Punishment: Consequentialism

David Wood*
University of Melbourne

Abstract

Punishment involves deliberating harming individuals. How, then, if at all, is it to be justified? This, the first of three papers on the philosophy of punishment (see also 'Punishment: Nonconsequentialism' and 'Punishment: The Future'), examines attempts to justify the practice or institution according to its consequences. One claim is that punishment reduces crime, and hence the resulting harms. Another is that punishment functions to rehabilitate offenders. A third claim is that punishment (or some forms of punishment) can serve to make restitution to victims, and a fourth is that it can strengthen social values. The paper examines these claims, and finally considers pluralist theories which combine retributive and harm-reductive or utilitarian considerations. (Retributive theories are examined in their own right in 'Punishment: Nonconsequentialism'.)

1. Introduction

Punishment involves the conscious and premeditated infliction of harm (Honderich 4; Boonin 213), and the deprivation of an individual's rights (Galligan 155; cf. Burgh 193). Some go further. Punishment 'cripples people' (Gross 375). It 'commits the state to the ''prima facie'' barbarity of deliberately damaging people's lives' (Gardner 2007: 203). Indeed, the 'punitive side' of the criminal law is 'an extraordinary abomination' (Gardner 1998: 32). Clearly, punishment – at least where it produces such consequences – stands in very strong need of justification.

In conjunction with 'Punishment: Nonconsequentialism', this paper sets out to introduce the important issues raised both by the need for justification and attempts to satisfy it, as canvassed in recent philosophical and jurisprudential literature. This paper is concerned with theories that attempt to justify punishment solely through its consequences, whereas the other considers theories that try to justify punishment in virtue of its innate worth, or some intrinsic characteristic it possesses. This paper also discusses – indeed, starts with – a matter relevant to all theories of punishment, namely the distinction between punishment as a public, state-administered institution (the main concern here), and as a private practice that takes place in, for instance, families, schools, and close-knit communities such as religious orders.

A further paper, 'Punishment: The Future', takes up the question of alternatives to punishment, and considers possible developments.

This, at least, is the intended division of labour. However, theorising about punishment resists neat packaging. Trying to enforce too rigid a demarcation between consequentialist and nonconsequentialist theories produces artificiality (but see Feinberg 2008: 624). Pluralist theories that are basically consequentialist, but include retributive and other constraints on the aims and ends they endorse ('side-constrained' consequentialist theories, as they have been called (Duff 2001: 11–4)), are better included in this paper. Some communicative theories, examined in 'Punishment: Nonconsequentialism' are put forward

not as nonconsequentialist theories, but rather as challenging, indeed 'transcending', the distinction between consequentialist and nonconsequentialist theories (Duff 1996: 3).

Inevitably, many developments must go unmentioned. In particular, no attempt is made to examine problems raised by particular forms of punishment. The paper totally sidesteps the huge literature generated, particularly in the United States, by the continued use of the death penalty. More generally, although undoubtedly '[p]hilosophers usually pay too little attention to sentencing' (Duff 1996: 57), no attempt to start reversing the trend can be made here.

Traditionally, the most influential consequentialist theory of punishment is utilitarianism. It claims, briefly, that punishment is justified according to whether it maximises happiness or utility, as against other available options. However, this paper will tend to discuss, more broadly, harm-reduction. The prevention or reduction of harm, in particular, the very serious harms that criminal conduct causes (murders, rapes, assaults, kidnappings, robberies, burglaries, etc.), is obviously a worthwhile social concern, and does not have to be seen in utilitarian terms (Wood n.1). One certainly does not have to be a utilitarian to endorse harm-reduction – on the contrary, any reasonable moral theorist will do so, and agree that, other things being equal, harm is morally bad, and consequently, the less crime (the fewer murders, rapes, and so on) the better. And one can say this making all due allowance to the controversial, 'contested' nature of the concept of crime. Obviously, the conduct in any particular jurisdiction that is criminalised, is not necessarily the conduct that *ought* to be criminalised.

Whereas harm-reduction is a general or community-oriented consequentialist aim (if 'community' sounds parochial, consider the global community), the paper turns next to the offender-related aim of rehabilitation. Many consider rehabilitation as the third major justification of punishment, in addition to harm-reduction and retribution, and some even its sole, genuine purpose. Following rehabilitation, the paper examines the victim-related aim of restitution (i.e., as a form of punishment, rather than a replacement for it). Perhaps surprisingly, restitution has not achieved the prominence of rehabilitation. Following these respectively offender-related and victim-related aims, the paper returns to a further purported general aim of punishment, namely strengthening social values.

However, it is not just 'monistic' theories of punishment that will be considered. A major trend in theorising about punishment has been the development of pluralist theories which combine both utilitarian or harm-reductive and retributive considerations. The earliest such theories, Rawls' (1955) and Hart's (1959–1960) (but see Ross 59–64; Quinton 140–2), had a dual-level structure, proposing utility-maximisation, and beneficial consequences more generally, as the basic aim of punishment, while imposing retributive constraints, and constraints of fairness more generally, on their pursuit. This paper finally considers such pluralist theories. (It may assist the reader to read first the section on 'Retribution' in 'Punishment: Nonconsequentialism'.)

As will be seen, none of the theories of punishment considered in this paper is without difficulties. Perhaps, the most fashionable approach is that taken by pluralist theories of punishment. Even these theories suffer from important drawbacks, however, which are discussed below.

2. The Concept of 'Punishment'

Before considering various attempts to justify punishment, it is best to pause briefly to consider what punishment *is*. The orthodox analysis of the concept of 'punishment' is that developed by H. L. A. Hart, who drew on the work of Antony Flew and Stanley

Benn. According to the Flew–Benn–Hart definition of 'punishment', as it is sometimes called (although Hart does not, in fact, claim to offer a definition of 'punishment', but only a description of its central or standard case), five conditions must be satisfied for an act to be an instance of 'punishment': '(i) it must involve pain or other consequences normally considered unpleasant; (ii) it must be for an offence against legal rules; (iii) it must be for an actual or supposed offender for his offence; (iv) it must be intentionally administered by human beings other than the offender; (v) it must be imposed and administered by an authority constituted by a legal system against which the offence is committed' (Hart 1959–1960: 4–5; cf. Flew 293–4; Benn; Scheid).

