Retribution, Proportionality and Retributivism

Robert Edward Martin

ProQuest Number: U642079

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest U642079

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

I consider a number of characterizations of retribution and reject all on the grounds that none of them substantively meets the requirement of simple proportionality, viz. that the more blameworthy an agent or the more immoral his action the greater the distress inflicted on him should be. I then formulate a concept of retribution based on equivalence where this latter relation is understood to hold between certain properties of blameworthinesses or immoral actions on the one hand and properties of distressing actions (putative retributions) on the other. These properties are given rankings within special sets, maximal sets. A large part of the thesis is then given over to an analysis of this suggested equivalence and its deficiencies. I then introduce further criteria for an equivalence relation definitive of retribution, much of which turns on giving precise conditions for when relations involving terms like 'much greater than,' 'quite a lot greater than' and 'a little greater than' are to hold. These conditions make much use of relations between certain subsets elements of which are ordered in a particular way. These relations pose problems, notably two: cases where the truth-value of claims involving them cannot be ascertained, and cases where the size of maximal sets cannot be estimated. Finally, with our concept of retribution fully

explicated, I consider certain retributivist accounts of the justification of retributive punishment, where this latter is understood as punishment which meets certain criteria given for retribution, but all of these are found to be unsatisfactory. Thus, somewhat parallel to Cottingham's demonstration that many retributivisms don't offer justifications of punishment based on the etymological root of retribution qua repayment, I show that such retributivisms don't offer justifications of retributive punishment where this latter concept is explicated in terms of what underpins the idea of repayment: equivalence.

Part I. Retribution.

Chapter I. The Concept of Retribution.

 (i) Some Standard Definitions and Conceptions.

I want to consider first what an act of retributive (state) punishment might be. I shall understand this as one which is both an act of retribution and of state punishment. The Concise Oxford Dictionary offers the following three definitions of retribution:

1. a recompense for evil.

2. vengeance.

3. requital of evil.

I shall supplement these with the following, not all of which are strictly definitions:

4. action in accordance with the lex talionis.

5. we mete out retribution against someone for his action against another where we deprive the former of the right he violated in the latter or where we deprive him of an equivalent right. In this context one right is equivalent to another when an average preference scale registers indifference between either the loss of the one or the other. (Goldman)

6. one act is retribution for another harmful action where its infliction satisfies the grievance originating in the harmful action. (Honderich. Punishment: The Supposed Justifications. 1989; pp. 28-9, 42-4, 233-4.)

7. an act is retributive if it apportions punishment according to the criminal's act-related illicit pleasure or wickedness or moral desert. (Davis, 1983.)

8. an act is retributive if it reforms the agent of an offence or it destroys or annuls the evil in one who commits it.

9. retribution is equal in magnitude to (r.H)-c, where c is the compensation made to the victim by the offender, H the wrong, and r the degree of responsibility with which H is performed. (Nozick. Philosophical Explanations. 1981; pp. 363-5.)

10. proportionate punishment in accordance with Bentham's rules. (An Introduction to the Principles of Morals and Legislation. 1948 Ed. Chapter 14, sections 1-26.)

11. punishment proportional to the desert of the offender (following Bedau's critique of classification-based sentencing.)

12. punishment which maximizes utility.

13. punishment in which the culpability of the offender is equivalent to the distress caused him.

I suggest that we subject characterizations of retribution to a test composed of four criteria: (i) that it be logically possible to mete out retribution for all immoral actions: (ii) that it be logically impossible to mete out retribution for actions which

are not immoral; (iii) that it be a response to an immoral action rather than some feature of character; (iv) that retribution for an immoral action meet the criterion of what I shall call simple proportionality (SP.) What this requirement means is that the more immoral an action is the greater must be the unpleasantness of the retribution for it.

Let us start with 1. and interpret 'evil' as immoral action. Now it seems that an individual recompenses another for something only if the former has received something of net or overall value from the latter. In other words a captain might recompense a crew member for saving a man overboard from sharks even though the rescued man lost his leg but might not recompense one who saved the leg and not the man. But in that case one may only recompense one for an immoral action if one has received net value from it. Yet since it is logically possible that there be morally wrong actions without beneficiaries it is logically possible that there be immoralities for which no recompense is possible. But then 1 fails to meet criterion (i.) Moreover, given that it might be thought a necessary condition of recompensing someone that one recompensed by way of something of roughly equal value, our acceptance of such a definition would have the bizarre consequence that we would mete out retribution to someone only where we gave them something of net value.

In fact we certainly do not make this a condition of retribution even if it may be hoped by some that the reformation consequent upon retribution has an effect of such net value.

An alternative rendering of 1 would involve interpreting 'recompense' as requiring that we do back to the agent of the immoral action something of value equal to that which the victim received from his action. But then if the victim actually lost nothing of value by the immoral action either there would be no retribution for this action, in which case there would be an immoral action for which there was no retribution, or the retributive act would involve doing something to the wrongdoer of no value to him and by which he lost nothing of value. But clearly the latter way of viewing the retributive action would commit us to an account of retribution which did not meet criterion (iv.) Certainly simple proportionality will break down where, say, some attempted homicide receives a retribution with no adverse effects while less heinous acts receive responses involving considerable harm. So on this interpretation of 'recompense' 1. fails to meet either criterion (i) or (iv) or both. We shall turn to definition 2.

I suggest the following three conditions on an action being one of vengeance. 1. it must be for an action in which an individual has been harmed or pained or

wronged. 2. it must be performed either by the person harmed or pained or wronged or by one who identifies with the victim in the sense that they regard themselves as having been harmed or pained or wronged by the action against the victim and it must be inflicted on the person who harmed or pained or wronged the victim or against one who, in being the target of the action, causes the individual who did harm or pain or wrong the victim to be harmed or pained or wronged. 3. it is roughly equal in respect of harm or pain or wronging to the action of harming or paining or wronging which grounds it.

It is immediately clear that, as a definition for retribution, vengeance cannot meet criterion (ii) that it not be logically possible to mete out retribution for an action which is not immoral. Vengeance may be exacted for an action in which someone has been harmed where such an action is not immoral. Also it may not meet criterion (i,) viz. it may be that there are immoral actions in which no one has been harmed, pained or wronged. Consider two cases: 1. X intends to hurt Y by action A. He fails to hurt Y by A and instead by A hurts Z. A is an immoral action. 2. X behaves negligently in performing action A. A is immoral. Any of Y, Z and W might have been pained or harmed by A, though in fact none of them were. If we suppose that only Z has been wronged in 1 I suggest we are

subscribing to a concept of wronging a necessary condition of which is that an individual wronged be harmed or pained. If we do this then **Y** hasn't been wronged in 1 and **Y,Z** and **W** aren't wronged in 2. But then since not all immoral actions will involve anyone being harmed or pained there may be immoral actions involving no wronging and so retribution qua vengeance will not be possible for such actions. Alternatively we might decide that in 1 and 2 at the least **W, Y** and **Z** have been wronged. But if we accept this there may still be cases in which we are uncertain that an immoral action involves wronging. If someone digs up the dead body of the unknown and unloved Smith and mutilates it has anyone been wronged? Surely if anyone has it must be Smith. So definition 2 of retribution is unsatisfactory.

The third definition of retribution is as a requital of evil. Now Juliet requites Romeo's love for Juliet if she loves Romeo which suggests that I may requite another's evil or immoral action towards me by acting evilly or immorally towards them. But apart from the problem that this would make it a conceptual necessity of all retribution that it be wrong, which would render its justification impossible ab initio, if I am killed by the evil Smith his evil cannot be requited because I am not around to do the requiting. But then such retribution does not meet criterion (i) for retribution

is not possible for all immoral actions. Alternatively, 'requital' might be interpreted more loosely as 'doing the same back'. But if that characteristic of the action which we are to copy is not moral it is hard to see what features of it are to be duplicated by way of retribution. If Jones drowns Smith in a butt of malmsey we are required to drown Jones in a butt by way of retribution. Yet this response, taking no account of degree of immorality of action to be requited, will inevitably fail to deliver a conception of retribution which will satisfy criterion (iv;) insensitivity to features like lack of responsibility, accidentalness, provocation or duress will lead to a failure of SP between immorality and retribution for it. But with what are we to replace mere replication of the action in its crudest terms as the appropriate retributory response? Certainly we would not wish to replicate by performing an equally wrong action for the reason already given: such retribution would be unjustifiable. Moreover, making the retribution as wrong as the grounding immorality would not necessarily mean that the greater the immorality the more unpleasant the retribution. A ranking of actions in respect of how wrong they are may not be the same as one in respect of how unpleasant they are for their target. So this third definition may be rejected on the above interpretations at least.

Retribution understood as action in accordance with the lex talionis need not detain us long either. The lex is a principle about physical injury. But if we merely replicate injury as retribution we will mete out retribution for actions which may not be wrong, so flouting condition (ii;) and we may be unable to mete it out for actions which are wrong where no physical or even psychological injury is involved. Lastly it goes without saying that such an account will not deliver a scale of retributions and immoralities which satisfies SP. Moreover, even if we stipulate that we apply the lex only in response to immoral actions which involve injury then the problem remains that physical or even psychological injury caused is not a good indicator of how wrong an action is, and so its application would again lead to adoption of a scale of retributions and immoralities which did not satisfy SP.

