PRINCIPLED SENTENCING

READINGS ON THEORY AND POLICY

Third Edition

Edited by

Andrew von Hirsch
Andrew Ashworth
and
Julian Roberts

• HART •
PUBLISHING

OXFORD AND PORTLAND, OREGON
2009

2.2

Punishment and Deterrence

JEREMY BENTHAM

2.2(a) The Aims of Punishment

When any act has been committed which is followed, or threatens to be followed, by such effects as a provident legislator would be anxious to prevent, two wishes naturally and immediately suggest themselves to his mind: first, to obviate the danger of the like mischief in future: secondly, to compensate the mischief that has already been done.

The mischief likely to ensue from acts of the like kind may arise from either of two sources—either the conduct of the party himself who has been the author of the mischief already done, or the conduct of such other persons as may have adequate motives and sufficient opportunities to do the like.

Hence the prevention of offences divides itself into two branches: *particular prevention*, which applies to the delinquent himself; and *general prevention*, which is applicable to all the members of the community without exception.

Pain and pleasure are the great springs of human action. When a man perceives or supposes pain to be the consequence of an act, he is acted upon in such a manner as tends, with a certain force, to withdraw him, as it were, from the commission of that act. If the apparent magnitude of that pain be greater than the apparent magnitude of the pleasure or good he expects to be the consequence of the act, he will be absolutely prevented from performing it. The mischief which would have ensued from the act, if performed, will also by that means be prevented.

With respect to a given individual, the recurrence of an offence may be provided against in three ways:

1. By taking from him the physical power of offending.
2. By taking away the desire of offending.
3. By making him afraid of offending.

In the first case, the individual can no more commit the offence; in the second, he

From 'The Principles of Penal Law', in J Bowring (ed), *The Works of Jeremy Bentham* (1838–43) 396.

no longer desires to commit it; in the third, he may still wish to commit it, but he no longer dares to do it. In the first case, there is a physical incapacity; in the second, a moral reformation; in the third, there is intimidation or terror of the law.

General prevention is effected by the denunciation of punishment, and by its application, which, according to the common expression, *serves for an example.* The punishment suffered by the offender presents to every one an example of what he himself will have to suffer, if he is guilty of the same offence.

General prevention ought to be the chief end of punishment, as it is its real justification. If we could consider an offence which has been committed as an isolated fact, the like of which would never recur, punishment would be useless. It would be only adding one evil to another. But when we consider that an unpunished crime leaves the path of crime open, not only to the same delinquent, but also to all those who may have the same motives and opportunities for entering upon it, we perceive that the punishment inflicted on the individual becomes a source of security to all. That punishment which, considered in itself, appeared base and repugnant to all generous sentiments, is elevated to the first rank of benefits, when it is regarded not as an act of wrath or of vengeance against a guilty or unfortunate individual who has given way to mischievous inclinations, but as an indispensable sacrifice to the common safety.

2.2(b) *The Quantum of Punishment*

Rule 1. The first object, it has been seen, is to prevent, in as far as it is worth while, all sorts of offences; therefore,

The value of the punishment must not be less in any case than what is sufficient to outweigh that of the profit of the offence.

If it be, the offence (unless some other considerations, independent of the punishment, should intervene and operate efficaciously in the character of tutelary motives) will be sure to be committed notwithstanding: the whole lot of punishment will be thrown away: it will be altogether *inefficacious.*

The above rule has been often objected to, on account of its seeming harshness: but this can only have happened for want of its being properly understood. The strength of the temptation, *ceteris paribus*, is as the profit of the offence: the quantum of the punishment must rise with the profit of the offence: *ceteris paribus*, it must therefore rise with the strength of the temptation. This there is no disputing. True it is that the stronger the temptation, the less conclusive is the indication which the act of delinquency affords of the depravity of the offender's disposition. So far then as the absence of any aggravation, arising from extraor-

From *An Introduction to the Principles of Morals and Legislation* (1789) ch 14.

dinary depravity of disposition, may operate, or at the utmost, so far as the presence of a ground of extenuation, resulting from the innocence or beneficence of the offender's disposition, can operate, the strength of the temptation may operate in abatement of the demand for punishment. But it can never operate so far as to indicate the propriety of making the punishment ineffectual, which it is sure to be when brought below the level of the apparent profit of the offence.

The partial benevolence which should prevail for the reduction of it below this level, would counteract as well those purposes which such a motive would actually have in view, as those more extensive purposes which benevolence ought to have in view: it would be cruelty not only to the public, but to the very persons in whose behalf it pleads: in its effects, I mean, however opposite in its intention. Cruelty to the public, that is cruelty to the innocent, by suffering them, for want of an adequate protection, to lie exposed to the mischief of the offence: cruelty even to the offender himself, by punishing him to no purpose, and without the chance of compassing that beneficial end, by which alone the introduction of the evil of punishment is to be justified.

Rule 2. But whether a given offence shall be prevented in a given degree by a given quantity of punishment, is never any thing better than a chance; for the purchasing of which, whatever punishment is employed, is so much expended in advance. However, for the sake of giving it the better chance of outweighing the profit of the offence,

The greater the mischief of the offence, the greater is the expense, which it may be worth while to be at, in the way of punishment.

Rule 3. The next object is, to induce a man to choose always the least mischievous of two offences; therefore

Where two offences come in competition, the punishment for the greater offence must be sufficient to induce a man to prefer the less.

Rule 4. When a man has resolved upon a particular offence, the next object is, to induce him to do no more mischief than what is necessary for his purpose: therefore

The punishment should be adjusted in such manner to each particular offence, that for every part of the mischief there may be a motive to restrain the offender from giving birth to it.

Rule 5. The last object is, whatever mischief is guarded against, to guard against it at as cheap a rate as possible: therefore

The punishment ought in no case to be more than what is necessary to bring it into conformity with the rules here given.

Rule 6. It is further to be observed, that owing to the different manners and degrees in which persons under different circumstances are affected by the same exciting cause, a punishment which is the same in name will not always either really produce, or even so much as appear to others to produce, in two different persons the same degree of pain: therefore

That the quantity actually inflicted on each individual offender may correspond to the quantity intended for similar offenders in general, the several circumstances influencing sensibility ought always to be taken into account.

Of the above rules of proportion, the four first, we may perceive, serve to mark out the limits on the side of diminution; the limits *below* which a punishment ought not to be *diminished*: the fifth, the limits on the side of increase; the limits *above* which it ought not to be *increased*. The five first are calculated to serve as guides to the legislator: the sixth is calculated, in some measure, indeed, for the same purpose; but principally for guiding the judge in his endeavours to conform, on both sides, to the intentions of the legislator.

2.2(c) Cases Where Punishment is Unjustified

All punishment is mischief: all punishment in itself is evil. Upon the principle of utility, if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil.

It is plain, therefore, that in the following cases punishment ought not to be inflicted.

1. Where it is *groundless*: where there is no mischief for it to prevent: the act not being mischievous upon the whole.
2. Where it must be *inefficacious*: where it cannot act so as to prevent the mischief.
3. Where it is *unprofitable*, or too *expensive*: where the mischief it would produce would be greater than what it prevented.
4. Where it is *needless*: where the mischief may be prevented, or cease of itself, without it: that is, at a cheaper rate.

From *An Introduction to the Principles of Morals and Legislation* (1789) ch 13.

2.4

Optimal Sanctions: Any Upper Limits?

RICHARD POSNER

[Posner's theory is based on the assumption that criminals in general behave as "national calculators" or, to be more precise, that a sentencing strategy which is based on this assumption will have the greatest preventive power. On this view, which is developed further in the sentencing context and more generally in Posner's *Economic Analysis of Law,* 2nd edn. (1977), ch. 7, and *passim*), crimes are committed because the expected benefits outweigh the expected costs; or, at least, significantly fewer crimes would be committed if the expected costs were known to exceed the expected benefits. In these calculations, benefits include any economic gain from the offence and other non-economic satisfactions (e.g., in crimes of passion or revenge); and costs include not only the expected punishment but also the opportunity costs of the criminal's time, expenses necessary to commit the crime, etc. These last points are important in Posner's theory because they suggest other possibilities of controlling crime apart from increasing punishments on convicted offenders (e.g., increasing the cost or scarcity of guns, or redistributing wealth). The extract below, however, concentrates on the potential of the criminal sanction for controlling crime—eds.]