A number of philosophers have supplemented this definition of 'punishment'. Joel Feinberg, for instance, has argued that, in addition to satisfying these conditions, it must also express 'attitudes of resentment and indignation, and...judgments of disapproval and reprobation, on the part either of the punishing authority himself or of those 'in whose name' the punishment is inflicted' (Feinberg 1965: 400). The expression of such 'attitudes' and 'judgments' inflicted on the offender, are sometimes referred to as the 'censure' or 'condemnation' and 'sanction' or 'hard treatment' or 'deprivation' elements of punishment, respectively (von Hirsch 1993: 9–13; Husak 2008b: 92–4; see 'Punishment: Non-consequentialism': 'Expression' and 'Communication and Retribution Again').

3. Public and Private Punishment

It is punishment as a public, state-administered institution (Lacey) that is the subject of these papers, rather than punishment as a private practice. This is not to deny that many of the same issues arise with both (Hildebrandt). However, it is the state which imposes the most severe forms of punishment, imprisonment and even death.

Even if a moral case for punishment can be made out in the abstract, there is the further step of showing that administering an official system of punishment is a fitting role for the state to perform (Duff 1992, 1996: 9). In particular, insofar as the justification of punishment lies in considerations of desert or retribution, the question is whether this is an end the state should pursue (Galligan 153).

The problem for liberal political theory is that a state that pursues retributive ends appears to engage in legal moralism, the view (on one account) that '[i]t can be morally legitimate to prohibit conduct on the ground that it is inherently immoral, even though it causes neither harm nor offense to the actor or others' (Feinberg 1985: 4; see also Duff 2007: 47–9, 83–9). (More extreme − or just thorough-going − liberals would delete the reference to 'offence'.) The problem is particularly acute where the state pursues these ends through punishment rather than other coercive means, such as taxation, let alone in non-coercive ways, for instance, public awareness campaigns. Unless punishment is to be justified merely instrumentally, for instance, as a means of reducing overall harm or increasing general happiness, the objection is that, in punishing, whatever further non-instrumental benefits this may provide, the state breaches Mill's harm principle, that 'the only purpose for which power can be rightfully exercised over any member of a civilised community against his will, is to prevent harm to others' (Mill, ch. 1, para 9; cf. Hampton 1994).

Indeed, retributive punishment presents an even greater problem for libertarian theory, since retribution does not appear to be among the permissible functions of the minimal state, these being limited, according to a well-known statement, to 'protection against force, theft, fraud, enforcement of contract, and so on' (Nozick 1974: ix). It seems surprising, then, that libertarians − notably Nozick (1981: 363–397) − tend to be strong

retributivists. (Randy Barnett, as seen below (see 'Restitution') is an obvious exception.) However, this list of the minimal state's legitimate functions appears to leave them only with consequentialist considerations to justify punishment (see Nozick 1974: ch. 4).

Perhaps a further right should be added to Nozick's list, namely that of crime victims to see perpetrators justly punished. Hampton's view, which sees the point of a retributive response as repairing the victim's 'moral injury' (Hampton 1992: 1666–85; and see 'Punishment: Nonconsequentialism': 'Retribution'), appears to support this. On the other hand, note Nozick's dismissal of a 'utilitarianism of rights' that merely replaces maximising happiness with minimising violations of rights, as 'the relevant end state in the utilitarian structure' (Nozick 1974: 28–30).

Punishment as a state institution creates further problems – not just for liberals and libertarians, but political theorists of all persuasions – and irrespective of whether retributive, or some other justification of punishment, is sought. Like any system of power and authority (which forms of private punishment may be as well, if generally on a smaller scale), a system of state punishment is liable to abuse and corruption (Husak 2000: 975, 2008a: 109). In the case of imprisonment, the punishment is not just the official loss of liberty, but 'typically' as well, 'impositions which are injurious, brutalising, often painful and deeply degrading' (Galligan 159; cf. Golash 2–4; Kleinig 273; Signler). There is the inevitable moral price of the susceptibility of a system of criminal law to wrongful convictions and hence punishments (Husak 2000: 975; Husak 2008: 109). Obviously, this is most alarming in the case of the death penalty, where DNA evidence has demonstrated a shocking incidence of conviction and (in the United States) execution of perfectly innocent persons. But there is also the massive commitment of human and economic resources that a system of state punishment involves (Husak 2000: 974–5, 2008a: 109, 1992: 451; cf. Berman 266). Why should a government pursue punishment, rather than health, education, or improving employment opportunities – especially given the further benefit of helping ameliorate the social causes of crime (Golash 155–60; Hampton 1992: 1700–1; and see 'Punishment: Nonconsequentialism', towards the end of 'Retribution'), and so reduce not just the incidence of crime, but likewise of punishment?

4. Harm-Reduction

The claim that punishment reduces harm by preventing or reducing crime raises numerous issues. Most obvious is the question of what counts as harm for its purposes, of the types of harm to be reduced through criminalising the responsible conduct. Harm–reductivists may obviously differ on how they identify different types of harm, and rank them (and individual instances of harm) for relative seriousness. Consider physical and psychological harm, harm to property and financial interests, and harm to public interests, such as the interest in a clean environment and in a functioning social and economic infrastructure. Consider also the moral harms created by the criminal law itself – not just cases of wrongful convictions and hence punishment, and of disproportionately harsh punishment, but forms of stigmatisation and ostracising treatment, and consequential material harms and losses (for instance, from discrimination in employment and rental markets). Consider further the more intangible social costs of 'so ambitious, intrusive and burdensome an institution as the criminal sanction' (von Hirsch 1993: 14; von Hirsch and Ashworth 24).