Clearly if retribution (5) is understood in terms of legal rights the characterization will not meet our conditions: some immoral actions may involve violation of no legal right, while some violations of legal right arguably involve no immoral action. Alternatively, if we are to understand the rights as moral then the characterization of retribution fails SP. The fact that we are indifferent between the loss of two rights **MR1** and **MR2** may mean that the harm their loss causes us is roughly equal but this doesn't mean that a violation of

MR1 is equally wrong with that of MR2. Much of the harm
caused in the violation of MR2 could be accidental as
when I try to frighten someone into giving me money
with a gun and the gun goes off and kills my victim.
But in punishing me by deprivation of MR1, let us say,
I may be punished as much as one who killed someone in
much less extenuating circumstances, call such a
violation MR3, simply in virtue of the equivalence of
MR2 and MR3 in respect of harm. But then retributory
moral right deprivation will not accord with the
requirement of SP.

Retribution (6) also will not serve us. The first
problem concerns what is actually meant by harm. If by
this we understand physical and psychological injury
and pain then retribution will not be possible for
those immoral actions which involve no harm.
Alternatively it might be argued that any immoral
action is harmful. I am harmed perhaps if someone takes
a shot at me and misses, even if I never get to hear
about it. But we have a perfectly good word for this
anyway in the form of 'wrong'. If we assume what has
not been shown, viz. that all immoral actions involve
wronging someone, and we mete out retribution for
wrongful actions, this at least avoids the problem of
retribution for harmful actions which are not immoral,
but we still face the difficulty that, where a
grievance is understood as a desire for the distress of

a certain agent, this grievance may be satisfied if it is believed that a certain agent has been made to suffer. Retribution equivalent to a wrongful action in so far as it just satisfies the grievance of the victim, or loved ones, will be equivalent if the victim or loved ones have certain beliefs only. But if retribution only required some systematic deception then there would be no reason to suppose that retributions and immoralities would satisfy SP. Of course, we might revise the characterization so that the distress satisfying the grievance had to be inflicted, but where a grievance is a desire for distress there need be no rise in unpleasantness of retribution with increasing immorality. There is no telling how much distress a victim may desire. And if we stipulate that a grievance must be a desire which increases with the immorality of the action which gives rise to it then it seems that we are not really talking about a desire any more and our idea of equivalence is no more than this: that the pain of the retribution be greater the more immoral the action it is for. But in that case the characterization just is one of our conditions. Of course it may be that no other account will do any better but for the moment we will continue our search.

Davis (1983) suggests a different way of understanding the impulse to apportion retribution to moral desert or

wickedness. Although he foresees certain difficulties in establishing a scale of crimes and penalties he suggests that the kind of pairing of crimes and penalties in ascending order of immorality or unpleasantness which would meet SP, at least for legal offences, is generated by a certain kind of equivalence. Roughly, the unfair advantage secured by commission of an offence may be determined by the price a licence to commit that crime would fetch on the open market and the equivalent penalty for the crime is one which is in some way analogous to the price for the licence to commit it.

Consider the Davis account in respect of SP alone. We may (falsely) assume for present purposes that crimes constitute all and only immoral actions. Now the greater the advantage a criminal takes by a crime the greater the penalty he deserves. So the penalty for a crime will be greater if the licence to commit that crime would fetch a higher price on the open market. Thus, if SP is to hold, the more immoral a crime is the greater must be the price of a licence to commit it at auction.

Certainly the fact that in Davis's auction scenario the number of licences available for the more serious crimes is heavily restricted would not, of itself, push the price up. Scarcity only raises a price given that demand exceeds supply. However, as will be noted later

in connexion with Morris's account, it could well fail
to be the case that, say, many would commit murder in
the absence of legal constraint not to do so. Perhaps
there would be more licences for murder than people who
wanted them. Rather Davis relies more on the idea that
because a licence to kill, say, is disruptive of social
order, people would be willing to pay someone not to
use it, in which case the licence would be worth at
least what people would be willing to pay for its non-
use. But if we are to use this for our present purposes
this amounts to the claim that the more disruptive of
social order a crime is, or perhaps the more we fear
it, the more expensive the licence to commit it will
be, the greater the advantage gained by its exercize
and, finally, the greater the deserved penalty for its
commission. Yet this looks curiously unlike a
retributivist claim since scaling offences by social
disruptiveness or how much they are feared begins to
look like something which usually functions in a
utilitarian calculus, the kind of calculus which
notoriously does not deliver offence and penalty scales
which meet SP. Moreover, if we scale offences by how
much they are feared frequency of crime or the
likelihood of its commission will figure in the
assessment. Burglary, for instance, may be much more
feared than kidnapping simply because relative social
order pushes the latter to the back of one's mind. It

might be objected here that what determines the price is the fear that would be aroused by a crime, not the likelihood of its happening. But that surely can't be right. In a situation in which resources are limited what a licence is worth will be determined by how much people will want to pay for its non-use. But if such licences are few enough then it may be rational to use purchasing power to diminish the likelihood of being the victim of the exercize of a licence to burgle. On the other hand, even if we could measure social disruptiveness, which in the final analysis would turn on whether agreement concerning relative factors could be reached, it is at least plausible that murder would figure less highly than, say, a systematic defrauding of a government body.

Secondly, it is unclear why the unfair advantage over law abiders of one who acts criminally is to be determined by the price of licences. Why should the fact that one crime is more socially disruptive, or more feared, than another show that the commission of the former gives me more unfair advantage than the latter. Clearly if I get more unfair advantage I get more advantage of some kind. But, to anticipate more detailed later discussion of the unfair advantage type of account, if we construe the greater advantage vis-a-vis other crimes of committing murder as greater freedom from the constraint imposed by legal rules it

remains to be seen why the constraint on our liberty is more restrictive or onerous where we may not murder than where, say, we may not steal something we are obsessed to possess and which will always be beyond our purchasing power.

So this account of the measurement of moral desert does not satisfy our criterion (iv.) However, it might be objected that a more sympathetic interpretation of criterion (iv) is available, one in which the degree of immorality of an action is determined by the Davis-style unfair advantage taken by it. But of course to see whether the auction model would suggest a ranking of crimes and penalties in accordance with simple proportionality on its new interpretation we would have to have some idea of the correct ranking of immoralities derived from something other than the auction model. Perhaps this is available since Davis (1986) is at pains to show that his procedure suggests a ranking such that attempted crimes are less serious than completed ones. Thus we could invalidate Davis's account from the viewpoint of revised SP (where the more immoral an offence in respect of the unfair advantage derived from it the more unpleasant the penalty for it must be) if it fails to suggest such a ranking of attempts vis-a-vis completed crimes.

One problem faced by the Davis account is that certain protective associations (roughly, formed from

individuals who share our fears) might bid up the price of licences to attempt certain crimes till they reached that of licences to commit completed crimes since the former licences, so it could be argued, would encourage completed crimes just as the latter would and so such associations would have good reason to keep such licences out of the hands of would-be criminals. But then if the prices for attempts and complete crimes were the same the punishment ought in fairness to be the same. Davis blocks this result by sacrificing the claim that the protective associations would in general determine the price of licences. He grounds this by supposing that in order to raise revenue the licence issuing body might issue more licences than the associations could buy up, in which case money could be made from would-be criminals as well. This would mean that criminals would determine the price of licences and so licences to attempt would be worth less than licences to commit complete crimes, for a criminal would prefer a licence to actually commit the burglary, say, to one just to attempt it. But then of course if the price of a licence is to be determined by the criminals we go back to the problem that moral gravity (though not of the Davis kind) would not necessarily determine licence price since the criminal may well prefer a licence to commit a crime which we would consider less grave than another. But in that case

although attempts would be of less value than completed crimes Davis has to embrace the possibility that he is committed to a scale of crimes and penalties significantly different from the one usually regarded as intuitively acceptable. This of course doesn't mean that the revised interpretation of SP in which immorality is measured by unfair advantage gained would not be met by a scaling of crimes and punishments congruent with the auction model but it does mean that a presumption of the Davis account can no longer bolster it. This presumption is that the auction model would deliver a scaling of crimes and penalties which we intuitively accept (i.e. one which meets SP on its original interpretation.) So this sympathetic rendering of SP would mean that the auction model would only deliver a scaling which met its requirements at the cost of that scaling being considerably different from one which met SP on its former interpretation. This cost is surely unacceptable to Davis.