We have seen that the main thing the criminal law punishes is the pure coercive transfer, or, as it might better be described in a case of tax evasion or price-fixing, the pure involuntary transfer, of wealth or utility. In discussing what criminal penalties are optimal to deter such transfers, I shall assume that most potential criminals are sufficiently rational to be deterrable—an assumption that has the support of an extensive literature.

We saw earlier that the sanction for a pure coercive transfer should be designed so that the criminal is made worse off by his act, but now a series of qualifications must be introduced. First, some criminal acts actually are wealth-maximizing. Suppose I lose my way in the woods and, as an alternative to starving, enter an unoccupied cabin and "steal" some food. Should the punishment be death, on the theory that the crime saved my life, and therefore no lesser penalty would deter? Of course not. The problem is that while the law of theft generally punishes takings in settings of low transaction costs, in this example the costs of transacting with the

From R Posner, 'An Economic Theory of Criminal Law' (1985) 85 *Columbia Law Review* 1193.

absent owner of the cabin are prohibitive. One approach is to define theft so as to exclude such examples; the criminal law has a defense of necessity that probably would succeed in this example. But defenses make the law more complicated, and an alternative that sometimes will be superior is to employ a somewhat overinclusive definition of the crime but set the expected punishment cost at a level that will not deter the occasional crime that is value-maximizing.

There is a related but more important reason for putting a ceiling on criminal punishments such that not all crimes are deterred. If there is a risk either of accidental violation of the criminal law or of legal error, an expected penalty will induce innocent people to forgo socially desirable activities at the borderline of criminal activity. The effect is magnified if people are risk averse and penalties are severe. If for example, the penalty for carelessly injuring someone in an automobile accident were death, people would drive too slowly, or not at all, to avoid an accidental violation or an erroneous conviction.

1. "Afflictive" Punishment

The foregoing analysis shows that there is a place in the criminal justice system, and a big one, for imprisonment; and perhaps for other non-monetary criminal sanctions as well. Since the cost of murder to the victim approaches infinity, even very heavy fines will not provide sufficient deterrence of murder, and even life imprisonment may not impose costs on the murderer equal to those of the victim. It might seem, however, that the important thing is not that the punishment for murder equal the cost to the victim but that it be high enough to make the murder not pay—and surely imprisoning the murderer for the rest of his life or, if he is wealthy, confiscating his wealth would cost him more than the murder could possibly have gained him. But this analysis implicitly treats the probability of apprehension and conviction as one. If it is less than one, as of course it is, then the murderer will not be comparing the gain from the crime with the loss if he is caught and sentenced; he will be comparing it with the disutility of the sentence discounted by the probability that it will actually be imposed. Suppose, for example, that the loss to the murder victim is one hundred million dollars, the probability of punishing the murderer is .5, and the murderer's total wealth is one million dollars and will be confiscated upon conviction Then his expected punishment cost when he is deciding whether to commit the crime is only $500,000—much less than his total wealth.

This analysis suggests incidentally that the much heavier punishment of crimes of violence than seemingly more serious white-collar crimes is not, as so often thought, an example of class bias. Once it is recognized that most people would demand astronomical sums to assume a substantial risk of death, it becomes apparent that even very large financial crimes are less serious than most crimes of violence. The same people who would accept quite modest sums to run very small risks of death would demand extremely large sums to run the substantial risks that many crimes of violence create, even when death does not ensue. This point holds even if the white-collar crime (say, violating a pollution regulation) creates a safety hazard, provided that the probability that the hazard will result in the death of any given person is low. Even if it were a virtual certainty that some people would die as a result of the crime, the aggregate disutility of many small risks of death may be much smaller than a single large risk of death to a particular person. This is the nonlinear relationship between utility and risk of death that I have stressed.[1]

By the same token the argument sketched above for capital punishment is not conclusive. Because the penalty is so severe, and irreversible, the cost of mistaken imposition is very high; therefore greater resources are invested in the litigation of a capital case. Indeed, if I am right in suggesting that the cost of death inflicted with a high probability (a reasonable description of capital punishment) is not just a linear extrapolation from less severe injuries, it is not surprising that the resources invested in the litigation of a capital case may, as one observes, greatly exceed those invested in litigation in cases where the maximum punishment is life imprisonment, even if there is no possibility of parole. The additional resources expended on the litigation of capital cases may not be justified if the added deterrent effect of capital punishment over long prison terms is small. But there is scientific evidence to support the layman's intuition that it is great.[2]

Capital punishment is also supported by considerations of marginal deterrence, which require as big a spread as possible between the punishments for the least and most serious crimes. If the maximum punishment for murder is life imprisonment, we may not want to make armed robbery also punishable by life imprisonment, for then armed robbers would have no additional incentive not to murder their victims. But arguments based on marginal deterrence for a differentiated penalty structure are inconclusive, particularly when the greater offence is a complement of the lesser one, as is often the case with murder. Moreover, the argument does not lead inexorably to the conclusion that capital punishment should be the punishment for simple murder. For if it is, then we have the problem of marginally deterring the multiple murderer. Maybe capital punishment should be reserved for him, so that murderers have a disincentive to kill witnesses to the murder, though again the number of such complementary murders may be less if the initial murder is punished severely.

An important application of this principle is to prison murders. A prisoner who is serving a life sentence for murder and is not likely to be paroled has no disincentive not to kill in prison, unless prison murder is punishable by death. Considerations of complementarity might argue for making out-of-prison murders capital also, since reducing the number of murders and the fraction of murderers in prison would reduce the occasions for prison murder. What makes little sense is to have capital punishment for neither out-of-prison nor prison murders, so that the latter becomes close to a free good. This is the present situation in federal law. Notice that varying the probability of apprehension and conviction cannot preserve marginal deterrence in this situation. The probability of apprehension and conviction in the prison murder case is close to one; the problem is that for the murderer already fated to spend the rest of his life in prison, there is no incremental punishment from being convicted of murder again.

Of course there is no realistic method of preserving marginal deterrence for every crime, although medieval law tried. It is a reasonable conjecture (if no more than that) that because more medieval than modern people believed in an afterlife, because life was more brutal and painful, and because life expectancy was short, capital punishment was not so serious a punishment in those days as it is today. Furthermore, because society was poor, severe punishments were badly needed and law enforcement was inefficient, so that devoting much greater resources to catching criminals would not have been feasible or productive. In an effort to make capital punishment a more costly punishment to the criminal, especially gruesome methods of execution (for example, drawing and quartering)[3] were prescribed for especially heinous crimes, such as treason. Boiling in oil, considered more horrible than hanging or beheading, was used to punish murder by poisoning; since poisoners were especially difficult to apprehend in those times, a heavier punishment than that prescribed for ordinary murderers was (economically) indicated.

The hanging of horse thieves in the nineteenth-century American West is another example of a penalty whose great severity reflects the low probability of punishment more than the high social cost of the crime. But the most famous example is the punishment of all serious (and some not so serious) crimes by death in pre-nineteenth-century England,[4] when there was no organized police force and the probability of punishment was therefore very low for most crimes.[5]

Death is not the only modern form of "afflictive" punishment. Flogging is still used by many parents and, in attenuated form, in some schools. The economic objection to punishing by inflicting physical pain is not that it is disgusting or that people have different thresholds of pain that make it difficult to calibrate the severity of the punishment—imprisonment and death are subject to the same problem. The objection is that it may be a poor method of inflicting severe but not lethal punishment. Just to inflict a momentary excruciating pain with no aftereffects might be a trivial deterrent, especially for people who had never experienced such pain; while to inflict a level of pain that would be the equivalent of five years in prison would require measures so drastic that they might endanger the life, or destroy the physical or mental health, of the offender. For slight punishments, fines will do. Incidentally, I do not mean, by omission, to disparage non-economic objections to "afflictive punishment". But this is an article about economics.