Another issue is the relation between the harm and the responsible conduct. Must the conduct cause the harm, or can a weaker connection suffice for criminal liability (and if so, under what conditions), such as the conduct's 'inducing', 'encouraging', or even 'not

discouraging' the harm? How far should the ambit of criminal liability reach beyond immediate perpetrators, to accomplices and instigators, and those more broadly who offer material or moral support?

A further question is how the relative seriousness of different types and instances of harm is to be gauged (von Hirsch and Jareborg 1991, in von Hirsch and Ashworth: Appendix 3). What makes one type or instance of harm more or less serious than another? One complication is that criminalised conduct may be beneficial, indeed, morally obligatory. Consider euthanasia (legally murder), or a popular insurrection against a tyrannical regime (legally treason).

Harm-reductivists may differ on whether they regard a particular type of harm as serious enough to warrant criminalising the responsible conduct – that is, where it is sufficiently culpable, and other conditions are satisfied, such as the crime's occurring within the relevant jurisdiction, and the defendant's being old enough to be criminally liable: '[t]he problem is to develop criteria for harmfulness and culpability that are more illuminating than simple intuition' (von Hirsch and Jareborg, in von Hirsch and Ashworth 186).

More precisely, harm-reductivists may differ on the amount of harm of each relevant type required to warrant criminalising the responsible conduct (again, subject to its meeting other conditions) – and also on how relatively seriously the conduct is to be regarded, on whether different categories or bands of crime seriousness should be recognised, and if so, the criteria for distinguishing them: 'the jurisprudence of crime seriousness has received comparatively little attention' (von Hirsch and Jareborg, in von Hirsch and Ashworth 186; see also 'Punishment: The Future': 'Looking Ahead').

Moving from harm to its reduction, both the material costs of criminal law intervention must be taken into account, the social resources that could be used elsewhere, and more importantly, the moral costs, such as the stigmatisation of criminal conviction, and most worryingly, the risk of wrongful conviction. The aim must be to reduce overall harm or loss – of the crime, the crime-reductive measures themselves, and the broader costs of the institutional arrangements to implement them (Duff 1996: 3).

Harm-reductivists may obviously differ on the extent to which they think an aim can be effectively pursued by a particular means, for instance, deterrence (whether specific, directed at actual offenders, or general, aimed at potential offenders), incapacitation or rehabilitation. (Indeed, harm-reductivists differ on what they hold *are* the available means). For instance, one harm-reductivist may place more weight on deterrence, a second (perhaps with a more optimistic view of human nature) on rehabilitation, and a third, on the combination of the capacity for prudential and moral motivation, merging them into the notion of 'general prevention' (Andenaes, esp. ch. I; von Hirsch and Ashworth 14–17).

Harm-reductivists could differ equally on the possibility and availability, relative effectiveness and moral justifiability, not just of harm-reductive punishment means, but harm-reductive measures, such as capital punishment, imprisonment, fines and other monetary penalties (for instance, restitution orders), community service and licence-cancellation. For instance, one harm-reductivist may reject capital punishment on the grounds that it is no more effective a deterrent than lengthy prison terms, a second accept it as a deterrent on the basis of different evidence, a third (at least, a pluralist, rather than pure, harm-reductivist) reject it on moral grounds, as contrary to human rights and human dignity (irrespective of whether it is an effective deterrent), and a fourth hold that it should be reserved for exceptional cases, such as the most heinous murders, or the murder of particular classes of victims, for instance, law-enforcement officers.

Harm-reductive means can usually draw upon different punishment measures. Any form of hard treatment can in principle serve deterrence purposes, as it can serve retributive purposes. Alternatively, a means may not be available because of the lack of a measure to implement it. The town gaol, the only viable place of incarceration, may have burnt down, or even more fancifully, capital punishment may not be available because no-one is prepared to take on the role of executioner.

Likewise, many punishment measures are capable of implementing more than one harm-reductive means. For instance, imprisonment serves the end of harm-reduction through deterrence, incapacitation, and according to optimists, rehabilitation. Licence-cancellation obviously may serve the means of both deterrence and incapacitation.

The harm-reductivist may be accused of not putting forward an account of punishment as a social institution or practice in its own right (note Feinberg's description of the criminal process as 'a great moral machine' (Feinberg 1987: 260)), but merely as another harm-reductive means, with no particular moral character or qualities. He may simply accept this characterisation of punishment, responding that the only real question is the extent of the required hard treatment. But this may be altogether too abrupt. The standard objection to 'pure' harm-reductivism remains (and equally, to purely utilitarian approaches), that the uncompromising pursuit of harm-reduction (or utility) may require punishment of the innocent (Boonin 31–52), and the disproportionately harsh punishment of the guilty (Boonin 54–8). It is not enough to appeal to the lack of likelihood of their arising in practice, critics claim − it is the possibility in principle that is objectionable (Lessnoff 142). But unless the harm-reductivist can adequately answer this objection, he must retreat from a pure to a pluralist position, by acknowledging constraints on the pursuit of harm-reduction through punishment, restricting its imposition to the guilty, and its severity according to the seriousness of the offender's crime or the extent of his culpability. The paper considers such pluralist retributive and harm-reductive or utilitarian theories later. (See 'Pluralist Retributive and Harm-Reductive or Utilitarian Theories'.)

5. Rehabilitation

Rehabilitation obviously appears a worthy, forward-looking, general social aim and policy concern. The pertinent issue is whether it has sufficient, and indeed appropriate, moral significance to justify the use of punishment (indeed, any coercive measures), and if so, punishment of what degree of 'hardness'. Corresponding to the three questions above, concerning harm-reduction as a punishment aim (and indeed, any consequentialist punishment aim), are the three following questions: *first*, of what actually rehabilitates, what are the available rehabilitative measures (or combination of measures); *secondly*, of which such measures are the most effective, which best rehabilitate; and *thirdly*, of the moral justifiability of the most effective, available measures, whether moral constraints must be placed on their use, either general constraints on their use at all, or specific constraints on their use in particular circumstances or under particular conditions. Less dramatically, there may also be questions of possible tradeoffs between effectiveness and moral justifiability, unless it is held that the relevant moral considerations are, in virtue of being such, overriding.