Secondly, Davis mistakes the scope of his auction model. Clearly he must rebut the claim that no one would wish to buy a licence to attempt since this would suggest that all attempted crimes would deserve the same penalty. He suggests that there is reason to buy one because if one tries but fails in one's attempt to commit the crime one will be protected from punishment if one has a licence to attempt. Moreover it is better

to have one of these licences to attempt than no
licence at all because commission of a crime with no
licence would constitute the crime of poaching for
which some price at a meta-auction would be set (thus
determining a penalty.) But this is clearly a dangerous
move. One who commits a crime does not commit two
crimes, one the crime itself and the other the failure
to buy a licence to commit it. If I commit the crime of
(real) poaching I haven't committed the crime of, say,
hunting deer and that of not having a licence to do so.
Of course it is open to Davis to object that only one
crime and one meta-crime have been committed, not two
crimes. But then it is unclear how we are to punish one
who has committed a crime. Is it to be determined by
the price of a licence to commit it, or by the price of
a (paradoxical) licence to poach, that is, to commit it
without a licence? Davis claims that poaching will be
severely punished. But surely not all poaching would
be; if it were crimes would no longer be punished in
proportion to the price of the licences to commit them.
Rather it is inductively reasonable to suppose that the
meta-licence to commit the meta-crime of committing
some crime without a licence will fetch the same price
at least proportionally as the licence to commit the
crime simpliciter. For if one wishes to prevent/perform
crime c and bids x for the licence it might be thought
that one would have the same motivation to

prevent/perform the crime c', that of performance of c without a licence. But even if the ripples of an infinite regress might be discounted the introduction of poaching is careless because unnecessary. One needn't commit a meta-crime of poaching in committing c without a licence but instead just commit the crime c, and this without a licence. This is all that is required to make it rational under certain circumstances to wish to purchase a licence to attempt. This criticism may be just a quibble but, in view of the earlier criticism, the ranking produced by the auction model revised to account for differential punishment of attempts over completed crimes is unattractive.

So the Davis account of a principle for determining retribution doesn't meet the requirement of simple proportionality, that is, criterion (iv,) and with the suggested modification of (iv) the suggested principle doesn't seem worth defending. However, in the context of the discussion of attempted and successful crimes we may extrapolate a different principle for determining what retribution requires. This comes from Duff.

Duff proposes that the punishment deserved for an attempted as opposed to a successful crime is less because one of the essential purposes of punishment is to communicate certain values to the offender. Now if, so the argument might run, we were to punish successful

and attempted crimes equally we would communicate the
message that, as far as the law or the community was
concerned, it was a matter of indifference whether,
say, harm was caused or merely attempted. But we do not
wish to communicate this message and so we punish
differently.

One difficulty concerns the moral import of this; it
may be true that effective communication of our valuing
the well-being of members of the community requires
differential punishment. But it seems some way from
this to the conclusion that differential punishment is
deserved as a matter of justice.

Secondly, if the purpose of punishment is something
like the emphatic denunciation by the community of a
crime, different conventions would suggest different
methods of achieving grades of denunciation. Now it
could be argued that greater penalties might signal
greater denunciation. (After all, it is one
comprehensible convention.) But the denunciation Duff
appears to have in mind is not one which should rise or
fall with culpability, nor, for that matter, with the
immorality of an offence (at least according to one
usual way of estimating it.) If **A** and **B** are equally
culpable, commit crimes which are equally immoral, but
A is guilty of attempted **X** and **B** of successful **X**, then
if **B**'s action results in greater harm as a matter of
chance then **B** ought to be more severely denounced and

so more severely punished. So on this view of how denunciation, and with it retribution, is to be apportioned SP or (iv) will not be met.

Alternatively, we might suppose that matters of luck concerning injury caused by an action should figure in the assessment of how morally wrong that action is. But if so this will not provide us with a helpful thumbnail sketch of how to determine what is appropriate qua retribution since if we merely denounce the more morally wrong offence more our criterion for estimating what retribution requires is simply that SP be met. Although denunciation may figure as an important part of the justification of punishment to say that retribution must involve a more unpleasant penalty the more immoral the offence for which it is meted out so that effective denunciation be achieved is to say no more than that retribution must meet SP. Of course a concept of retribution part of which is that it must meet SP will meet SP but this has hardly advanced us.

However other remarks which provide substance are available from Duff. One of the important features of differential denunciation is that it communicates values, which includes of course not just what we value but how much we value it. One point of differential denunciation of attempts and completed crimes is that it encourages the criminal to repent and be relieved that he has not caused the harm he intended (and not

just for himself but for the victim.) But clearly a grading of retributive punishments so that each punishment effectively reforms or best reforms the agent of the offence for which it is administered would fail to meet SP. To begin with moral reformation concerns character, while retribution is for an action. To get an evil genius to repent of a minor misdeed may take all of man's and heaven's wrath, but clearly this would not be consistent with an apportionment of punishment in line with the immorality of actions.

Secondly, even if one were to suppose a congruence between acts and character such that the more immoral the offence the more evil the one committing it it still wouldn't follow that the punishment which effected repentance would have to be greater the greater the immorality of the offence for which it was inflicted. The idea that a penalty can somehow destroy evil, and a greater penalty will destroy a greater evil seems to draw its inspiration from the same source as this. Apparently, a greater penalty will be necessary to effect repentance for what is a greater evil. That such a psychological generalization is true is a precarious presupposition and prima facie is not a good ground for supposing that the account will meet SP.

Finally, and more generally, if retribution is simply that which appropriately denounces a crime, then the conventional punishment responses needn't meet SP for

obvious reasons. Infliction of pain is surely not part of the concept of denunciation.

Turning now to Nozick we are to assess a retributive matching penalty by reference to the product of r and H where r is the degree of responsibility with which H, the moral wrong, is performed. Now one might adopt different conventions for determining what is relevant to the assessment of how immoral an action is. For example, compare two situations: 1. the half-witted Smith kills Jones, mistakenly believing that he is being taunted by him; 2. the intellectually normal Brown kills Jones, mistakenly believing that he is being taunted. It seems that we may either say that the actions in 1 and 2 are equally morally wrong, but Brown is more blameworthy than Smith, or that Smith's action is less morally wrong than Brown's simpliciter. Now if SP is understood in accordance with the second convention, that is, assessment of immorality of action takes into account degree of responsibility with which the action is performed, then Nozickean retribution will meet it, for the product of r and H will ensure that Brown is punished more than Smith. On the other hand if we understand the immorality of an action as Nozick does, that is, as taking no account of degree of responsibility, then Nozickean retribution will not meet SP on this interpretation of wrongdoing, since on

such a reading of SP if Brown is punished more than Smith he is punished more for the same wrong.

And what are we to make of the idea that retributive punishment is to be equal in magnitude to the product of r and H? If magnitude is something for which there will be units then clearly wrongs and punishments do not possess magnitude. For one thing where units are available we can talk of one thing being half or twice another. Yet even if a ranking of wrongs were available the wrong ranked at position 14 wouldn't be twice that ranked at position 7. For instance how would we know that the degree of difference between the wrongs at 8 and 7 was the same (or more or less than) those at positions 11 and 12? Or how could two wrongs at 7 be equal in wrongness to one at 14?

On the practical level it is little use to us that Nozick tells us that where r=1 we perform where feasible the same action against the offender which he performed against the victim. By the same action he can't mean a wrong action, but a non-moral description of the action which will allow us to correctly characterize it morally would need to be so specific as to make duplication practically impossible. Also if we were to adopt this policy SP would not be met. This is because the H value doesn't stand for a harm but a wrong. Hence H may be very wrong but not particularly harmful, or very harmful but not particularly wrong.

But then not immorality but harm done would determine what we did to the offender. Nor does this policy of retributive punishment square with his own suggestion about how much we are to punish. If we are to punish equal to the maximum of the disutility the victim or the offender could be expected to incur there is something rather arbitrary in this. If Smith is terribly sensitive but expects the thick-skinned Jones not to be caused much disutility by his action, and Smith's expectations are in no way unrealistic or biassed, why should we penalize Smith above and beyond what he expected by making him undergo what he would have undergone if the victim of his own action, especially if Jones does indeed incur no more disutility than one might expect? So Nozickean retribution is not satisfactory for our purposes.

Bentham offers (roughly) the following guidelines for the proportioning of a punishment to an offence: punishment shall 1. outweigh the profit of the offence; 2. venture more against a great offence than a small one; 3. cause the least of two offences to be preferred; 4. be for each particle of the offence; 5. be in no degree without special reason; 6. attend to circumstances influencing sensibility; 7. be such that want of certainty of punishment shall be made up for; 8. be for the habit where an act is indicative of a habit; 9. be increased in quantity for the sake of

quality; 10. deviate from such rules for a moral
lesson; 11. sometimes be diminished because
unprofitable; 12. in some cases be deviant from
proportionality on a minor scale for simplicity's sake.

I want to consider three questions a utilitarian
might need to answer. (i) If the aim of punishment is
to maximize utility does it make any difference whether
we always punish so that each punitive act taken
individually maximizes utility or instead adopt a
system of punishment which maximizes utility? (ii) If
we punish in accordance with the rules 1-12 above do we
in fact maximize utility? (iii) If we punish in
accordance with 1-12 do we meet the requirement of
simple proportionality?

(i) I shall try to show that if a system or
institution of punishment maximizes utility then
individual acts of punishment within that system must
be such that taken individually they too maximize
utility. If this proof is successful it should be
possible to see how one could establish the converse
result, viz. that if all the punitive acts of an
institution taken individually maximize utility then
the institution itself maximizes utility.