The infliction of physical pain is not the only way in which the severity of punishment can be varied other than by varying the length of imprisonment. Size of prison cell, temperature, and quality of food could also be used as "amenity variables". It may seem very attractive from a cost-effectiveness standpoint to reduce the length of imprisonment but compensate by reducing the quality of the food served the prisoners; the costs of imprisonment to the state, but not to the prisoners, would be reduced. The problem is that this would make information about sanctions very costly, because there would be so many dimensions to evaluate. Time has the attractive characteristic of being one-dimensional, and differs from pain in that it has more variability. But as a matter of fact, society does vary the amenities of prison life for different criminals. Minimum security prisons are more comfortable than intermediate security prisons, and the latter are more comfortable than maximum security prisons. Assignments to these different tiers are related to the gravity of the crime, and in the direction one would predict.

2. Imprisonment

If society must continue to rely heavily on imprisonment as a criminal sanction, there is an argument—subject to caveats that should be familiar to the reader by now, based on risk aversion, overinclusion, avoidance and error costs, and (less clearly) marginal deterrence—for combining heavy prison terms for convicted criminals with low probabilities of apprehension and conviction. Consider the choice between combining a .1 probability of apprehension and conviction with a 10-year prison term and a .2 probability of apprehension and conviction with a five-year term. Under the second approach twice as many individuals are imprisoned but for only half as long, so the total costs of imprisonment to the government will be the same under the two approaches. But the costs of police, court officials, and the like will probably be lower under the first approach. The probability of apprehension and conviction, and hence the number of prosecutions, is only half as great. Although more resources will be devoted to a trial where the possible punishment is greater, these resources will be incurred in fewer trials because fewer people will be punished, and even if the total litigation resources are no lower, police and prosecution costs will clearly be much lower. And notice that this variant of our earlier model of high fines and trivial probabilities of apprehension and conviction corrects the most serious problem with that model—that is, solvency.

But isn't a system under which probabilities of punishment are low "unfair", because it creates ex post inequality among offenders? Many go scot-free; others serve longer prison sentences than they would if more offenders were caught. However, to object to this result is like saying that all lotteries are unfair because, ex post, they create wealth differences among the players. In an equally significant sense both the criminal justice system that creates low probabilities of apprehension and conviction and the lottery are fair so long as the ex ante costs and benefits are equalized among the participants. Nor is it correct that while real lotteries are voluntary the criminal justice "lottery" is not. The criminal justice system is voluntary: you keep out of it by not committing crimes. Maybe, though, such a system of punishment is not sustainable in practice, because judges and jurors underestimate the benefits of what would seem, viewed in isolation, savagely cruel sentences. The prisoner who is to receive the sentence will be there in the dock, in person; the victims of the crimes for which he has not been prosecuted (because the fraction of crimes prosecuted is very low) will not be present—they will be statistics. I hesitate, though, to call this an economic argument; it could be stated in economic terms by reference to costs of information, but more analysis would be needed before this could be regarded as anything better than relabelling.

There is, however, another and more clearly economic problem with combining very long prison sentences with very low probabilities of apprehension and conviction. A prison term is lengthened, of course, by adding time on to the end of it. If the criminal has a significant discount rate, the added years may not create a substantial added disutility. At a discount rate of 10 per cent, a 10-year prison term imposes a disutility only 6.1 times the disutility of a one-year sentence, and a 20-year sentence increases this figure to only 8.5 times; the corresponding figures for a five per cent discount rate are 7.7 and 12.5 times.

Discount rates may seem out of place in a discussion of non-monetary utilities and disutilities, though imprisonment has a monetary dimension, because a prisoner will have a lower income in prison than on the outside. But the reason that interest (discount) rates are positive even when there is no risk of default and the expected rate of inflation is zero is that people prefer present to future consumption and so must be paid to defer consumption. A criminal, too, will value his future consumption, which imprisonment will reduce, less than his present consumption.

The discounting problem could be ameliorated by preventive detention, whereby the defendant in effect begins to serve his sentence before he is convicted, or sometimes before his appeal rights are exhausted. The pros and cons of preventive detention involve issues of criminal procedure that would carry us beyond the scope of this article, and here I merely note that the argument for preventive detention is stronger the graver the defendant's crime (and hence the longer the optimal length of imprisonment), regardless of whether the defendant is likely to commit a crime if he is released on bail pending trial.

The major lesson to be drawn from this is that criminal sanctions are costly. A tort sanction is close to a costless transfer payment. A criminal sanction, even when it takes the form of a fine, and patently when it takes the form of imprisonment or death, is not. And yet it appears to be the optimal method of deterring most pure coercive transfers—which are therefore the central concern of the criminal law.

Notes

1. This point is overlooked in "radical" critiques of criminal law. See, e.g., S. Box, *Power, Crime, and Mystification* (1983), 9.
2. See, e.g., D. Pyle, *The Economics of Crime and Law Enforcement* (London, 1983); Ehrlich, "The Deterrent Effect of Capital Punishment: A Question of Life and Death", (1975) 65 *Am. Econ. Rev.* 397; Ehrlich and Gibbons, "On the Measurement of the Deterrent Effect of Capital Punishment and the Theory of Deterrence", (1977) 6 *J. Legal Stud.* 35; Layson, "Homicide and Deterrence: A Re-examination of the U.S.

Time-Series Evidence" (August 1984) (unpublished manuscript on file at the offices of the *Columbia Law Review*). The evidence has not gone unchallenged, of course. See D. Pyle, above, ch. 4, for discussion and references.

3. This punishment was still "on the books" in eighteenth-century England. For the grisly details, see W. Blackstone, *Commentaries* 92.
4. See, e.g., Langbein, "Shaping the Eighteenth-Century Criminal Trial: A View from the Ryder Sources", (1983) 50 *U. Chi. L Rev.* 1, 36–49.
5. Many capital sentences, however, were commuted to banishment to the colonies.

4.2

Proportionate Sentences: a Desert Perspective

ANDREW VON HIRSCH

Criminologists' interest in desert dates from the mid-1970s, with the publication of a number of works arguing that this notion should be seen as the central requirement of justice in sentencing. Once broached, the idea of desert quickly became influential. A number of American states' sentencing-guidelines systems (most notable, those of Minnesota and Oregon) have explicitly relied on it; some European sentencing-reform efforts (particularly those of Finland, Sweden and, somewhat less systematically, England) have done likewise, although these latter schemes make use of statutory statements of guiding principle, rather than specific, numerical guidelines.[1]

The groundwork for this revival of interest in desert was laid already in the post-Second World War literature of analytical moral philosophy. These writings supplied a principled critique of purely instrumental ways thinking about social and penal issues, suggesting how such reckonings were capable of sacrificing individual rights to serve majority interests. The philosophical literature also began exploring the conception of desert, suggesting how it constitutes an integral part of everyday moral judgements.[2]

The movement toward a proportionality-based sentencing theory began, perhaps, in 1971 with the publication of the Quaker-sponsored American Friends Service Committee report, *Struggle for Justice* (1971). The report recommended moderate, proportionate punishments, and opposed deciding sentence severity on predictive or rehabilitative grounds. The Friends Committee report did not rely explicitly on the idea of desert as the basis for its proposals; that was left to subsequent writings, including the Australian philosopher John Kleinig's *Punishment and Desert* (1973), my own *Doing Justice* (1976), and the British philosopher RA Duff's *Trials and Punishments* (1986). A number of influential British and Scandinavian penologists have also contributed to this literature, including AE Bottoms, Andrew Ashworth, Martin Wasik and Nils Jareborg.[3] The present essay is designed to summarise recent writing in this area, including subsequent work of Duff's and mine (von Hirsch, 1993; Duff, 2001; von Hirsch and Ashworth, 2005).

Desert theories for sentencing have had the attraction that they purport to be about *just* outcomes: the emphasis is on what the offender should fairly receive for

Originally published in the previous (first) edition of this book as Selection 4.4.

his crime, rather than how his punishment might affect his future behaviour or that of others. It also seems capable of providing more guidance: the sentencer, instead of having to address elusive empirical questions of the crime-preventative effect of the sentence, can address matters more within his or her ken, concerning the seriousness of the criminal offence—how harmful the conduct typically is, how culpable the offender was in committing it (see Selection 4.5 below).

Censure and Penal Desert

There have been a variety of retributive or desert-based accounts of punishment, ranging from intuitionist theories (see Selection 4.1 above), to talionic notions of requiting evil for evil, to conceptions that see punishment as taking away the 'unjust advantage' over others which the offender obtains by choosing to offend (see Finnis, 1980, 263–4).[4] The desert-based conception examined in this essay, however, relies on a different account: one emphasizing the communicative features of punishment.