Rehabilitation had its heyday in the 1950s and 1960s (Allen) a period of considerable faith in people's capacity (rather misplaced as it turned out), with the aid of the then rapidly developing human and social sciences, to solve individual and social problems: '[t]he hope of providing humane and constructive treatment for offenders had a powerful hold on the public imagination' (Golash 18).

However, the 'medicalisation' of criminality, and the push to replace punishment with treatment (Wootton ch. 2), led to the extensive use of forms of therapy and other techniques that often were far more harsh than punishment based upon – and more importantly limited by – retributive considerations. Far harder treatment was liable to be imposed on individuals under the guise that they were not being punished at all. Stanley Kubric's provocative film *A Clockwork Orange*, parodying the use of aversion therapy, was one dramatic example. But there were even more disturbing real life examples (Golash 190).

Retributive-minded critics appealed to the idea of 'just deserts' (von Hirsch 1976) to limit the extent of harsh, but supposedly rehabilitative, treatment. According to this idea, offenders ought to receive only what they morally *deserve*. Typically, for such critics, this is punishment whose severity is proportionate to the seriousness of the crime. To impose harsher punishments on offenders in the name of treatment or therapy is to give offenders more than their 'just deserts'. Some philosophers of punishment, such as Herbert Morris (but see most notably G. W. F. Hegel 126–7), even suggested that offenders have 'a right to be punished', which is entailed by their right to be treated as responsible moral agents (Morris, see also Deigh). The thought here is that to refrain from punishing offenders, for instance, by imposing treatment or therapy on them instead, is to fail to treat them as responsible moral agents, capable of responding to the moral message communicated by punishment. This use of the slogan 'a right to be punished' must be contrasted with the stronger, but more dubious idea of the right of the offender to have his wrongdoing annulled (see the second retributive view in 'Punishment: Nonconsequentialism': 'Retribution'), and hence the moral burden of the crime lifted from him. Underlying the notions of 'just deserts' and 'a right to be punished', and perhaps most influential of all, was the growing recognition and influence of liberal ideals of human dignity and individual rights.

However, the fate of the 'rehabilitative ideal' was sealed more at the level of practice than theory, with the growing realisation that rehabilitation was failing in its own terms. If it was not quite the case that 'nothing worked' (Martinson), it was difficult to see what worked and why, which programmes and methods produced worthwhile results, and for what reasons. There was evidence of some success with specialised programmes directed at specific offender sub-groups (Rex; Zedner 97), but the incarceration required for most rehabilitative programmes was itself profoundly 'dehabilitative'.

The 'nothing works' thesis also had resonance with deterrence, although there the problem was more one of trying to determine the marginal effectiveness of different sentencing measures, in particular, capital punishment as against imprisonment, and custodial as against non-custodial sentences. In addition, sharing with retribution the requirement of hard treatment, deterrence was able to piggy-back on the revival of retributive justifications of punishment (if not to the extent of justifying disproportionate sentences, beyond the ceiling imposed by the offender's 'just deserts'). Deterrence therefore was better placed than rehabilitation in the face of empirical doubts.

The failure of 'minimal' or 'thin' rehabilitation, as it could be called – rehabilitation as evidenced merely by 'non-recidivism', and hence 'operationally' indistinguishable from specific deterrence – suggests turning to richer forms of rehabilitation, concerned more deeply and ambitiously with the moral reform of offenders, indeed, their moral education, with their becoming better citizens and community members. Strictly, this is '*pre*habilitation' rather than rehabilitation, in that it is intended to result in a stronger capacity to resist criminal impulses or motivation, or better still, prevent their arising in the first place (see 'Strengthening Social Values').

The important issue here (apart from paternalistic dangers: Zedner 98) concerns the standards by which persons subjected to such measures are to be judged rehabilitated or 'prehabilitated'. Are such aims to be assessed according to the prevailing, positive standards of the community in question, the standards of 'the morality actually accepted and shared by a given group' (Hart 1963: 20), or alternatively according to critical moral standards, 'the general moral principles used in the criticism of actual social institutions including positive morality' (*Ibid*)? Insofar as the aims are to be assessed according to the latter, the question arises of what *are* the applicable moral standards. Are they, for instance, to be utilitarian or some other consequentialist standards, nonconsequentialist standards, or some combination of the two types of standards?

6. Restitution

Restitution is an aim of private law, obviously of the law of restitution or unjust enrichment, but more generally of tort law, in the broad sense that encompasses making good a loss, and so goes beyond the technical sense of returning – or 'disgorging' – a gain. Restitution in this broad sense is a quite general social policy concern, which may be supported by taxation and insurance. It has increasingly been recognised as an aim of the criminal law, through restitution orders as alternatives to fines, and the establishment of criminal compensation tribunals.

Although restitution is considered in this paper on consequentialist theories of punishment, it is not purely or exclusively forward-looking – to look forward to what must be done to restore the victim to the 'status quo ante', it must look back at what that position was, and determine the effect of the crime on the victim. Perhaps a parallel point could be made in relation to rehabilitation – although 'rehabilitation' is a misnomer where, as is often the case, the 'rehabilitee' never was in the supposed 'habilitated' state, and hence there is no question of any return to such a state. This point, about including both forward- and backward-looking considerations, is perhaps more plausible in relation to restorative justice (see 'Punishment: The Future': 'Condemnation without Hard Treatment'). However, while there is danger in trying to push the point too far (otherwise, one could even say that retributivism is likewise forward, as well as backward-looking, in that it looks forward to a situation where just retribution is inflicted on the offender), one must not suppose that the distinction between consequentialist and nonconsequentialist theories is necessarily clearcut. The warning in the 'Introduction' that trying to enforce too rigid a demarcation between consequentialist and nonconsequentialist theories produces artificiality, needs to be kept in mind.