Suppose we have a system of punishment S made up of
the individual acts of punishment **p1,p2...pn**. Suppose
also that S maximizes utility in the restricted sense
that no other system creates as much utility as S (i.e.

there may be something which is not a system which has
greater utility.) We might try to show that each of
p1...pn taken individually must be maximizing of
utility in the following way: suppose that p3 is not
utility maximizing; then some other act pu is; but then
p1..pu...pn, where pu replaces p3, it might be said,
would have greater utility overall than p1...pn. One
might try to rebut this by saying that p1..pu..pn
wouldn't necessarily have greater utility because the
substitution of pu for p3 might change the utility of
individual p values p1...pn from which p3 is absent.
P1...pn has a utility as a whole different from the sum
of the utilities of its acts. This utility overall
could well change as a result of the suggested
replacement. But this can't be right. Pu will be
performed before, after or contemporaneously with other
p values. But then at some point we can either perform
pu or p3. But if pu will have the result that or
influence things in such a way that another set of acts
is performed which, taken overall, will mean that
utility is not maximized, then pu can't be utility
maximizing taken individually. Nor can we avoid this
conclusion by talk of just the consequences of pu or
p3. If I leave open the door so that ten hours later
the intruder can walk in and murder me the utility of
my act will be determined by how it fits in with other
acts which it facilitates or fits in with rather than

simply by those acts which it causes. So the utility of
pu will be determined by how it fits in with other **p**
actions. So if **pu** is maximizing of utility this means
that **p1,..pu ...pn** is utility maximizing. But then S
doesn't maximize utility which is contrary to
supposition, so by reductio individual **p** acts in S must
be utility maximizing.

So if S maximizes utility so do **p1...pn** taken
individually. A similar proof would establish the
converse. In view of this I shall think of a system of
punishments which maximizes utility as a group of
individual punishments each of which maximizes utility.
This means that if one punitive act could be found
which did not maximize utility then the system from
which it was drawn could not be utility maximizing.
However it also means that the test of whether a
particular punitive act is utility-maximizing would
amount to a test of whether it occupied a place in a
system of punishment which itself maximized utility.
But then this is surely to rule out the practicality of
a test procedure to determine whether a particular
punitive act maximizes utility. Suppose the act to be
tested is **pt**. Then **pt** must take its place in some
system **p1..pt..pn** which, incidentally, would almost
certainly include punishments imposed before **pt**. But
suppose that the system S itself has a certain
flexibility, in which case if we imagine **p1..p6** as the

acts performed before **pt**, different series of acts
simultaneous with or successive to **pt** will be possible
as part of **S**, we will then have to consider not only whether
pt + p8...pn is utility maximizing given that **p1...p6**
has been performed but whether **pt +** some other series
of actions given performance of **p1..p6** maximizes
utility. Such a question will be practically impossible
to determine in which case question ii is not one which
I shall try to tackle.

Let us turn to the third question I posed, namely,
whether punishment in accordance with rules 1-12 will
meet SP. I think it is immediately clear that such
punishment will not. Bentham tries to make commission
of a crime unattractive to a rational agent by making
the cost of punishment for it greater than the gain of
committing it (without detection.) But one consequence
of this is that we punish more for a crime which is
likely to go undetected than for one which is not
simply because the rational agent facing the
possibility of committing a crime will estimate the
cost of committing it as a function not only of the
magnitude of the punishment but also of the probability
of receiving it. This appears in rule number 7. But
clearly greater punishment for that which it is more
difficult to get convictions for is not necessarily
going to be congruent with greater punishment for more
immoral offences than for less. Secondly, it is to be

questioned what is meant by a great offence here. If the general principle at work in the rules is to make the crime unattractive to a rational agent then punishment must be greater the greater the gain from the crime, but it needn't be that the greater the gain from the offence the greater the immorality of it. The crimes owed the worst punishments on this account will be those from which personal utility for the criminal is highest, not those most grave or heinous. And for this reason if magnitude of offence in rules 2 and 3 is determined by degree of immorality then these rules are in conflict with rule 1. In the absence of a decision procedure for resolving such conflicts it is unclear whether SP is met. Again rule 10 may come into conflict with 3 if 3 requires that we attach the lesser penalty to the less wrong offence, since doing this may be waived pour encourager les autres. And does rule 11 require that Kant's convicted murderer from the island community about to disband should be set free? In general, the more the system of punishment approaches utility maximization, and individual punitive actions maximize utility, the more doubt there is that SP or some such retributive principle will be met.

Turning now to definition 11, in his article on classification-based sentencing Bedau considers a retributive punishment proportional to the desert of the offender where this desert is determined by the

harmfulness of the offence and the fault in the offender. I shall consider only two ways in which we might understand the fault of the offender, firstly, as culpability and secondly, as given by criminal record. In turn we might understand culpability either as (i) a function of how wrong an action is and the moral responsibility of its agent or (ii) a function of how harmful an offence is and how morally responsible the agent is. Lastly, it will be remembered that SP may be understood in a narrow or an extended sense. In the narrow sense an action is assessed morally without taking account of the agent's moral responsibility for it. Thus, for example, things like the intention with which the action was performed, e.g. Jones's desire to kill Smith because he was provoking him, would be included in the assessment of the immorality of the action but the fact that Jones was a simpleton would not. In the extended sense it would be part of our assessment of how morally wrong the action was that Jones didn't really have a complete understanding of what was going on in a way relevant to moral responsibility as we see it.

Clearly if we adopt sense (ii) of culpability offender desert will simply be determined by fault in the offender (and not fault + harm,) since this already takes into account the harmfulness of the offence. But if desert is simply a matter of

culpability in sense (ii) it could be argued that on the extended interpretation of SP this account of retribution will not meet it (i.e. SP) since the definition uses too thin a set of criteria for determining culpability. An example may help. It is usually accepted that factors such as provocation and temptation reduce culpability. But if someone acts under great temptation this may not make them any less responsible for their action. But obviously temptation cannot figure in the assessment of culpability under the title of harm of offence so it is dubious that it can actually be included in an estimation of culpability on (ii) at all. However if we determine how wrong an offence is by reference to factors such as blameworthiness then where all factors are equal except degree of temptation present though the two acts will be morally wrong to different degrees (as far as extended SP is concerned) on the other hand as far as desert qua culpability in sense (ii) is concerned they may yield similar readings (for culpability) and so determine the same retributive punishment. Hence on sense (ii) of culpability definition 11 doesn't meet SP in its extended sense. Likewise on the narrower interpretation of SP (where wrongness of action is not determined by reference to moral responsibility) in the case above of two actions with different moral evaluations in virtue of their difference with regard

to temptation, according to narrower SP there should be a difference in retributive punishment for the two actions, but definition 11 will not require a differential response because ⁻ as far as culpability is concerned in sense (ii) the agents are the same.

Alternatively, if we understand culpability in sense (i) it would appear that harm done by an offender would be counted twice in assessment of desert, once in how wrong the offence was to determine culpability, and once more in the overall calculation of desert in accordance with the definition (i.e. where desert is a function of fault and harm.) Clearly this would not do as it would introduce some consideration when, ex hypothesi, its moral relevance had already been drained. However if one were to ignore the harm condition in the general characterization of desert then desert would become simply a matter of culpability. But then, of course, SP in its extended sense would be met by definition 11 since the definition would just require that the more culpable an offender the more punishment he should get, and because of the way this culpability is assessed it would follow that the more morally wrong an action the more punishment to be meted out, so satisfying extended SP. SP in its narrower sense would not, of course, be satisfied by retributive punishment since in this sense

of SP moral responsibility is not relevant. We may
conclude then that even where the definition meets SP
it offers nothing new.

Lastly, following the Pennsylvania penal code, we
might interpret fault as something to be measured by
reference to past offences (for which the individual
was convicted.) This means that punishment is to be
determined barely by the harm caused by the offence and
the criminal record of the offender. But this is
clearly inadequate to meet SP since the gravity of an
offence will not be determined by reference to features
relevant to SP on either of its interpretations:
intention, recklessness, provocation, temptation,
maliciousness, etc.. Hence it will not do.

Definition 12 is often criticized on the grounds that
a penalty which maximizes utility may not be consistent
with SP. Now Benn proposes that the utilitarian can
consistently adopt a penalty schedule which is
proportional in the sense that minor offences will be
accorded lenient punishments and major offences harsh
ones. His idea is that we select a penalty at which
'the aggregate of suffering caused by crimes actually
committed and punishments actually inflicted would be
the smallest possible.'

Benn may mean either that (i) where p_1 represents a
type of penalty and c_1 a type of crime we must choose a
p_1 such that $p_1 + c_1$ in all its instances amounts to

the least suffering or (ii) that over the range of penalty and crime types we must choose **p1** so that **(p1 + c1) + ...(pn + cn)** amounts to the least overall suffering.

The first point to note about either option is that no account is taken of the amount of happiness produced, in which case it might be thought that we have not taken into account everything relevant to a utilitarian calculation in the first place: a little bit more suffering may be worth a lot more happiness. More importantly, choosing **p1** in accordance with (i) seems to leave out of the account its effect on other crimes. It might turn out that **p1**, though extremely effective as a deterrent of **c1** by its very mode of operation caused racial tensions leading to a wave of crime of type **c2**, which in turn could only be quelled by a very harsh and hence utility-reducing punishment **p2**.