The criminal sanction conveys censure: punishing consists of imposing a deprivation on someone, because he purportedly has committed a wrong, under circumstances and in a manner that conveys disapprobation of him for his wrong. Treating the offender as a wrongdoer, Richard Wasserstrom (1980) has pointed out, is central to the idea of punishment. The difference between a tax and a fine, for example, does not rest in the material deprivation imposed—which is money in both cases. It consists, rather, in the fact that with a fine, money is taken in a manner that conveys disapproval or censure; whereas with a tax, no disapproval is implied.

A sanction that treats the conduct as wrong—that is, not a 'neutral' sanction—has two important moral functions that are not reducible to crime prevention. One is to recognise the importance of the rights that have been infringed. The censure in punishment conveys to victims and potential victims the acknowledgement that they are wronged by criminal conduct, that rights to which they properly are entitled have been infringed. The other (and perhaps still more important) role of censure is that of addressing the offender as a moral agent, by appealing to his or her sense of right and wrong. This is not just a crime-prevention strategy, however, for otherwise there would be no point in censuring actors who are repentant already (since they need no blame to make them regret their actions and to try to desist in future) or who seemingly are incorrigible (since they will not change despite the censure). Any human actor, this communicative perspective suggests, should be treated as a moral agent, having the capacity (unless clearly incompetent) of evaluating others' assessment of their conduct. A response to criminal wrongdoing that conveys blame gives the individual the opportunity to respond in ways that are typically those of an agent capable of moral deliberation: to recognise the wrongfulness of action; feel remorse; to make efforts to desist in future—or try to give reasons why the conduct was not actually wrong. What a purely 'neutral' sanction not embodying blame would deny, even if no less effective in preventing crime, is precisely this recognition of the person's status as a moral agent. A neutral sanction would treat offenders and potential offenders much as beasts in a circus—as beings which must be restrained, intimidated, or conditioned into submission because they are incapable of understanding that predatory conduct is wrong (von Hirsch, 1993; and, more fully, von Hirsch and Ashworth, 2005, ch 2).

Relying on this idea of censure helps remove some of the seeming mysteriousness of penal desert judgements: censure or blaming involves everyday moral judgements used in a wide variety of social contexts, of which punishment is just one. This account also helps address another objection traditionally raised against retributive penal theories, namely, their seeming harshness—their apparent insistence on an eye for an eye. Once the paying back of evil for evil is not seen as the underlying idea, penal desert does not demand visitation of suffering equal to the harm done. What is called for instead is punishments that are *proportionate* to the seriousness of the criminal conduct. Proportionate punishments—even if not involving harm-for-harm equivalence—would suffice to convey blame for various crimes according to their degree of reprehensibleness. Indeed, several advocates of the desert perspective (including myself) have advocated substantial reductions of penalty levels (see, eg Singer, 1979; Duff, 1986; von Hirsch, 1993, ch 2; Ashworth, 2005, ch 9; Jareborg, 1995).

Can the institution of punishment be explained purely in terms of censure? Punishment does convey blame, but does so in a special way—through visitation of deprivation ('hard treatment') on the offender. That deprivation is, of course, the vehicle through which the blame is expressed. But why use this vehicle, rather than simply expressing blame in symbolic fashion? Some adherents of the communicative view of desert, most notably RA Duff, hold that the hard-treatment component of the penal sanction can itself be explained in desert terms: Duff treats the deprivation involved in punishment as providing a kind of secular penance (see Selection 4.3).[5] I have my doubts, however. The reason for having the institution of punishment (that is, for expressing disapproval through hard treatment, instead of merely censuring) seems to have to do with keeping predatory behaviour within tolerable limits. Had the criminal sanction no usefulness in preventing crime, there should be no need to visit material deprivation on those who offend. True, we might still wish to devise another way of issuing authoritative judgements of blaming, for such predatory behaviour as occurs. But those judgements, in the interest of keeping state-inflicted suffering to a minimum, would no longer be linked to purposive infliction of suffering (von Hirsch, 1993, 14; von Hirsch and Ashworth, 2005, ch 2).

If the criminal sanction thus serves to prevent crime as well as censure, how is this consistent with treating offenders and potential offenders as moral agents? The hard-treatment in punishment, in my view, serves as a prudential reason for obedience to those insufficiently motivated by the penal censure's moral appeal. But this should *supplement* rather than replace the normative reasons for desisting from crime conveyed by penal censure—that is, it provides an *additional* reason for compliance to those who are capable of recognising the law's moral demands, but who are also tempted to disobey them. The law thus addresses *ourselves*, not a distinct 'criminal' class of those considered incapable of grasping moral appeals. And it addresses us neither as perfectly moral agents (we are not like angels) nor as beasts which only can be coerced through threats, but, rather, as moral but fallible agents who need some prudential supplement to help us resist criminal temptation (see von Hirsch and Ashworth, 2005, ch 2; see also Narayan, 1993). However, this account (as will be discussed further below) calls for moderation in the overall severity in punishment levels. The harsher the penalty system is, the less plausible it becomes to see it as embodying chiefly a moral appeal rather than a system of bare threats.

The Rationale for Proportionality

In a minimal sense, proportionality always had a role in sentencing policy: penalties that were grossly excessive in relation to the gravity of the offence were perceived as unfair. Statutory maximum sentences reflected that understanding, and it also had a constitutional dimension: some jurisdictions have adopted a constitutional bar against grossly excessive punishments.[6] This, however, gave the notion of proportionality only the outer, constraining role of barring draconian sanctions for lesser offences. Short of these (rather high) maximum limits, proportionality had small weight in theories about how sanctions should be determined, with consequentialist concerns (about rehabilitation, incapacitation and deterrence) counting chiefly instead.

What is distinctive about contemporary desert theory is that it moves notions of proportionality from this merely peripheral to a central role in deciding sanctions. The primary basis for deciding quanta of punishments, under this theory, is the principle of proportionality or 'commensurate deserts', requiring the severity of the penalty to be proportionate to the gravity of the defendant's criminal conduct. The criterion for deciding the quantum of punishment is thus retrospective rather than consequentialist: the seriousness of the offence for which the defendant stands convicted.

What is the basis for this principle? The censure account, just discussed, provides the explanation. If punishment embodies blame, then how much one punishes will convey how much the conduct is condemned. If crime X is punished more severely than crime Y, this connotes the greater disapprobation of crime X. Punishments, consequently, should be allocated consistently with their blaming implications. When penalties are arrayed in severity according to the gravity of offences, the disapprobation thereby conveyed will reflect the degree of reprehensibleness of the conduct. When punishments are arrayed otherwise, this is not merely inefficient (who knows, it might sometimes work), it is also unfair; offenders are being visited with more or less censure than the comparative blameworthiness of their conduct would warrant (von Hirsch, 1993, ch 2; von Hirsch and Ashworth, 2005, ch 9).

Equity is sacrificed when the proportionality principle is disregarded, even when this is done for the sake of crime prevention. Suppose that offenders A and B commit and are convicted of criminal conduct of approximately the same degree of seriousness. Suppose B is deemed more likely to reoffend, and therefore is given a longer sentence. Notwithstanding the possible preventative utility of that sentence, the objection remains that B, through his more severe punishment, is being treated as more to blame than A, though their conduct had the same degree of blameworthiness.

A possible objection (see Dolinko, 1992) to this argument might run as follows: if the component of 'hard treatment' in the criminal sanction serves a crime-preventative as well as a purely censuring function (as argued earlier in this essay), then why cannot one allocate the relative severities of punishment in part on preventative grounds rather than purely on the basis of offence seriousness? The reply is that punishment's deprivations and its reprobative connotations are inextricably intermixed: it is the threatened penal deprivation that expresses the degree of censure. If the deprivations visited on a given type of crime are increased, even for preventative reasons, this (necessarily) increases the severity of the punishment. But changing the severity, relative to other penalties, alters the implicit censure—which would not be justified if the seriousness of the conduct is itself unchanged (von Hirsch, 1993, 16–17; von Hirsch and Ashworth, 2005, ch 9).

Desert as 'Determining' or 'Limiting'?