That restitution has not achieved the status of harm-reduction and rehabilitation amongst consequentialist punishment aims (indeed, until recently, it was virtually ignored by punishment theorists), may reflect a general tendency to ignore victims of crime, notwithstanding the best efforts of vociferous victims' rights groups, and the rapid development of victimology as a sub-discipline of criminology. However, there seems no reason to treat 'material', as opposed to 'moral' restitution (through annulling the victim's moral harm as a result of the crime, as it could be put), as anything more than a peripheral aim of the criminal law, however, important it may be in private law (see the discussion of restorative justice in 'Punishment: The Future': 'Condemnation without Hard Treatment'). But if punishment, and the criminal law more broadly, is to be concerned as much with victims as offenders, moral restitution must be given as important a role as moral rehabilitation.

Barnett has proposed that restitution should replace punishment as a new paradigm, or less radically, at least supplement punishment (Barnett; Boonin: 216, 224–6). The latter is hardly contentious (Boonin 216). However, material restitution is scarcely important enough to play such a role. Only the far more ambitious and challenging notion of moral restitution, mirroring that of moral rehabilitation of the offender, has the potential to be a new paradigm, or even a major punishment aim. But Barnett discusses material rehabilitation only. (See further Boonin 216–8, the discussion of restorative justice in 'Punishment: The Future', and for development of the idea of restitution, Boonin ch. 5.)

In seeing an appropriate retributive response as compensating and repairing the victim's moral injury, Hampton could be seen as a proponent of a moral restitutionary theory of punishment. Certainly, such a view serves as a much needed antidote to the preoccupation of theorists such as Duff (see 'Punishment: Nonconsequentialism': 'Communication and Retribution Again'), with the offender and his moral reform.

7. Strengthening Social Values

It is sometimes asserted that punishment is justified because it strengthens social values. The key claim in relation to this justification of punishment is that promoting strong social values is an intrinsically worthwhile end, irrespective of any crime-reductive benefits, or the contribution it makes to social cohesion and solidarity (Gould and Scholz; Scholz), or to use Durkheim's term, the 'collective conscience' (Durkheim: 120–2, 226–35, 240–2, 329–30). Whereas rehabilitation and restitution attempt to justify punishment by improvements (whether material or moral) to particular classes of individuals – offenders and victims respectively – this aim, like harm-reduction, attempts to justify punishment through how it benefits society generally.

An immediate question is whether strengthening social values is a worthwhile aim irrespective of what the values are – or only subject to their meeting some critical moral standard. Is there merely a formal or functional relation between the aim and the relevant social values, or more importantly, a substantive relation that brings in their content? Why strengthen a society's values simply because they *are* its values?

Some suggest that a society should take an active role in developing its values, adopt a positive programme of moral education and reform, indeed, that it is the role of the state, as the society's moral representative, to implement such a programme. Consider, at a relatively uncontentious level – appeal being more to prudential than moral motivation – publicly-funded advertising campaigns exhorting people (to take some recent examples from Australian television) to drive safely, stop smoking, drink responsibly, and even to be alert to overhead power lines when boating. Into more contentious territory, are campaigns to encourage racial and cultural tolerance, and acceptance of religious and other minorities.

Cynics may ridicule the notion of the moral development of an individual, and likewise the parallel notion of the moral development of a society or community. And both notions are certainly open to abuse, the latter even more so than the former. There are obvious paternalist dangers in coercing or manipulating individuals to hold 'proper' values. Certainly, any legitimate process of rational persuasion must respect individuals' moral autonomy and independence, must 'take seriously the distinction between persons' (Rawls 1999: 24). Nevertheless, it is perfectly meaningful to talk of a society as a whole, and not just individuals and particular social institutions and organisations, with better and worse values, however, contentious it may be to decide which values are better and which worse.

Consider a society with a strong sense of justice (Rawls 1963, 1999: ch. VIII) and deep commitment to humanitarian values and human rights (in substance, irrespective of the terminology used), that genuinely respects the equal worth of each individual person, and palpably demonstrates a sincere, non-paternalistic, concern that the basic needs (liberally interpreted) of all its members are satisfied, and that all have the realistic opportunity for a meaningful and fulfilling life.

But even if these ideals are enthusiastically endorsed, it is not easy to see how punishment could usefully help to implement such a moral agenda, and in any case, how this could justify such punishment. Punishment to maintain a minimal level of social life and survival is one thing – but punishment for such aspirational purposes, seems quite another. On the other hand, however, this aim may well support institutions and practices very different from punishment as it presently exists.

8. *Pluralist Retributive and Harm-Reductive or Utilitarian Theories*

A number of consequentialist punishment aims have been examined so far – harm-reduction, rehabilitation, restitution and strengthening social values. For some time, however – at least since the publication in the 1950s of two seminal articles, John Rawls' 'Two Concepts of Rules' (Rawls 1955) and H. L. A. Hart's 'Prolegomenon to the Principles of Punishment' (Hart 1959–1960) – most punishment theorists (Husak 1992: 452) have been pluralists of one kind or another (other terms include 'dualist', 'hybrid', 'mixed', 'synthetic', 'compromise', 'intermediate' and 'middle-of-the-road'), holding that no single type of consideration, such as utility, harm-reduction, desert or retribution, can offer a satisfactory justificatory account of punishment on its own. Instead, they have been engaged in the project of trying to show how such considerations, and more recently, other considerations like the moral reform of the offender, can be coherently combined.