But more interesting than this is Benn's idea that where a crime is minor it will not be cost-effective to introduce a harsh penalty whereas if it is major a harsh penalty will be; the harshness of the penalty will be counterbalanced by the reduction in crimes which have individually very great disutility. Yet if we leave to one side the qualm that this does not seem to dispose of the possibility that it would be a good idea to introduce flogging for parking offences, since

with such a penalty there might be almost no offences
of this kind, if we explore the consequences of the
Benn idea in relation to (i) it comes to look far from
obvious that it would be cost-effective to introduce
harsh penalties only for major offences. For suppose we
have a minor offence type c_1 and two possible penalty
types, p_1, which is lenient, and p_2 which is harsh and
represents p_1 with the addition of harm h. If Benn is
right and it would not be worthwhile to use p_2 it must
be that the reduction in crimes of type c_1 produced by
p_2 but not by p_1 would amount to a saving of less
suffering than was introduced by the penalty schedule
itself. But if we simplify and suppose that no cases of
c_1 go unpunished, then on the assumption that there are
n occurrences of c_1 when we use p_1 the aggregate of
suffering if we use p_1 will be given by 1. $n(p_1 + c_1)$
while the aggregate of suffering if we use p_2,
supposing there will be y fewer occurrences of c_1, will
be given by 2. $(n-y)(p_2 + c_1.)$ This latter however is
equal to 3. $(n-y)((p_1+h) +c_1)$ which, expanded, gives
4. $np_1+nh-yp_1-yh+nc_1-yc_1$, which is equal to 5.
$np_1+nc_1+nh-y(p_1+h+c_1)$. Appropriate bracketing then
yields 6. $n(p_1+h+c_1)-y(p_1+h+c_1)$, which is the product
of the sum of suffering involved in the harsh penalty
for the crime and that crime itself and the number of
times the offence would have been committed under the
old penalty minus the product of this sum and the

number of times less the offence would be committed under the harsh penalty. But since $n(p_1+c_1)$ expanded becomes np_1+nc_1, and this sum is at the beginning of 5. we can see that part of Benn's supposition about suffering (i.e. that 2 involves more suffering than 1,) will only be correct if $nh-y(p_1+h+c_1)$ yields a positive amount of suffering. This will only be so where $y(p_1+h+c_1) < nh$. But then if we suppose that p_1 and c_1 are minor, that is that their sum is small, coupled with the assumption that $y < n$, it may be reasonable to suppose that the inequality holds, since the inequality will approximate to $yh < nh$, and certainly yh will be smaller than nh.

Conversely it must be shown that it is plausible to suppose that it creates more suffering to have minor penalties for grave offences than to have harsh penalties for such offences. Supposing this time that p_1 is a harsh penalty, h is the reduction in that penalty to produce a lenient one, and c_1 is a grave offence, it must be the case that if c_1 is committed n times with a harsh penalty in position and $n+y$ times with a lenient one, that the aggregate of suffering given by 1. $n(p_1+c_1)$ is less than that given by 2. $(n+y)((p_1-h)+c_1)$. From 2 we have 3. $np_1+nc_1-nh+y(p_1-h+c_1)$. Hence 1 will amount to less suffering than 3 only if $y(p_1-h+c_1)-nh$ yields a positive amount of suffering. But this will be so only if $y(p_1+c_1)-h(n+y)$

yields a positive amount of suffering. This will be the case where $y(p_1+c_1) > h(n+y)$. But one is surely at a loss to see why this last inequality is compelling. Certainly there seems to be nothing about the structure and the assumptions we have used which makes it so. But then the plausibility of Benn's claim is surely not established.

Moreover if we adopt reading (ii) of the Benn claim then our calculations will have to be more complicated, since we are not to compare the sum of two kinds of punishments and the crimes for which they are imposed but rather such different sums when calculated along with other sums of (different) punishments and the (different) crimes for which they are imposed. The complexity of this may be glimpsed as follows: if we have an existing aggregate of suffering produced by a punishment system we will need a calculation not simply of the form $(p_1+c_1)+...+(p_n+c_n)$ where there are supposedly n instances of each $p+c$ summation and p and c values with different suffices designate different punishment and crime types. Rather we will need to know the number of instances for each $p+c$ summation because the introduction into our comparison of a heavier or more lenient punishment may well affect the number of occurrences of other p and c values. Thus the aggregates to be compared, where n and m values designate numbers of crimes (and punishments,) would

take the form of, say, $n1(p1+c1)+...+nn(pn+cn)$ and, with a harsher penalty in place, $m1(p1+h+c1)+...mm(pn+cn)$. We do not know that, say, $m2$ will be the same as $n2$. The more complicated this calculation the less obvious it will seem that the Benn contention necessarily holds. In this context we will have to compare aggregates of suffering given by $n1p1+n1c1+n2p2+n2c2+.. ..+nnpn+nncn$ and $m1p1+m1h+m1c1+ . ..+mmpn+mmcn$ etc.. For this reason I think we can conclude that much more needs to be said on the subject of utilitarian proportionality to convince us that the utilitarian penalty schedule will meet SP.

Definition 13 requires that the distress the offender is caused be equivalent to his culpability. Honderich claims that there can be no equivalence between distress and culpability as the two terms of the relation are not commensurable, that is, there are no common units available for them. (Honderich, op. cit., pp. 27-8, 210-11.)   It might be objected to this however that things may be of equal value even though there are no common units of measurement, simply because there are no units of measurement. Against this it might be said that even where two things are of equal value there are common units of measurement. Perhaps x and y are of equal value to me because I would pay as much for one as for the other. But this will not scupper the objection for two reasons. Firstly although in the case of two things being of equal value

the two terms of the equivalence relation have common units it is not in virtue of the thing which these units are of that the equivalence holds. The two bricks may be of equal weight because they both weigh 2 kilos, that is, because the common units of weight are available. But pounds sterling are units of price, of a certain kind of value even, but clearly not necessarily of the value at issue, that is, of the value in terms of which the equivalence holds. The point about common units being a necessary condition of equivalence, if it is to have plausibility, must be one about common units of whatever it is in virtue of which the equivalence is to hold. So the objection that there can be equivalence without common units remains valid.

Secondly, Honderich cannot be maintaining that cash value would serve as a ground of equivalence since the example of things he gives which couldn't stand in the equivalence relation is one of tunes and buns. But surely tunes and buns can be related in respect of cash value. What in fact we can't say is that there are degrees of tuneness and bunness in virtue of which an equivalence relation may hold (or fail to hold.) But then if this is the point about tunes and buns it will not establish that there cannot be equivalence between the items at issue, for clearly there can be degrees of distress and culpability.

(ii) Retribution based on equivalence.

Perhaps we can formulate some concept of retribution
on the basis of one of the central features of some of
the foregoing characterizations, viz. some kind of
equivalence or sameness: equivalence in respect of
value, harm, pain or wrong, or sameness of action (in
the case of requital for example.) Let us say that r is
retribution for i only if i is an immoral action and r
is in some respect equivalent to i. Now one possibility
is that we explicate this equivalence as a requirement
that the disvalue of r to its recipient must equal the
moral disvalue of i. However, if we interpret the moral
disvalue of an immoral action as simply how wrong it
is, and we measure the disvalue of r to its recipient,
the agent of i, in respect of how it reduces his level
of well-being, then we have an equivalence between how
wrong i is and how much well being the agent of i loses
by r. One may note here that the ranking of different
r's to some one individual in respect of their effect
on his well-being may change for that individual
depending on his age and experiences. For example, a
young person might prefer a relatively short
incarceration with considerable physical hardship to a
longer but more physically comfortable one. He might

wish to get the imprisonment over with as quickly as possible in order to get out into the world to make his/her mark. However, in later life such an individual might develop a taste for scholarship and shy away from the shorter imprisonment in favour of a longer one which permitted study. In other words in early life the longer incarceration would probably cause greater loss of well being than the shorter; in later life vice versa. A fortiori, given that differences of appetite and constitution may affect the ranking of the disvalues of different retributions for a single individual, one would have reason to expect the rankings to be different for different people.

A further problem suggests itself in view of the fact that it is in respect of effect on well-being that we measure disvalue of a retribution. It seems that we would do well to measure disvalue caused by a retribution in respect of the distress it causes rather than the effect it has on the well-being of the target of that retribution. The reason for this is that we can imagine a punitive act effecting moral regeneration, or some other state which had the consequence that, on balance, the punitive act enhanced the well-being of its target. Indeed one argument for retribution might well be that it effects such regeneration. But one consequence of such a way of measuring disvalue might be that if someone did undergo moral regeneration as a

result of what was thought to be retribution, it might
turn out that, because of that regeneration what was
thought to be retribution was not retribution at all;
indeed in order to mete out retribution it might be
necessary to bring back the morally regenerated
individual and punish him some more. If we were to
measure disvalue in respect of distress caused by the
putative retributive act, on the other hand, we would
not face this problem since effects on well-being,
especially those to society's advantage, would not
figure in the disvalue relevant to retribution. But
even this states the matter improperly for the disvalue
in respect of distress is apt to conjure an idea of an
assessment of distress in respect of the things of
value lost and gained by it. This sort of calculation
would obviously open the door again to questions of
well-being acquired as a result of moral regeneration,
etc.. Nor do we mean by disvalue in respect of distress
just a calculation of the things of value lost as a
result of distress experienced. If distress caused us
to lose something of value which didn't actually figure
in distress and this figured in our calculation of
disvalue then clearly things other than distress are
entering into the assessment of the disvalue of the
retribution. This is not to say that such a way of
measuring the disvalue of retribution is not
acceptable, rather it means that there will not

necessarily be a straightforward correspondence between disvalue and distress of a retribution if we choose such a method of determining the disvalue of a retributive act. We could have two retributive actions r1, r2 such that although r1 caused its target I more distress than r2, r1 was nonetheless of   less disvalue in respect of distress than r2. In view of this I suggest that we stick to a (slightly) more straightforward means of measuring a retribution. Thus the equivalence we are to consider is that between the distress a retribution r causes its target and how wrong the immoral action i is to which that retribution is a response.