If the principle of proportionality is so important, is it a 'determining' or merely a 'limiting' principle? While our sense of justice tells us that criminals should be punished as they deserve, there do not seem to be definite quanta of severity associated with our desert-judgements. Armed robbers have committed a serious offence, deserving of substantial punishment, but it is not apparent whether that should consist of two years' confinement, three years, or some shorter or longer period.

One response to this problem has been Norval Morris's: to say that desert is merely a limiting principle (see Selection 4.4 below by Richard Frase, elaborating on Morris's views). It tells us, he asserts, not how much robbers deserve, but only some broad limits beyond which their punishments would be *un*deserved. Within such limits, the sentence can be decided on other (for example, predictive) grounds. This view, however, would mean that persons who commit similar crimes could receive quite different amounts of punishment. If punishment embodies blame as a central characteristic, it becomes morally problematic to visit such different degrees of severity, and hence of implicit blame, on comparably blameworthy transgressions.

A conceivable opposite response, but scarcely a plausible one, would be the heroic intuitionist stance: that if we only ponder hard enough we will perceive deserved quanta of punishments: that robbers ordinarily deserve so-and-so many months or years of confinement, and so forth. Our intuitions, however, fail to provide such answers.

The way out of this apparent dilemma is to recognise the crucial difference between the comparative ranking of punishments on one hand, and the overall magnitude and anchoring of the penalty scale on the other. With respect to comparative rankings, *ordinal* proportionality provides considerable guidance: persons convicted of similar crimes should receive punishments of comparable severity (save in special aggravating or mitigating circumstances altering the harm or culpability of the conduct in the particular circumstances); and persons convicted of crimes of differing gravity should suffer punishments correspondingly graded in onerousness. These ordinal-proportionality requirements are no mere limits, and they are infringed when equally reprehensible conduct is punished markedly unequally in the manner that Frase, following Morris, suggests (von Hirsch, 1986, ch 4; von Hirsch and Ashworth, 2005, ch 9).

Desert provides less constraint, however, on the penalty scale's overall dimensions and anchoring points. This is because the censure expressed through penal deprivations is, to a considerable degree, a convention. When a penalty scale reflects the comparative gravity of crimes, making pro rata decreases or increases in the prescribed sanctions constitutes a change in that convention.

This distinction helps resolve the dilemma just mentioned. The leeway which desert allows in fixing the scale's overall degree of onerousness explains why we cannot perceive a single right or fitting penalty for a crime. Whether X months, Y months or somewhere in between is the appropriate penalty for robbery depends on how the scale has been anchored and what punishments are prescribed for other crimes. Once those anchoring points are decided, however, the more restrictive requirements of ordinal proportionality apply. This explains why it would be inappropriate to give short prison terms to some robbers and long ones to other robbers on the basis of (say) predictive factors not reflecting the degree of seriousness of the criminal conduct.

Does this purported solution still leave the anchoring of the scale too wide open? Could it not permit a very severe penalty scale, as long it is not so harsh as to impose drastic penalties on manifestly trivial crimes? My suggested answer to this question has been that high overall severity levels are inconsistent with the moral functions of penal censure. Through punishments' censuring features, the criminal sanction offers a normative reason for desisting to human beings seen as moral agents: that doing certain acts is wrong and hence should be refrained from. Punishments' material deprivations can then be viewed (as noted earlier) as providing a supplemental disincentive—as providing humans (given human fallibility and the temptations of offending) an additional prudential reason for complying with the law. The higher penalty levels rise, however, the less the normative reasons for desisting supplied by penal censure will count and the more the system becomes in effect a bare system of threats (in Hegel's apt words, a stick that might be raised to a dog). Very high penalty levels thus would 'drown out' the moral message embodied in penal censure. To the extent this argument is accepted, it points towards keeping penalties at moderate levels (von Hirsch, 1993, ch 5).[7]

Inclusion of Crime-control Aims?

Desert theory sets priorities among sentencing aims: it assumes that a just system accords greater importance to proportionately ordered sanctions than to the pursuit of other objectives—say, incapacitating those deemed higher risks. This understandably evokes discomfort: why cannot one seek proportionality and pursue other desired ends, whether they be treatment, incapacitation or something else?

To an extent, a desert model permits consideration of other aims: namely, to the degree this is consistent with the proportionate ordering of penalties. Thus, when there is a choice between two non-custodial sanctions of approximately equivalent severity (say, a unit-fine of so many days' earnings and intensive probation for a specified duration), proportionality constraints are not offended when one of these is chosen over the other on (say) treatment grounds. Desert theorists thus have come forward with schemes for scaling intermediate, non-custodial penalties; these sanctions would be ranked in severity according to the gravity of the crime, but penalties of roughly equivalent onerousness could be substituted for one another when treatment or feasibility concerns so indicate (see Wasik and von Hirsch, 1988). Nevertheless, a pure desert model remains a somewhat constraining one: ulterior aims may be relied upon only where these do not substantially alter the comparative severity of penalties. Giving substantial extra prison time to persons deemed high risks would thus breach the model's requirements. Why not, then, relax the model's constraints to allow greater scope to such other aspirations?

A possibility—sometimes referred to as a 'modified' desert model— would be to relax the constraints to a limited degree. Proportionality would ordinarily determine comparative punishment levels, but deviations would be permitted in case of the gravest risks of crime (see Robinson, 1987; von Hirsch, 1993: 48–53). Here, the idea is that avoiding extraordinary harms is so important a goal as to warrant some sacrifice of fairness. This position differs from ordinary penal consequentialism, however, in that departures from desert requirements could be invoked only exceptionally, to deal with threats of an extraordinary nature (see Selection 3.3 above). Alternatively, deviations could more regularly be permitted, but these would be restricted ones: say, a deviation of no more than 10 or 15% from the deserved sentence. While departures from proportionality involve a sacrifice of equity, the extent of that sacrifice depends on the degree of the deviation from desert constraints. Limited deviations, it might be argued, would permit the pursuit of ulterior objectives without 'too much' unfairness (von Hirsch, 1993: 54–56; von Hirsch and Ashworth, 2005: appendix 2).

These mixed approaches still make desert the primary determinant for the ordering of penalties, but give some extra scope for ulterior purposes. Even such schemes remain constraining, however: especially dangerous offenders might be given substantial extra prison time, but not the ordinary potential recidivist; some extra leeway might be granted to suit a non-custodial penalty to the offender's apparent treatment needs, but not a great deal.

Could still more scope be given to non-desert considerations? In a hybrid rationale, either desert will predominate or something else will. If—in the ordinary case—the seriousness of the crime is the penalty's primary determinant, the system remains desert-dominated. If other (say, crime-preventive) aims are given the greater emphasis, however, that creates a system dominated by those aims. That will reintroduce the familiar problems of consequentialist sentencing schemes—for example, those relating to equity among offenders – and those of insufficient systematic knowledge of preventive effects (see Chapters 1–3 above).

In assessing these alternatives, it needs to be borne in mind that even a 'purely' desert-based sentencing scale is likely to have collateral crime-prevention benefits—in such deterrence as its penalties achieve, and in the possible incapacitative effects of the prison sentences it prescribes for serious crimes. Departing from proportionality for the sake of crime prevention, then, will call not just for a showing that preventive effects might be achieved (for a desert-based system may achieve these too); instead, it would call for a showing that departures are likely to yield *enhanced* preventative effects—which is no easy matter to establish. And here one is likely to confront a fairness/effectiveness trade-off: because crime rates tend to be rather insensitive to small variations in punishment, modest departures from proportionality are likely to have relatively little impact; large departures might possibly work better, but these precisely are the ones that are most troublesome on moral grounds (see Selection 3.4 above).

Other Issues: Severity and Social Deprivation

Must desert lead to harsh penalties? As the theory emerged and became influential at a time when penalty levels rose in many jurisdictions, some critics have argued that the theory must in part be responsible for such increases (Braithwaite and Pettit, 1990: 15–16; Garland, 2001: 9). However, desert theory itself does not require a severe sentencing policy— indeed, as noted earlier, it permits (indeed, arguably points toward) considerable penalty reductions. Moreover, the sentence-reform schemes which rely explicitly on notions of desert tend not to be severe ones: the Minnesota and Oregon sentencing guidelines, for example, call for relatively modest penalties by American standards; European desert-orientated sentencing standards, such as those of Finland and Sweden, are likewise associated with penal moderation (von Hirsch and Ashworth, 2005: ch 6). Measures which most clearly call for tougher sanctions tend to utilise criteria inconsistent with proportionality: mandatory sentences, for example, select particular offence categories for harsh treatment, without regard to the gravity of the offence involved, or the penalties imposed for other offences (*idem*).