For instance, C. L. Ten adopts 'a pluralist approach to punishment, believing that in the practice of punishment we are often confronted with a number of different considerations, each not reducible to the others, and each having a contribution to make' (Ten 6). Ten accordingly identifies as his main target, 'any theory of punishment, whether it be utilitarian or retributive, which seeks to justify punishment in terms of just one ultimate value' (*Ibid*). Nicola Lacey defends a conception of punishment which is 'pluralistic...in that it identifies punishment as pursuing and respecting a set of different values which cannot necessarily be reduced to each other, or into some overriding common currency such as general utility' (Lacey 199). Hampton describes herself as 'a convert to the pluralist approach to [the] justification of punishment' (Hampton 1992: n.2).

This is not to say that all theorists have succumbed to pluralism. Michael Moore is one of the leading advocates of 'pure retributivism', according to which retribution (and retributive considerations alone) not only sets limits on when punishment may be justifiably imposed, but also provides (and alone provides) the state with a positive reason for punishing offenders. Moore's theory, as a 'pure retributivist' theory, does not require harm-reduction or any other consequentialist good to serve as part of, or be included in, what Hart calls the 'general justifying aim' of punishment (Hart 1959–1960; and see shortly below). As Moore states: '[r]etributivism...is truly a theory of justice such that, if it is true, we have an obligation to set up institutions so retribution is achieved' (Moore 91). On the other hand, pure harm-reductivists and utilitarians (who concede no constraint on the pursuit of harm-reduction or utility respectively) seem more numerous than pure retributivists, certainly amongst economic theorists.

However, it is with Rawls and Hart that we should start. Rawls' rule-utilitarian theory distinguishes between the justification of a practice, and of a particular act falling under it. The practice of punishment is to be justified by appeal to utility, whereas particular instances of punishment are to be justified by appealing to retributive, and maybe other considerations, for instance, rehabilitation. Interestingly enough, although Rawls famously went on to reject utilitarianism, and with it his own version of rule-utilitarianism, and develop a general theory of justice (indeed, a general body of normative political philosophy) to oppose the utilitarian tradition, he did not revisit the topic of punishment or revise his theory in 'Two Concepts of Rules' (Rawls 1955).

Hart's theory similarly involves a distinction between two stages, first, as just mentioned, of 'the general justifying aim' of punishment, and secondly of distribution, which divides into two further questions, of 'liability' (who to punish?) and 'amount' (how much to punish?) (Hart 1959–1960). He holds that utilitarian considerations are relevant to the first issue, of why have an institution of punishment at all, whereas retribution, and considerations of justice and fairness more generally, are relevant to the second issue, and function to restrict the imposition of punishment to those properly convicted of a criminal offence, and its severity according to the seriousness of their crimes or the extent of their culpability. Hart makes this move to avoid an important criticism of 'pure' utilitarian theories of punishment, namely, that such theories imply that it is morally permissible to punish the innocent, or to punish the guilty disproportionately, if doing so maximises utility, or brings about some other consequentialist good. According to Hart's theory of punishment, the retributive considerations that are relevant to the question of liability ensure that we never punish the innocent, and that the retributive considerations that are relevant to the question of amount ensure that we never impose disproportionately harsh punishment on the guilty.

Three objections to the endeavour of combining utilitarian or harm-reductive and retributive or desert considerations in a pluralist theory of punishment are worth noting. First, in confining retribution to the second level, and so restricting it to grounding constraints on the pursuit of utilitarian or harm-reductive goals, such pluralist theories are accused of not treating it seriously enough. Rather than merely another crime-reductive institution or practice, punishment is special in having retribution as its pivotal notion: '[t]o many people there is a deeply embedded intuition that part at least of the general purpose of criminal justice is the correction of wrongs, quite apart from any consequential contribution to crime control' (Galligan 151).

Ten similarly holds that retributive considerations can provide 'positive reasons in support of punishment' (Ten 84). According to his 'Compromise Theory', there are separate retributive and utilitarian conditions on punishment, and both 'enter into the justification of every case of punishment. Each condition is necessary, and the two conditions are jointly sufficient for the justification of punishment' (Ten 79). (Although Ten modifies both conditions to yield his 'Revised Compromise Theory' (Ten 80), the changes do not affect the relation between them.)

Moreover, there seems no way of accommodating Galligan's 'deeply embedded intuition' within a 'structured hybrid' (Gardner 1998: n. 3) or 'split-level' (Wood 304–7) approach. Keeping the two levels separate requires confining retributive considerations to the second level, and according to the objection, this is to assign them a lesser role. Taking the above intuition seriously requires a different type of pluralist theory, one that recognises retribution as itself a goal or an end of punishment, and not just a constraint on the pursuit of other punishment goals or ends, such as utility-maximisation or harm-reduction. In short, this intuition requires retribution as a positive reason for punishing

(Mackie 4), and not just as grounding or supporting constraints on other reasons for punishing, such as harm-reduction.

The second objection is that 'structured hybrid' or 'split-level' pluralists fail to establish the connection between the two levels – in the case of Rawls, the rules of the rule-utilitarian theory in question and individual acts, and with Hart, the general justifying aim of punishment and the principle of distribution. To focus on Hart's theory, this objection holds that it is arbitrary to restrict appeal to utilitarian or harm-reductive considerations to the level of the general justifying aim, and retributive considerations to that of distribution. To move from one type of consideration to the other in this way is like changing horses mid-stream: '[s]o the reasons given for punishing at all, and the guiltiness of the offender float, as it were, side by side, without being able to make contact' (Galligan 150). The split-level pluralist appears opportunistic, appealing alternatively to harm-reductive or utilitarian considerations, and retributive or desert considerations, as convenient. It is incumbent on him to show that the questions which are to be answered appealing respectively to harm-reductive or utilitarian and retributive or desert considerations, can genuinely be kept separate.

Thirdly, while pure retributivists hold that retribution alone can be a sufficient justification for punishment, more pluralistically minded retributivists object that it has insufficient weight on its own to justify punishment as a social, and indeed state institution. Retributive considerations may offer a partial justification, but it is unduly moralistic to create a system of punishment for the sole purpose of exacting retribution from offenders. There is insufficient reason (some would say no reason at all) to systematically pursue retribution as an end in its own right, let alone for the state to do so through a system of criminal law and punishment. (Contrast this with the pluralist view that even though retribution does not by itself offer a sufficient justification of punishment as a social institution, it may do so once harm-reductive, and possibly other consequential benefits, are taken into account (Wood 307).)