Our next question then is whether there can be some kind of equivalence between how wrong an action is and the distress caused by it. Prima facie, there cannot be an equivalence between two actions, i and r, in respect of the qualities mentioned. There are at least two distinct equivalence relations: numerical and qualitative. Certainly numerical equivalence is a non-starter, since r and i would only be numerically equivalent if they were one and the same action. Now it is true that we can imagine circumstances in which an immoral action backfired on us and constituted its own retribution. Suppose that in aiming to kill Smith the gun backfired and shot me in the leg leaving me permanently crippled. But certainly we would not wish

this to be a general condition which retribution must satisfy. For one thing it would rule out the possibility of anything like a state punishment system with a retributive basis.

So we are left with qualitative equivalence. But this can't be the relation which holds between r and i either since qualitative equivalence holds between two objects in respect of a single quality, but what we have is some supposed equivalence between how much distress is caused on the one hand and how wrong an action is on the other, in other words between two qualities, distressingness and wrongness. Of course, we might drop the idea of an equivalence based on moral wrongness and distress and opt for a characterization of retributive equivalence in respect of one or other or both these aforementioned qualities. But then although we would have an equivalence of which we can make sense we would also have a relation such that SP need not hold between pairings of r and i values. To see this suppose that the retributive act $r1$ and the immoral act $i1$ are equivalent in respect of the distress they cause to some individual I. Suppose also that $r2$ is in the same respect equivalent to $i2$. Now if SP obtains it must be that the greater the immorality of an action the more unpleasant, or as we shall now put it the more distress must be caused by the retribution for it. Yet it is possible for us to

conceive of circumstances in which r1 causes I more distress than r2 and yet in which i1 is less morally wrong than i2. Likewise where equivalence holds between r1, i1, on the one hand and r2, i2, on the other, in respect of moral wrongness, it may nonetheless be the case that i1 is worse than i2, while r2 is worse than r1. Finally, where r1=i1 and r2=i2 in respect both of moral wrongness and distress to I, SP still need not obtain. Suppose i1 > i2 in respect of immorality: it needn't be that r1 > r2 in respect of distress. It may be that r1 < r2 in respect of distress, in which case i1 < i2 in respect of distress. And there is nothing inconsistent in the idea that an immoral action i1 may be more immoral than i2 and yet the former may cause less distress to I than the latter. So on any of these interpretations of retributive equivalence the characterization of retribution is unsatisfactory because it does not meet SP. However we might still interpret retributive equivalence as a qualitative equivalence between two qualities or properties of actions.

Now a is equivalent to b in some respect q where a specific q value is both a property of a and b. For example, a is equivalent to b in respect of weight (q) where a and b share the property of having some specific weight. I suggest that we adopt a useful piece of nomenclature here. We will distinguish between a

property and a quality. An example will make the distinction clear I hope. Weight is a quality, but weighing ten pounds is a property. In the light of this distinction we may say that qualitative equivalence between **a** and **b** is in virtue of a quality, but only exists in respect of that quality where the same property derivable from that quality may be predicated truly of both **a** and **b**. For example, weighing ten pounds is a property derivable from the quality of weight. Now we are supposing that retributive equivalence may be explicated in terms of an equivalence between two properties of two actions, one a retribution the other an immorality. The idea is that there may be some quality which the properties of the two actions share, where some property derivable from that quality may be truly predicated of both properties of the actions in question.

The first point one might make here is to raise the question of whether it is possible for a property to have a property. But this is soon dealt with as the answer is clearly yes. Being coloured is a quality a derivable property of which is being red. Now we may say meaningfully and doubtless truly that being red is not a property which may be attributed to a mental event. Thus the property of being red has a property, viz. that of not being predicable of a mental event. Closer to the present context it might be said for

example that the moral wrongness of an action is only perceivable by intuition. But if this were true then the property of being heinous would have a certain epistemic property.

Secondly, the fact that a and b share some quality does not necessarily mean that they can be equivalent in respect of that quality. For example, a and b may share the property derivable from the quality of having colour that they are both light blue. But we would not usually say that they are equivalent in respect of light blueness. An observation which is perhaps related to this is that if a and b are to be equivalent in respect of some quality q then the contents of derivable properties of that quality can be ranked vis-a-vis one another, though units for this ranking may not always be available, as, for example, where we are ranking in respect of value, or distress. Another way of putting this is that it appears to be a necessary condition of a and b being qualitatively equivalent in respect of q that the contents of properties derivable from q are linearly orderable by the '>' relation. The point about contents of properties is motivated by the following observation: if we are to rank, say, the properties of being of certain weights, the property of weighing one ton will not itself be linearly orderable vis-a-vis the property of weighing half a ton since the properties themselves have no weight at all. Rather the

content of the first property is orderable in relation
to the second.

In the light of these remarks we must ask of our
putative retributive equivalence in respect of some
quality q, whether there is such a quality derivable
properties of which (call these q-properties) might be
truly predicated of both the property of being morally
wrong to a specific degree and being distressing to a
certain degree and where such q-properties have
contents that are linearly orderable by the '>'
relation. I shall say that where a quality has
derivable properties contents of which may be ranked
vis-a-vis one another, it is a scaling notion. So, for
example, weight is a scaling notion because 1. it is a
quality and 2. contents of properties derived from it
(e.g. weighing six tons,) are linearly orderable by the
'>' relation.

Thus we are looking for some quality which is a
scaling notion and which is shared by the properties of
being wrong and distressing to a specific degree. One
possibility is that these properties share the quality
of what I shall dub 'scalability'. Let us say that a
property is scalable iff its content may be ranked vis-
a-vis other contents of properties derived from the
same quality. (For example weighing six tons is a
scalable property.) But is scalability a scaling
notion? To be such it must be that contents of

properties derived from the quality of scalability can be linearly ordered. But what is a property derived from scalability? It is (the property of) having a particular content which can be ranked vis-a-vis the contents of properties derived from the same quality. But this is only a scaling notion if properties derived from it have contents which are '>' orderable. Yet it is unclear what is meant by a property from this quality. While the quality/property distinction seems intuitively graspable via the having colour/being red example, in the present instance one is uncertain that an example of a property derived from the quality of scalability would be, say, the property of having a content of weighing six tons. But in any case even if it were correct to regard this as such a property the content of it, i.e. 'having a content such-and-such,' is not '>' orderable anyway. So scalability is a poor choice for a scaling notion.

However things are not perhaps the same with the slightly different quality of rankability, as I shall call it. The quality of rankability is that of occupying some position within a ranking scheme. Rankability will be a scaling notion iff property contents from it can be linearly ordered. But in this case such property contents will be things like occupying position six in a ranking scheme. But in view of this it can be seen that rankability will not pass

muster as a scaling notion either. The problem is not that we might specify the position of some object in a ranking scheme in a way which does not make use of a number. For example where we have a ranking in which y is placed between x and z we might specify the position of y as the position between x and z. In such a case the content of the property of occupying that position, viz. occupying the position between x and z, does not look even remotely orderable by the '>' relation. Rather the problem persists even if ranking positions are numerically specified, for the point is that although 6 and 11 are linearly orderable by '>', 'occupying position 6' (which is the content of the property in question) and 'occupying position 11' are not so orderable. Clearly occupation of position 11 is not greater than occupation of position 6. So rankability does not satisfy our criteria for being a ground of qualitative equivalence between our two properties of wrongness and distressingness to specific degrees.

I suggest here that we regard occupation of the same ranking position as a ground of equivalence, even though it is not a ground of qualitative equivalence. However if we do this I suggest that we build in the proviso that the ranking itself is carried out in accordance with a convention. For example, I may have two equinumerous groups of objects, one ranked by

height, the other by weight. Now provided that the ranking is such that, the higher the numerical position assigned to an object in the ranking the higher the degree in which it possesses that property, or the higher the degree in which the content of the property possesses that property, then occupation of the same ranking position will be taken to be a ground of equivalence. What this means is that the property of weighing six tons is to be assigned a lower ranking position than the property of weighing sixty tons, while an object weighing six tons is to be assigned a lower ranking position than one weighing sixty. This is the convention for ranking which will be assumed in what follows. It will be clear, of course, that this convention would not enable us to rank contents of properties like 'occupying position 6 in a ranking scheme.' Our justification in doing this is perhaps no greater than that it is inductive reasoning which leads to the criteria for qualitative equivalence suggested, but then the equivalence suggested in respect of occupation of numerically the same ranking position is itself data for the inductive process.