Another issue is that concerning just punishment and social deprivation (see Murphy, 1973). Many offenders live in grim social environments that restrict their opportunities for living tolerable and law-abiding lives. Should such persons be punished differently? The penal law is a poor instrument for rectifying social ills: it is social policy, rather than criminal policy, that is the appropriate instrument for addressing problems of social deprivation. But the question remains disturbing, nevertheless. If social policy fails to alleviate poverty and deprivation, how should the deprived offender be sentenced?

It has been pointed out that desert theory at least does not *add* to the punishment imposed on deprived persons—whereas penal consequentialism would do so, to the extent that social deprivation is a sign (say) of greater dangerousness (von Hirsch, 1976, ch 17). But the questions remain whether such persons should have their penalties *reduced*, in view of the obstacles they face in leading law-abiding lives. And if they are to be reduced, puzzling questions remain concerning the grounds for reduction: whether on the basis of diminished culpability (see Hudson, 1998), or for compassion-related reasons (as Andrew Ashworth and I have recently suggested).[8] Granting such mitigation would also create a host of practical and political difficulties—so that the perplexity remains.

Notes

1. For an analysis of the Oregon and Minnesota Guidelines, see von Hirsch *et al* (1987), ch 5 and von Hirsch (1995); for the Swedish sentencing scheme, see Jareborg (1995); for the English system after the Criminal Justice Act 1991, see Ashworth (2005). Proportionality-based systems of sentencing guidance are now being developed in New Zealand, and have been proposed for South Africa (South African Law Commission, 2000).
2. For a critique of utilitarianism, see Williams (1973); for earlier philosophical writings on the idea of desert, see Armstrong (1961) and Morris (1968).
3. See Suggestions for Further Reading, this chapter, for references.
4. For a critique of the 'unjust advantage' theory, see von Hirsch (1986), 57–9; Duff (1986), ch 8.
5. In Duff's view, the hard treatment in punishment should serve to bring the criminal to understand, and repent of, his wrongdoing—and also to provide a vehicle which will enable him to work through and express his penitent understanding.
6. The US Supreme Court formerly held that grossly excessive punishments violated the Constitutional ban on cruel and unusual punishments (see *Weems v US* [(1910) 217 US 349]), but the Court later overruled that doctrine (see *Rummel v Estelle* (1980) 445 US

263). The German Constitutional Court has adopted doctrines barring gross disproportionality of sentence in relation to the seriousness of the crime. For fuller discussion of this question of excessive punishments, see van Zyl Smit and Ashworth (2004).

7. There exists a conceptually separate further reason for keeping penalty levels low—namely, the idea of 'parsimony'—of keeping state-inflicted suffering to a minimum; see von Hirsch and Ashworth (2005), 142.
8. For a fuller recent discussion of such compassion-based grounds, see von Hirsch and Ashworth (2005), ch 5.

References

American Friends Service Committee, (1971) *Struggle for Justice* (New York: Hill & Wang).

Armstrong, KG, (1961) 'The Retributivist Hits Back' 70 *Mind* 471.

Ashworth, A, (2005) *Sentencing and Criminal Justice*, 4th edn (Cambridge, Cambridge University Press).

Braithwaite, J, and Pettit, P, (1990) *Not Just Deserts* (Oxford, Oxford University Press).

Dolinko, D, (1992) 'Three Mistakes of Retributivism' 39 *UCLA Law Review* 1623.

Duff, RA, (1986) *Trials and Punishments* (Cambridge: Cambridge University Press).

—— (2001) *Punishment, Communication and Community* (Oxford, Oxford University Press).

Finnis, J, (1980) *Natural Law and Natural Rights* (Oxford, Oxford University Press)

Garland, D, (2001) *The Culture of Control: Crime and Social Order in Contemporary Society* (Oxford, Oxford University Press).

Hudson, B, (1998) 'Doing Justice to Difference' in A Ashworth and M Wasik (eds), *Fundamentals of Sentencing Theory: Essays in Honour of Andrew von Hirsch* (Oxford, Oxford University Press) ch 9.

Jareborg, N, (1995) 'The Swedish Sentencing Reform', in: C.M.V. Clarkson and R. Morgan (eds), *The Politics of Sentencing Reform* (Oxford, Oxford University Press).

Kleinig, J, (1973) *Punishment and Desert* (The Hague, Martinus Nijhoff).

Morris, H, (1968) 'Persons and Punishment' 52 *The Monist* 475.

Murphy, JG, (1973) 'Marxism and Retribution' 2 *Philosophy and Public Affairs* 217.

Narayan, U, (1993) 'Adequate Responses and Preventive Benefits' 13 *Oxford Journal of Legal Studies* 166.

Robinson, P, (1987) 'Hybrid Principles for the Distribution of Criminal Sanctions' 82 *Northwestern Law Review* 19.

Singer, R, (1979) *Just Deserts* (Cambridge, MA, Ballinger Publishing Co).

South Africa Law Commission (2000) *Report: Sentencing—a New Sentencing Framework* (Pretoria, Law Commission).

von Hirsch, A, (1995) 'Proportionality and Parsimony in American Sentencing Guidelines' in CMV Clarkson and R Morgan (eds) *The Politics of Sentencing Reform* (Oxford, Oxford University Press).

von Hirsch, A, (1993) *Censure and Sanctions* (Oxford, Oxford University Press).

von Hirsch, A, (1986) *Past or Future Crimes* (Manchester, Manchester University Press).

von Hirsch, A, (1976) *Doing Justice* (New York, Hill & Wang).

von Hirsch, A and Ashworth, A, (2005) *Proportionate Sentencing: Exploring the Principles* (Oxford, Oxford University Press).

von Hirsch, A, Knapp, K and Tonry, M, (1987) *The Sentencing Commission and Its Guidelines* (Boston, MA, Northeastern University Press).

van Zyl Smit, D and Ashworth, A, (2004) 'Disproportionate Sentences as Human Rights Violations' 67 *Modern Law Review* 541.

Wasik, M and von Hirsch, A, (1988) 'Non-Custodial Penalties and the Principles of Desert', *Criminal Law Review* 555.

Wasserstrom, R, (1980) 'Punishment' in R Wasserstrom, *Philosophy and Social Issues: Five Studies* (Notre Dame, IN, University of Notre Dame Press).

Williams, B, (1973) 'A Critique of Utilitarianism' in JSC Smart and B Williams, *Utilitarianism: For and Against* (Cambridge, Cambridge University Press).

cannot rectify racial inequalities and social injustices; it will do well if it does not exacerbate them. It is proper that predictions of violence should figure in many decisions in applying the criminal law, and if they are applied within principles that I am seeking to tease out, that is all that can be expected. My submissions may be in error, but if they are, then anyone seeking to apply predictions should offer alternative predictions. We cannot properly close our eyes to the different threats that criminal X and criminal Y pose to the community. But it is of first importance that we base our decisions about their respective dangerousness on validated knowledge and not on prejudice, particularly racial prejudice, and hence that we insist on the most careful validation of such stereotypes of dangerousness; my submissions are an effort to define what is required to achieve such validation. We must insist, if predictions are to be used, that they be reliable.

To conclude. As is so often the case with issues of justice, procedural and evidentiary issues become of central importance. Let me put the point curtly and again. Clinical predictions of dangerousness unsupported by actuarial studies should never be relied on. Clinical judgments firmly grounded on well-established base expectancy rates are a precondition, rarely fulfilled, to the just invocation of prediction of dangerousness as a ground for intensifying punishment.

I must admit that, if my submissions are accepted, I doubt the availability of sufficient knowledge to meet the necessary preconditions of just sentencing based on express predictions of violence. Further that gap in our knowledge should make us sceptical about our present widespread reliance on implicit and intuitive predictions of dangerousness in exercising discretion—in situations where we do not declare that usage as we do in the situations I have been discussing.