Such criticisms notwithstanding, some commentators have sought to modify Hart's characterisation of the 'general justifying aim' of punishment in developing their own accounts. Lacey, for instance, broadens Hart's utilitarian understanding of this aim to include the protection of the values of autonomy and welfare (Lacey ch. 6; Wood 314–7). The general justifying aim in John Braithwaite's and Philip Pettit's republican theory is 'the maximisation of the dominion of individual people' (Braithwaite and Pettit 54). Alvin Goldman goes much further in proposing desert as 'the central justifying factor' in punishment, with utilitarian justification required 'for the infliction of punishment in individual cases', so acting as 'a limitation on the desert principle' (Goldman 55). John Tasioulas replaces Hart's utilitarian 'general justifying aim' with that of the communication of censure for wrongdoing, conceived as a formal end, but one which encompasses the substantive ends of retributive justice, mercy and crime prevention (Tasioulas 285).

Indeed, inverting the order of the two levels has become something of a new orthodoxy, with the development of what could be called 'communicative-retributive' theories (see 'Punishment: Nonconsequentialism': 'Communication and Retribution Again'). Interestingly, Douglas Husak suggests that retributive justice should be the central goal or aim of punishment, with crime-reduction off-setting three major drawbacks with any system of state punishment ('Public and Private Punishment', final paragraph). These drawbacks are the inevitability of wrongful convictions, the liability as with any system of authority to the abuse of power, and sheer cost, the consumption of resources which could be devoted to other social purposes (Husak 2000: 974–5, 1992: 451, 2008a: 109).

9. Conclusion

Punishment involves deliberately harming individuals, and hence stands in very strong need of justification. This paper examined attempts to justify punishment that appeal to its claimed beneficial consequences – whether to offenders (rehabilitation), victims (restitution) or the community generally (harm–reduction, and also strengthening social values). The paper also considered pluralist theories that attempt to combine harm–reductive or utilitarian considerations on the one hand, and retributive considerations on the other. Two further papers consider respectively nonconsequentialist theories of punishment ('Punishment: Nonconsequentialism'), and the future of punishment as a state institution as well as alternatives to punishment ('Punishment: The Future').

Short Biography

David Wood teaches legal philosophy and criminal law in the Law Faculty, University of Melbourne. He is co-editor of *Essays on Philosophy in Australia* (Kluwer, 1992), author of *Judicial Ethics: A Discussion Paper* (Australian Institute of Judicial Administration, 1996), and a co-author of *Criminal Law in Australia* (Federation Press, 2006). He has published on punishment and sentencing, judicial ethics, business ethics, constitutional theory and legal education, in such journals as *Australasian Journal of Philosophy*, *Ethics*, *Criminal Law Review*, *Journal of Business Ethics*, *Oxford Journal of Legal Studies*, *Scandinavian Studies in Law*, and *Utilitas*.

Note

* Correspondence: Faculty of Law, University of Melbourne, Victoria 3010, Australia. Email: david.wood@unimelb.edu.au.

Works Cited

Allen, Francis A. *The Decline of the Rehabilitative Ideal*. New Haven: Yale University Press, 1981.
Andenaes, J. *Punishment and Deterrence*. Ann Arbor: University of Michigan Press, 1982.
Barnett, Randy. 'Restitution: A New Paradigm of Criminal Justice.' *Ethics* 87 (1977): 279–301.
Benn, S. I. 'An Approach to the Problems of Punishment.' *Philosophy* 33 (1958): 325–41.
Berman, Mitchell N. 'Punishment and Justification.' *Ethics* 118 (2008): 258–90.
Boonin, David. *The Problem of Punishment*. Cambridge: Cambridge University Press, 2008.
Braithwaite, John and Philip Pettit. *Not Just Deserts: A Republican Theory of Criminal Justice*. Oxford: Oxford University Press, 1990.
Burgh, R. W. 'Do the Guilty Deserve Punishment?' *Journal of Philosophy* 94 (1982): 193–210.

Deigh, John. 'On the Right to be Punished: Some Doubts'. *Ethics* 94 (1984): 191–211.

Duff, R. A. 'Penal Communications: Recent Work in the Philosophy of Punishment.' *Crime and Justice: A Review of Recent Research*. 20 Ed. Michael Tonry. Chicago: University of Chicago Press, 1996. 1–97.

Duff, R. A. *Answering for Crime: Responsibility and Liability in the Criminal Law*. Oxford: Hart, 2007.

Durkheim, Emile. *The Division of Labour in Society*. Trans. W. D. Halls. Basingstoke: Macmillan, 1984.

Feinberg, Joel. 'The Expressive Theory of Punishment.' *Monist* 49 (1965): 397–423.

——. *Offence to Others: The Moral Limits of the Criminal Law, Vol. Two*. Oxford: Oxford University Press, 1985.

——. 'Some Unswept Debris from the Hart-Devlin Debate.' *Synthese* 72 (1987): 249–75.

——. 'The Classic Debate.' *Philosophy of Law*. 8th ed. Ed. Joel Feinberg, Jules Coleman. Belmont: Thomson, 2008. 624–9.

Flew, Antony. 'The Justification of Punishment.' *Philosophy* 29 (1954): 291–307.

Galligan, D. J. 'The Return to Retribution in Penal Theory.' *Crime, Proof and Punishment: Essays in Memory of Rupert Cross*. Ed. C. F. H. Tapper. London: Butterworths, 1981. 144–71.

Gardner, John. 'Crime: in Proportion and Perspective.' *Fundamentals of Sentencing Theory: Essays in Honour of Andrew von Hirsch*. Ed. Andrew Ashworth, Martin Wasik. Oxford: Clarendon, 1998. 31–52.