So far then we are proposing that r and i are retributively equivalent where the property of r of being distressing to a specified degree is equivalent in respect of rankability, that is, would occupy the same position in a ranking, as the property of i of

being immoral to some specific degree where the aforementioned ranking convention is adopted. Suppose we call the property of r in question **pr** and that of i **pi**. Now it is presumably clear that it is not entailed by **pi** and **pr** being equivalent in respect of rankability that i and r are when the latter are ranked according to how wrong and distressing they are respectively, nor that, conversely, i and r being equivalent in respect of rankability entails that **pi** and **pr** are, even where i and r are ranked in accordance with their possession of the properties **pi** and **pr** respectively. We are at liberty to include as many elements in the sets from which the rankings of the respective entities are taken as we please. However, in the next chapter we shall make use of the idea of a set which is maximal in respect of a certain property, i.e. a set to which no individual can be added which possesses the property in respect of which the set is maximal in a degree different from that already possessed by some member of the set. Now if **pr, pi** are ranked in maximal sets **PR, PI** respectively, where **PR** is maximal in respect of the property of having a content of being wrong to a specified degree, **PI** is maximal in respect of the property of having a content of being distressing to a certain degree, and r and i are ranked in the maximal sets **R** and **I** respectively, where **R** is maximal with respect to the property of being distressing to a

specified degree and I is maximal in respect of the property of being immoral to some specified degree, and we adopt the convention that in such sets any object possessing the property to the same degree be ranked at the same position, then if some **pr** is retributively equivalent to **pi** it is entailed that **r** (the same **r** of **pr**) is equivalent in respect of rankability to **i** (the **i** of the **pi** in question,) and vice versa. For this reason in much of what follows I shall speak of rankings of actions rather than properties of those actions, the reason being that the mutual entailment briefly mentioned above gives us confidence that, with the addition of certain assumptions left to the next chapter, such rankings and equivalences of actions only hold where the retributive equivalence proper between properties of those actions in respect of rankability holds also.

Finally, it will be recalled that two interpretations of SP have been suggested: the first in which the immorality of an action is determined without reference to the moral responsibility of the agent performing it, the second in which the moral responsibility with which the action is performed does figure in its moral evaluation. Now   the second interpretation of SP amounts to the requirement that the greater an individual's blameworthiness in the performance of an action the greater his distress for it should be, where

blameworthiness is understood as a function of the
wrongness of an action (wrongness according to the
first interpretation of SP) and the agent's moral
responsibility for it. This suggests a second
characterization of retribution as follows: r is
retributively equivalent to a blameworthiness b where r
and b are equivalent in respect of rankability. The
attentive reader will note two things here. Firstly, we
must relax the condition that retribution be for an
immoral action alone, one reason for the introduction
of which was that retribution be not for some aspect of
a person's character. Now it will be objected that
blameworthiness is surely part of someone's character.
Yet I think that the relaxation of the condition is
justified in the present case because blameworthiness
is tied to an action or set of actions. What I sought
to avoid was retribution as a response to character
irrespective of action. I see no impropriety in the
idea of an evil man who has never actually done
anything which has harmed others than himself. This
sort of person would not be a fit subject of
retribution as I am understanding it. Retribution must
be in some sense a response to action, though this may
be taken as no more than stipulative by those who feel
that it may encompass more. And it is because
blameworthiness is tied to action that it seems unduly

restrictive not to admit it as a proper subject for a
retributive response.

Secondly, the equivalence between a blameworthiness
and an action is one between things rather than their
properties. However, in view of considerations entirely
analogous to those given above concerning equivalence
in respect of rankability between actions where such
actions are drawn from special maximal sets, it turns
out that equivalence in respect of rankability between
a blameworthiness and an action causing distress holds
in virtue of rankings of these objects in maximal sets
iff properties of these objects, like how distressing
they are and how morally serious they are, are also
equivalent in respect of rankability where rankings of
properties too are taken from maximal sets of a certain
kind.

Of course these remarks need to be clarified and
considerably amplified in certain ways and it is to
this project of amplification and clarification which
we will turn our attention in the next chapter.

Chapter 2. Retributive criteria

(i) Cardinal rankability and maximal sets.

Let us proceed then to give an account of a ranking
in which locations within it are assigned numerical
expressions. In line with our expressed strategy we
shall look at the case of ranking actions.

Consider what we might call a cardinal ranking.
Suppose we have two groups **X, Y,** of actions and these
actions are ranked respectively in respect of
immorality and how distressing they are to their
target. Then by assigning a number (the same number) to
the least wrong and distressing actions in the ordering
and successive numbers to the successively greater
actions, greater in respect of their immorality and
unpleasantness to their target, such that if some act
ak is of greater unpleasantness/immorality than some
other **ai** to which we have assigned the number **n** and
there is no other retributive/immoral act of the group
aj such that **aj** is greater than **ai** but smaller than **ak,**
then we assign to **ak** the number **s(n),** where **s(n)**
denotes the successor of **n,** where **n** and **s(n)** are

natural numbers. Where **X** and **Y** have six members apiece
we could get something as follows:

	X	**Y**
1	i1	r1
2	i2	r2
3	i3	r3
4	i4	r4
5	i5	r5
6	i6	r6

Clearly here the numbers which succeed the r's and
i's are chosen once their rank ordering position is
known but we could have chosen other names which would
not presuppose such knowledge.

Now it seems possible on the strength of this sort of
simple ordering scheme to assert truly of two actions
on the different scales that they occupy the same
relative position on their ranking scales. (We can
think of the **X** column as a ranking scale of actions in
respect of immorality, the **Y** column as a ranking scale
of actions in respect of distress to target.) Hence
where i3 and r3 have the quality of rankability, that
is they occupy specific positions in certain rankings,
then they could be equivalent in respect of rankability
where the property of occupying a particular position
on some scale (though not necessarily the same one)

which may be truly predicated of them is the same. Let us call this relation equivalence in respect of cardinal rankability.

But this equivalence is a rather vacuous relation. Consider the case in which we have the two rank orderings of putative retributions or distressing actions and immoral actions as follows:

immoral acts	retributions
1 i1	r1
2 i2	r2
3 i3	r3
4 i4	r4
5 i5	
6 i6	

Now we are at liberty to select whichever sets of putative retributions and immoral acts we like provided that the greater the number assigned to the act from a particular set the greater the immorality/distress of that action within that set and that the numbering within the sets adheres to the criterion that the successor function determine the numerical location of the next greatest object in the set. But then if we suppose that r1 represents a fine of £10 and r4 represents hanging while i1 represents a parking offence and i4 a fourth parking offence

(assuming the latter to be morally worse than the former) then there is an equivalence in respect of cardinal rankability between committing a fourth parking offence and hanging someone.

Moreover there is no second term of an equivalence relation available for i6 (suppose this to be the act of not declaring all one's earnings to the tax man.) So the quality of rankability generates in some cases the ground for a rather 'peculiar' equivalence, in others a ground for no equivalence at all.

The second kind of problem, i.e. that of there not being a second term for an equivalence relation in respect of cardinal rankability for some terms, would not be removed if we were to specify that the equivalence between immoral actions and retributions had to be based on equivalences of ranking positions within maximal sets of immoral actions and retributions in respect of some quality. Let us say that a set of actions is maximal in respect of some quality q where any action which could be added to the set would occupy, in virtue of the possession of the q to the same degree as some object already within the set, an already existing number position in the ranking for that set in respect of q and its addition would not displace any other members of the set from their number position. (This involves the assumption of course that more than one object may occupy the same ranking

position.) This definition of a maximal set entails that we may have different sized maximal sets of objects in respect of some quality, that is, we may have maximal sets of objects ranked in respect of the same quality with different numbers of members, though it is not possible that two maximal sets of objects in respect of the same quality q should have different numbers of ranking positions. For example, the maximal set of actions in respect of q might contain one or one thousand objects ranked at position 15 and yet it may be true of the set in either case that it is maximal simply because no addition of an element can be made which will affect the location of objects within the existing ranking scheme. In view of this I suggest we make a distinction between a maximal set simpliciter and a restricted maximal set. The latter will be a maximal set which comprises objects such that only one object occupies each ranking position within it. In what follows I shall assume that we are talking about restricted maximal.

So we might specify that equivalence in respect of cardinal rankability between an immorality and a supposedly retributive act is only of interest to us where the rankings of the actions in respect of the qualities of moral wrongness and distress respectively is within maximal sets in respect of these qualities. But it still might happen that the number of ranking

positions in a set of actions which was maximal in respect of unpleasantness would be different to the number of ranking positions in a set of actions which was maximal in respect of moral wrongness. So the introduction of the idea of sets which are maximal in respect of certain qualities does not remove the problem of the possibility of cases in which the equivalence relation will not hold. What this means is that we cannot rule out the possibility of there being actions causing distress to which no morally wrong action ranked in respect of its wrongness is equivalent in respect of cardinal rankability or the possibility of there being actions of a moral wrongness or gravity to which no action distressing to a target individual is equivalent. Perhaps this is morally irrelevant. For example, it may be that according to a cardinal ordering on the sets of actions maximal in respect of unpleasantness and moral wrongness there is no immorality which, in virtue of its cardinal number placement within the set of actions maximal in respect of immorality, is equivalent to the act of torturing a target individual for thirty years.

Secondly, the fact that we are using maximal sets of actions will not guarantee that putative retributions and immoralities equivalent in respect of cardinal rankability will strike one as intuitively correct. For example it may be that in the set of immoral actions

there are many shades of minor immorality which are distinguishable in rank ordering according to moral wrongness and yet few discernible grades of distress at the lower end of the scale. But this could mean that i50 was something like petty pilfering while r50 was 15 years in prison. (Whatever the intuitive relation is thought to be something like r50 and i50 would not be related by it.)