3.6

Selective Incapacitation

JAMES Q. WILSON

When criminals are deprived of their liberty, as by imprisonment (or banishment, or very tight control in the community), their ability to commit offences against citizens is ended. We say these persons have been "incapacitated", and we try to estimate the amount by which crime is reduced by this incapacitation.

Incapacitation cannot be the sole purpose of the criminal justice system; if it were, we would put everybody who has committed one or two offences in prison until they were too old to commit another. And if we thought prison too costly, we would simply cut off their hands or their heads. Justice, humanity, and proportionality, among other goals, must also be served by the courts.

But there is one great advantage to incapacitation as a crime control strategy—namely, it does not require us to make any assumptions about human nature. By contrast, deterrence works only if people take into account the costs and benefits of alternative courses of action and choose that which confers the largest net benefit (or the smallest net cost). Though people almost surely do take such matters into account, it is difficult to be certain by how much such considerations affect their behaviour and what change, if any, in crime rates will result from a given, feasible change in either the costs of crime or the benefits of not committing a crime. Rehabilitation works only if the values, preferences, or time-horizon of criminals can be altered by plan. There is not much evidence that we can make these alterations for large numbers of persons, though there is some evidence that it can be done for a few under certain circumstances.

Incapacitation, on the other hand, works by definition: its effects result from the physical restraint placed upon the offender and not from his sub-

From James Q. Wilson, "Dealing with the High-Rate Offender", (1983) 72 *The Public Interest* 52 with some textual changes. Reprinted by permission of the author and publisher.

jective state. More accurately, it works provided at least three conditions are met: some offenders must be repeaters, offenders taken off the streets must not be immediately and completely replaced by new recruits, and prison must not increase the post-release criminal activity of those who have been incarcerated sufficiently to offset the crimes prevented by their stay in prison.

The first condition is surely true. Every study of prison inmates shows that a large fraction (recently, about two-thirds) of them had prior criminal records before their current incarceration; every study of ex-convicts shows that a significant fraction (estimates vary from a quarter to a half) are rearrested for new offences within a relatively brief period. In short, the great majority of persons in prison are repeat offenders, and thus prison, whatever else it may do, protects society from the offences these persons would commit if they were free.

The second condition—that incarcerating one robber does not lead automatically to the recruitment of a new robber to replace him—seems plausible. Although some persons, such as Ernest van den Haag, have argued that new offenders will step forward to take the place vacated by the imprisoned offenders, they have presented no evidence that this is the case, except, perhaps, for certain crimes (such as narcotics trafficking or prostitution) which are organized along business lines.[1] For the kinds of predatory street crimes with which we are concerned—robbery, burglary, auto theft, larceny—there are no barriers to entry and no scarcity of criminal opportunities. No one need wait for a "vacancy" to appear before he can find an opportunity to become a criminal. The supply of robbers is not affected by the number of robbers practicing, because existing robbers have no way of excluding new robbers and because the opportunity for robbing (if you wish, the "demand" for robbery) is much larger than the existing number of robberies. In general, the earnings of street criminals are not affected by how many "competitors" they have.

The third condition that must be met if incapacitation is to work is that prisons must not be such successful "schools for crime" that the crimes prevented by incarceration are outnumbered by the increased crimes committed after release attributable to what was learned in prison. It is doubtless the case that for some offenders prison is a school; it is also doubtless that for other offenders prison is a deterrent. The former group will commit more, or more skillful, crimes after release; the latter will commit fewer crimes after release. The question, therefore, is whether the net effect of these two offsetting tendencies is positive or negative. In general, there is no evidence that the prison experience makes offenders as a whole more criminal, and there is some evidence that certain kinds of offenders (especially certain younger ones) may be deterred by a prison experience. Moreover, interviews with prisoners reveal no relationship between the number of crimes committed and whether the offenders had served a prior prison term. Though there are many qualifications that should be made to this bald summary, there is no evidence that the net effect of prison is to increase the crime rates of ex-convicts sufficiently to cancel out the gains to society resulting from incapacitation.

To determine the amount of crime that is prevented by incarcerating a given number of offenders for a given length of time, the key estimate we must make is the number of offences a criminal commits per year free on the street.[2] If a community experiences one thousand robberies a year, it obviously makes a great deal of difference whether these robberies are the work of ten robbers, each of whom commits one hundred robberies per year, or the work of one thousand robbers, each of whom commits only one robbery per year. In the first case, locking up only five robbers will cut the number of robberies in half; in the second case, locking up one hundred robbers will only reduce the number of robberies by 10 per cent.

In the late 1970s, researchers at the Rand Corporation had been interviewing prisoners (first in California, then in other states) to find out directly from known offenders how much crime they were committing while free.[3] No-one can be certain, of course, that the reports of the convicts constitute an accurate record of their crimes, undetected as well as detected, but the Rand researchers cross-checked the information against arrest records and looked for evidence of internal consistency in the self-reports. Moreover, the inmates volunteered information about crimes they had committed but for which they had not been arrested. Still, it is quite possible that the self-reports were somewhat inaccurate. However, it is reasonable to assume that inmates would be more likely to conceal crimes they did commit rather than admit to crimes they did not commit. Thus, any errors in these self-reports probably lead to an underestimate of the true rate of criminality of these persons.

The Rand Group learned that the "average" individual offence rate was virtually a meaningless term because the inmates they interviewed differed so sharply in how many crimes they committed. A large number of offenders committed a small number of offences while free and a small number of offenders committed a very large number of offences. In statistical language, the distribution of offences was highly skewed. For example, the median number of burglaries committed by the inmates in the three states was about 5 a year, but the 10 per cent of the inmates who were the highest-rate offenders committed an average of 232 burglaries a year. The median number of robberies was also about 5 a year, but the top 10 per cent of offenders committed an average of 87 a year. As Peter W. Greenwood, one of the members of the Rand group, put it, incarcerating one robber who was among the top 10 per cent in offence rates would prevent more robberies than incarcerating eighteen offenders who were at or below the median.[4]

All the evidence we have implies that, for crime-reduction purposes, the most rational way to use the incapacitative powers of our prisons would be to do so selectively. Instead of longer sentences for everyone, or for persons who have prior records, or for persons whose present crime is especially grave, longer sentences would be given primarily to those who, when free, commit the most crimes.

But how do we know who these high-rate, repeat criminals are? Knowing the nature of the present offence is not a good clue. The reason for this is quite simple—most street criminals do not specialize. Today's robber can be tomorrow's burglar and the next day's car thief.[5] When the police happen to arrest him, the crime for which he is arrested is determined by a kind of lottery—he happened to be caught red-handed, or as the result of a tip, committing a particular crime that may or may not be the same as either his previous crime or his next one. If judges give sentences based entirely on he gravity of the present offence, then a high-rate offender may get off lightly because on this occasion he happened to be caught snatching a purse. The low-rate offender may get a long sentence because he was unlucky enough to be caught robbing a liquor store with a gun.

Prosecutors have an understandable tendency to throw the book at persons caught committing a serious crime, especially if they have been caught before. To a certain extent, we want to encourage that tendency. After all, we not only want to reduce crime, we want to see criminals get their just deserts. Society would not, and should not, tolerate a system in which a prosecutor throws the book at purse snatchers and lets armed robbers off with a suspended sentence. But while society's legitimate desire for retribution must set the outer bounds of any sentencing policy, there is still room for flexibility within those bounds. We can, for example, act so that all robbers are punished with prison terms, but give, within certain relatively narrow ranges, longer sentences to those robbers who commit the most crimes.

If knowing the nature of the present offence and even knowing the prior record of the offender are not accurate guides to identifying high-rate offenders, what is? Obviously, we cannot ask the offenders. They may cooperate with researchers once in jail, but they have little incentive to cooperate with prosecutors before they go to jail, especially if the price of

co-operation is to get a tougher sentence. But we can see what legally admissible, objective attributes of the offenders best predict who is and who is not a high-rate offender. In the Rand study, Greenwood and his colleagues discovered, by trial and error, that the following seven factors, taken together, were highly predictive of a convicted person being a high-rate offender: he (1) was convicted of a crime while a juvenile (that is, before age 16), (2) used illegal drugs as a juvenile, (3) used illegal drugs during the previous two years, (4) was employed less than 50 per cent of the time during the previous two years, (5) served time in a juvenile facility, (6) was incarcerated in prison more than 50 per cent of the previous two years, and (7) was previously convicted for the present offence.