——. 'The Functions and Justifications of Criminal Law and Punishment.' *Offences and Defences*. John Gardner. Oxford: Oxford University Press, 2007. 201–21.

Golash, Deirdre. *The Case Against Punishment*. New York: New York University Press, 2005.

Goldman, Alvin. 'The Paradox of Punishment.' *Philosophy and Public Affairs* 9 (1979): 42–58.

Gould, Carol C. and Sally J. Scholtz, eds. 'Solidarity.' Special issue of *Journal of Social Philosophy* XXXVIII.1 Oxford: Blackwell (2007). 1–203.

Gross, Hyman. *A Theory of Criminal Justice*. New York: Oxford University Press, 1979.

Hampton, Jean. 'Correcting Harms versus Righting Wrongs: The Goal of Retribution.' *UCLA Law Review* 39 (1992): 1659–702.

——. 'Liberalism, Retribution and Criminality.' *In Harm's Way: Essays in Honour of Joel Feinberg*. Ed. Jules L. Coleman, Allen Buchanan. Cambridge: Cambridge University Press, 1994. 159–82.

Hart, H. L. A. 'Prolegomenon to the Principles of Punishment.' *Proceedings of the Aristotelian Society* 60 (1959–1960): 1–26.

Hart, H. L. A. *Law, Liberty and Morality*. Oxford: Oxford University Press, 1963.

Hegel, G. W. F. *Elements of the Philosophy of Right*. Trans. N. B. Nisbet, ed. Allen W. Wood. Cambridge: Cambridge University Press, 1991.

Hildebrandt, Mirelle. 'Justice and Police: Regulatory Offences and the Criminal Law.' *New Criminal Law Review* 12 (2009): 43–68.

von Hirsch, Andrew. *Doing Justice: The Choice of Punishments (Report of the Committee for the Study of Incarceration)*. New York: Hill and Wang, 1976.

——. *Censure and Sanctions*. Oxford: Clarendon Press, 1993.

—— and Andrew Ashworth. *Proportionate Sentencing*. Oxford: Oxford University Press, 2005.

—— and Jareborg, Nils. 'Gauging Criminal Harm: A Living Standard Analysis.' *Oxford Journal of Legal Studies* 11 (1991): 1–38. Republished with minor changes as 'Appendix 3: Gauging Crime Seriousness: A 'Living Standard' Conception of Criminal Harm.' von Hirsch and Ashworth. *Proportionate Sentencing*. Oxford: Oxford University Press, 2005. 186–219.

Honderich, Ted. *Punishment: The Supposed Justifications Revisited*. London: Pluto, 2006.

Husak, Douglas. 'Why Punish the Deserving?.' *Nous* 26 (1992): 446–64.

——. 'Retribution in Criminal Theory.' *San Diego Law Review* 37 (2000): 959–86.

——. Why Criminal Law: A Question of Content?' *Criminal Law and Philosophy* 2 (2008a): 99–122.

——. *Overcriminalisation: The Limits of the Criminal Law*. Oxford: Oxford University Press, 2008b.

Kleinig, John. 'The Hardness of Hard Treatment.' *Fundamentals of Sentencing Theory: Essays in Honour of Andrew von Hirsch*. Ed. Andrew Ashworth, Martin Wasik. Oxford: Clarendon, 1998. 273–98.

Lacey, Nicola. *State Punishment: Political Principles and Community Values*. London: Routledge, 1988.

Lessnoff, Michael. 'Two Justifications of Punishment.' *Philosophical Quarterly* 21 (1971): 141–8.

Mackie, J. L. 'Morality and the Retributive Emotions.' *Criminal Justice Ethics* 1 (1982): 3–10.

Martinson, Robert. 'What Works? Questions and Answers about Prison Reform.' *Public Interest* 35 (1974): 22–56.

Mill, J. S. *On Liberty*. London: Dent, 1972 (originally published 1859).

Moore, Michael. *Placing Blame: A General Theory of Criminal Law*. Oxford: Clarendon, 1997.

Morris, Herbert. 'Persons and Punishment.' *Monist* 52 (1968): 475–501.

Nozick, Robert. *Anarchy, State and Utopia*. Oxford: Blackwell, 1974.

——. *Philosophical Explanations*. Oxford: Oxford University Press, 1981.

Quinton, Anthony M. 'On Punishment.' *Analysis* 14 (1954): 134–42.

Rawls, John. 'Two Concepts of Rules.' *Philosophical Review* 64 (1955): 3–32.

——. 'The Sense of Justice.' *Philosophical Review* 62 (1963): 281–305.

——. *A Theory of Justice*. Revised ed. Oxford: Oxford University Press, 1999.

Rex, Sue. 'A New Form of Rehabilitation.' *Principled Sentencing: Readings on Theory and Policy*. 2nd ed. Ed. Andrew von Hirsch, Andrew Ashworth. Oxford: Hart, 1998. 34–9.

Ross, W. D. *The Right and the Good*. London: Oxford University Press, 1930.

Scheid, Don E. 'Note on Defining 'Punishment.' *Canadian Journal of Philosophy* 10 (1980): 453–62.

Scholz, Sally J. *Political Solidarity*. University Park: Pennsylvania State University Press, 2008.

Signler, Mary. 'By the Light of Virtue: Prison Rape and the Corruption of Character.' *Iowa Law Review* 91 (2006): 561–607.

Tasioulas, John. 'Punishment and Repentance.' *Philosophy* 81 (2006): 280–322.

Ten, C. L. *Crime, Guilt and Punishment*. Oxford: Clarendon, 1987.

Wood, David. 'Retribution, Harm Reduction, and the Justification of Punishment.' *Oxford Journal of Legal Studies* 22 (2002): 301–21.

Wootton, Barbara. *Crime and the Criminal Law*. London: Stevens, 1963.

Zedner, Lucia. *Criminal Justice*. Oxford: Clarendon, 2004.