish, we should punish in the light of the criteria determining moral guilt. Morality operates as a control not only by prescribing or prohibiting acts, but also by conditioning character (and therefore conduct in general). We blame men for being bad tempered; we punish them only for assault. Furthermore, punishment is administered through formal machinery of investigation, proof, conviction, sentence, and execution; blame by informal and personal procedures which may well take account of evidence of character that might nevertheless be rightly inadmissible in a court of law. To the extent that the techniques are analogous, they may be expected to employ similar criteria; but the analogy cannot be pushed all the way.

[1] Consider, in this connection, the difficulty of distinguishing the genuine survivor of a suicide pact, who has been unable to carry out his side of the bargain, from the cheat who relies on a counterfeit pact to evade the maximum penalty for murder. (See the Report on Capital Punishment, referred to above, §§ 163–176.) The same applies to "mercy-killing": "How, for example, were the jury to decide whether a daughter had killed her invalid father from compassion, from a desire for material gain, from a natural wish to bring to an end a trying period of her life, or from a combination of motives?" (Ibid., § 179). Nevertheless, where we feel reasonably sure that the motive was merciful, we expect leniency. A mercy-killing is not in the same class as a brutal murder for profit, and we may feel justified in tolerating a few examples rather than inflict the maximum penalty on this type of offender.

## *Conclusion*

The quarrel between retributivist and utilitarian is primarily about procedures of justification, about how to go about defending or attacking punishment, in general or in particular, about the formal criteria that together form a schema to which any justification must conform. I have maintained that when what is wanted is a justification of a rule, or an institution, of punishment in general, or of the scale of punishments assigned to different classes of offence, it must be sought in terms of the net advantages gained or mischiefs avoided. When the particular sentence is in question, the first consideration is guilt, without which punishment in a strict sense is impossible, but which once established constitutes a prima facie case for it. The second consideration must be the legally prescribed limits, within which the penalty must fall. Beyond that, decision must be made in the light of criteria tending to mitigate if not totally defeat the presumption in favour of the maximum penalty. These criteria must themselves be justified in terms of the net advantages, or mischief avoided, in adopting them as general principles.

These are formal principles only. To make out a substantial justification, we must postulate first the sort of advantages we expect from punishment as an institution. I have assumed that its principal advantage is that it secures conformity to rules (though others might conceivably be offered, e.g. that it reformed criminal characters, which could be regarded as a good thing in itself; or that it gave the injured person the satisfaction of being revenged). Further, I have assumed that it operates primarily by way of deterrence. These are in part assumptions of fact, in part moral judgments. I maintain that these being given, the criteria by which the prima facie case for punishment may be defeated, wholly or in part, are generally justifiable in utilitarian terms; that they do not weaken the deterrent threat, that they avoid inflicting suffering which would not be justified by the resultant additional degree of conformity. Further, the total assimilation to the system of punishment of criteria tending to defeat or mitigate blameworthiness, is unjustifiable in theory and is not made in practice. We do not punish men because they are morally guilty, nor must we *necessarily* refrain because they are are morally guiltless, nor mitigate the punishment in the same degree for all the same reasons that we mitigate blame. This is not to say that the justifications sought are not *moral* justifications; it is simply that they must be made in the light of criteria different from those governing blame, since however close the analogy may be between the two techniques of control, there are still significant differences between them.

*The University, Southampton.*



https://doi.org/10.1017/S0031819100055017 Published online by Cambridge University Press