Using this scale, Greenwood found that 82 per cent of those predicted to be low-rate offenders in fact were, and 82 per cent of those predicted to be medium- or high-rate offenders also were. To understand how big these differences are, the median California prison inmate who is predicted to be a low-rate offender will in fact commit slightly more than one burglary and slightly less than one robbery per year free. By contrast, the median California inmate who is predicted to be a high-rate offender will commit ninety-three burglaries and thirteen robberies per year free. In other states, this prediction scale may be more or less accurate.

Opinions differ as to the effect on the crime rate and prison population of making sentences for high-rate offenders longer than those for low-rate ones. Greenwood applied his scale to California and found that if all low-rate robbers received two-year prison terms (most now receive longer ones) and all high-rate robbers received seven-year terms (most now receive shorter ones), the number of robberies committed in the state would drop by an estimated 20 per cent with no increase in the prison population.

Obviously, a policy of reducing crime by selective incapacitation (that is, by adjusting prison terms to reflect predicted individual offence rates) raises a number of issues. Though these issues are important, one must bear in mind that they cannot be resolved by comparing selective incapacitation to some ideal system of criminal justice in which everyone receives exactly his just deserts. No such system exists or ever will. One must compare instead the proposed policy with what exists now, with all its imperfections, and ask whether the gains in crime reduction are worth the risks entailed when we try to make predictions about human behaviour.

The first issue is whether it is permissible to allow crime-control to be an objective of sentencing policy. Some persons, such as Andrew von Hirsch, claim that only retribution—what he calls "just deserts"—can be a legitimate basis for sentencing.[6] To some extent, he is undoubtedly correct. Even if we were absolutely certain that a convicted murderer would never murder again, we would still feel obliged to impose a relatively severe sentence in order to vindicate the principle that life is dear and may not be unlawfully taken without paying a price. Moreover, the sentences given low-rate offenders must reflect society's judgment as to the moral blame such behaviour deserves, and the sentences given high-rate offenders ought not exceed what society feels is the highest sentence appropriate to the crime for which the offenders were convicted. And low-rate offenders should get a sufficiently severe sentence to help persuade them, and others like them, not to become high-rate offenders. Still, after allowing for all of these considerations, there will inevitably remain a range of possible sentences within which the goal of incapacitation can be served. The range will exist in part because there is no objective way to convert a desire for retribution into a precise sentence for a given offence and in part because legislatures will almost invariably act so as to preserve some judicial discretion so that the circumstances of a case which cannot be anticipated in advance may affect the sentence. Among those circumstances is a concern for protecting society from the threat that a given offender represents.

The second issue is whether our prediction methods are good enough to allow them to influence sentence length. The answer to that question depends on what one will accept as "good enough". Absolute certainty will never be attainable. Moreover, criminal justice *now*, at almost every stage, operates by trying to predict future behaviour. When a prosecutor decides how much plea bargaining he will allow, he is trying to predict how a judge or jury will react to his evidence, and he is often trying to guess how dangerous an offender he has in his grasp. When a judge sets bail, he is always making a prediction about the likelihood of a person out on bail showing up for his trial and is frequently trying to predict whether the person, if out on bail, will commit another crime while free. When a defence attorney argues in favour of his client being released on his own recognizance, without bail, he is trying to persuade the judge to accept his prediction that the accused will not skip town. When the judge passes a sentence, he is trying, at least in part, to predict whether the convicted person represents a future threat to society. When a parole board considers a convict's application for early release, it tries to predict—often on the basis of a quantitative system, called a "base expectancy table"—whether the person will become a recidivist if released. Virtually every member of the criminal justice system is routinely engaged in predicting behaviour, often on the basis of very scant knowledge and quite dubious rules of thumb. The question, therefore, is this: are the kinds of predictions that

scholars such as Greenwood make about future criminality better (more accurate) and thus fairer than the predictions prosecutors and judges now make?

A third issue is tougher. Is it fair for a low-rate offender who is caught committing a serious crime to serve a shorter sentence (because he is not much of a threat to society) than a high-rate offender who gets caught committing a relatively minor offence? Probably not. Sentences would have to have legal boundaries set so that the use of selective incapacitation could not lead to perverse sentences—armed robbers getting one year, purse-snatchers getting five.

Finally, there is bound to be a debate about the legal and even ethical propriety of using certain facts as the basis for making predictions. Everyone would agree that race should not be a factor; everyone would probably agree that prior record should be a factor. I certainly believe that it is proper to take into account an offender's juvenile as well as his adult record, but I am aware that some people disagree. But can one take into account alcohol or drug use? Suppose the person claims to be cured of his drinking or his drug problem; do we believe him? And if we do, do we wipe the slate clean of information about these matters? And should we penalize more heavily persons who are chronically unemployed, even if unemployment is a good predictor of recidivism? Some people will argue that this is tantamount to making unemployment a crime, though I think that overstates the matter. After all, advocates of pretrial release of arrested persons, lenient bail policies, and diverting offenders away from jail do not hesitate to claim that having a good employment record should be counted in the accused's favour. If employment counts in favour of some, then obviously unemployment may be counted against others. Since advocates of "bail reform" are also frequent opponents of incapacitation, selective or collective, it is incumbent on them to straighten out their own thinking on how we make use of employment records. Nonetheless, this important issue deserves thoughtful attention.

On one matter, critics of prison may take heart. If Greenwood and the others are correct, then an advantage of selective incapacitation is that it can be accomplished without great increases (or perhaps any increases) in the use of prisons. It is a way of allocating more rationally the existing stock of prison cells to produce, within the constraints of just deserts, greater crime-control benefits. Many offenders—indeed most offenders—would probably have their sentences shortened, and the space thereby freed would be allocated to the small number of high-rate offenders whom even the most determined opponents of prison would probably concede should be behind bars.

Notes

1. Ernest van den Haag, *Punishing Criminals* (New York: Basic Books, 1975), 52–60.
2. Scholars who study incapacitation call the numberr of crimes committed per offender per year free "lambda", or λ. To avoid technical terminology, I will refer to it as the "individual offence rate".
3. Mark A. Peterson and Harriet B. Braiker, *Doing Time: A Survey of California Prison Inmates* (Santa Monica, Calif.: Rand, 1980), vii, 32.
4. Peter W. Greenwood, *Selective Incapacitation* (Santa Monica, Calif.: Rand, 1982), 43–4.
5. Alfred Blumstein and Jacqueline Cohen, "Estimation of Individual Crime Rates", from Arrest Records", 70 *Journal of Criminal Law and Criminology* 581 (1979); Peterson and Braiker, *Doing Crime*, above n. 3, at 40; Marvin Wolfgang *et al.*, *Delinquency in a Birth Cohort* (Chicago: University of Chicago Press, 1972), 206.
6. Andrew von Hirsch, *Doing Justice: The Choice of Punishments*, a report of the Committee for the Study of Incarceration (New York: Hill and Wang, 1976). See also Selection 4.4 below.

3.7

Selective Incapacitation: Some Doubts

ANDREW VON HIRSCH

Prediction research in criminology has, by and large, focused on characteristics of offenders. Various facts about criminals are recorded: age, previous arrests and convictions, social history, and so forth. It is then statistically determined which of these factors are most strongly associated with subsequent offending.[1] The result is a "selective" prediction strategy: among those convicted of a given type of offence, some will be identified as bad risks and others not.

Traditional Prediction Methods

Traditional statistical prediction techniques pursued this selective approach. Generally, they found that certain facts about an offender—principally, previous criminal history, drug habits, and history of unemployment—were to a modest extent indicative of increased likelihood of recidivism.[2]

These techniques did not, however, distinguish between serious and trivial recidivism. Both the offender who subsequently committed a single minor offence and the individual who committed many serious new crimes were lumped together as recidivists. Moreover, the techniques offered no promise of reduced crime rates, as they did not attempt to estimate aggregate crime-prevention effects. Locking up the potential recidivist thus assured only that he or she would be restrained; since other criminals remained at large, it did not necessarily diminish the overall risk of victimization. By the 1970s, these limitations reduced penologists' interest in traditional prediction techniques.

This was published for the first time in the first edition of this book and has been edited somewhat for the present edition.