So equivalence in respect of cardinal rankability, though coherent enough, may fail to deliver equivalences between immoral actions and actions distressing to someone which have the seal of approval of intuitions, even where the actions are ranked within maximal sets in respect of the qualities of moral wrongness and distressingness to target individual respectively. Henceforth we shall refer to equivalence between two objects in respect of cardinal rankability, where the property of rankability, i.e. the ranking position of the action in question, is assigned to an object in respect of its position within a maximal set in respect of possession of some property, as maximal cardinal rankability, blessedly, MCR equivalence for short. One may note in passing here that the fact that two objects may be MCR equivalent and yet not possess the intuitive fit spoken of already is not to say that what we have termed simple proportionality may not hold between them. In fact simple proportionality must hold

between objects from the maximal sets of actions in respect of moral wrongness and distress to target which are MCR equivalent. If we take two immoral actions, **ix**, **iy**, ranked appropriately in the maximal set of actions in respect of moral wrongness, and two actions of unpleasantness to their targets, **rx**, **ry**, also ranked appropriately in the maximal set of actions in respect of unpleasantness to target of action, then where we substitute the same numerical expressions for 'x' and 'y' it will be true, in virtue of the way we designate such actions, that if, in respect of immorality **iy** > **ix** then, in respect of distress, **ry** > **rx**. This is because if we substitute the same numbers for **x**, **y** it will be true that **y** > **x**. This is because we know that **iy** is more immoral than **ix**, so the number substituted for **y** must be numerically greater than that substituted for **x**. But then **ry** > **rx** in respect of distressingness to target given the way that the numerical suffixes of **ry**, **rx** are determined.

So where immoral actions and their putative retributions are MCR equivalent simple proportionality obtains. Moreover it is in a sense true that the equivalence is one between how bad an action is and how unpleasant the putative response is, and for some at least this is part of what retributive equivalence is about. To see this one need only note that it is precisely how wrong one action in the pair is and how

distressing is its counterpart which determine the ranking position in respect of which their equivalence obtains. Let us assume then that MCR equivalence of a certain kind between actions may serve as a characterization of so-called retributive equivalence. It may not yet be perfect but it is at least worthy of further consideration. Further, we will assume that the immoral actions we have in mind so far are ranked as a result of a moral evaluation which takes no account of the moral responsibility of the agent performing the action in question. Thus the account of retributive equivalence we have meets SP on its first interpretation. (Of course if moral wrongs are evaluated in line with the second interpretation of SP retributive equivalence between wrongs so evaluated and distressing actions will also meet the requirement of SP on its second interpretation.) In view of this interpretation of moral wrongs room is left for a characterization of retributive equivalence in respect of MCR equivalence where blameworthinesses and distressing actions are ranked respectively according to their degrees of badness and unpleasantness vis-a-vis the other members of maximal sets of blameworthinesses in respect of badness or seriousness of blameworthiness and of distressing actions in respect of how unpleasant they are.

A few words are perhaps in order here about the
ascertaining of the maximal cardinal ranking position
or property of maximal cardinal rankability of some
object o. Firstly, and this may be obvious, o may have
more than one maximal cardinal ranking position
depending on how the ranking in the maximal set within
which it is ranked proceeds. The ranking of the maximal
set is in respect of some quality but obviously some
other quality might be chosen with which to construct
and rank members of a maximal set. Moreover members of
maximal sets of x's in respect of different qualities
may not even be equinumerous. Secondly, a maximal
cardinal ranking, call it MCR, is the result of the
construction of a set of objects which are such that in
respect of some quality q no object can be added to the
set and ranked in respect of its possession of q and
displace any other object in the set from its previous
ranking position (it will be remembered that more than
one object may occupy the same ranking position.) In
theory we may construct an MCR from any set the members
of which may be ranked vis-a-vis one another.

Clearly if retributions qua distressing actions are
to be MCR equivalent to blameworthinesses where both
take their rankings from their positions vis-a-vis the
other members of the relevant maximal sets then the
problems concerning items for which there is no MCR
equivalent counterpart and the possibility of the

equivalences not getting the seal of approval of
intuitions (whatever this amounts to) face us just as
they did where we were dealing with equivalences
between immoral and distressing actions. In the next
section we must turn our attention to the intuitive
fittingness which supposedly obtains between
retribution proper on the one hand and a
blameworthiness or immoral act on the other.

(ii) Intuitive criteria of retribution and the maxims
of crude proportionality.

What criteria might be involved where a putative
retribution and immoral act or blameworthiness have the
seal of approval of intuitions? A term often used to
describe the relation of intuitive fit between such
terms is proportionality. It is often supposed that
although there can be no interesting equivalence
between a blameworthiness or moral wrongness on the one
hand and a distressing action on the other there can at
least be a proportionality between such terms. As we
have seen, in the case of sets of morally wrong and
unpleasant actions, proportionality may be construed as
the requirement that the more morally wrong an action
is the greater must be the distress to its target of
the retribution which accompanies it. It may also be
thought of as the analogous requirement that the

greater the blameworthiness of an agent in the performance of a morally wrong action the greater the putative retribution for it must be. (Strictly, of course, this obtains where SP on its second interpretation is met.) However if we have two sets of actions, or one of blameworthinesses, the other of actions, such that each member of one set is MCR equivalent to a member of the other then for any pair of pairs of objects which are MCR equivalent simple proportionality holds between the pairs. So, given the blameworthinesses **bx, by,** and the retributive actions, **rx, ry,** if **bx** > **by** then **rx** > **ry**, where we substitute the same numerical expression for **x, y,** throughout. This will be so in virtue of the way in which the rankings of objects which are MCR equivalent has been constructed. But, as was pointed out, the fact that two objects are MCR equivalent does not ensure that the intuitive fit required obtains between these objects. Clearly if the relation we have described as 'intuitive fit' is to be thought of as the relation of proportionality this latter cannot be conceived of as simple proportionality.

Consider first as a replacement for simple proportionality the relation of numerical proportionality which holds between the two sets of values A and B:

```
A    1    2    3    4      n-1   n

B    2    4    6    8      2n-2  2n
```

Line no.   (i) (ii)(iii)(iv)   (n-1) (n)

We may think of (n-1) and (n) where they appear in
brackets as line no. indicators as expressions
denoting, respectively, the numerical expressions 'n-1'
and 'n' in the symbolism of Roman numerals. According
to this table we can identify values in it by citing
their reference no. the first part of which will be
either A or B to indicate that the value in question is
in the A row or B row, the second part of which will be
a line no. to indicate that the value in question is to
be found in that line or column. Now we will say that A
is numerically proportional to B because the following
relation holds between individual A and B values:

  A(i) : B(i) = A(ii) : B(ii) .... = A(n) : B(n), where
X : Y represents the ratio of X to Y. Call this NP1.
But where NP1 obtains it follows, mathematically, that
the following relation, NP2, holds also:

  A(i) : A(ii) = B(i) : B(ii), A(ii) : A(iii) =
B(ii) : B(iii),....A(n-1) : A(n) = B(n-1) : B(n). The
reason for this is that if, e.g. A(n-1) : B(n-1) =
 A(n) : B(n) then A(n-1)/B(n-1) = A(n)/B(n) and so
A(n-1)/A(n) = B(n-1)/B(n), but then since ratios may be
expressed as fractions this entails that

A(n-1) : A(n) = B(n-1) : B(n). Also where X,Y,Z and

A,B,C designate numbers if X : Y = A : B and Y : Z =

B : C then X : Y : Z = A : B : C. But then since

A(i) : A(ii) = B(i) : B(ii), A(ii) : A(iii) =

B(ii) : B(iii) then clearly A(i) : A(ii) : ... A(n) =

B(i) : B(ii) : ...B(n). Call this NP3. Moreover where

either of NP3 or NP2 holds then numerical

proportionality in the form of the relation NP1 holds

also.

But now suppose we have a ranking of

blameworthinesses and distressing actions in the X and

Y rows respectively:

X	B1	B2	(...)	B(n-1)	B(n)
Y	D1	D2	(...)	D(m-1)	D(m)

L no.	(i)	(ii)		(m-1)/(n-1)	(m)/(n)

Let us suppose that X and Y are maximal sets of

blameworthinesses and distressing actions respectively.

This means that we could not take an action causing its

target a certain distress and rank the unpleasantness

of that action in the Y list and change the ranking

position of any distressing action already in the

table. At most the only effect of this would be to have

an extra expression designating a distressing action at

some ranking position already occupied. Likewise for a

blameworthiness, we could not select an action of a certain kind performed with a certain degree of moral responsibility under certain circumstances such that the blameworthiness of the agent thereof would affect the ordering of the ranking **X**. Secondly, a word about the line numbering. We do not know that n designates the same number as **m** and so **B(n)** and **D(m)** may not be at the same line number. This also applies to the preceding B and D values for **B(n)** and **D(m)**, that is, **B(n-1)** and **D(n-1)**. In view of this no single line number is given for such values. Lastly, we may suppose that the greater the line number the greater the blameworthiness or distress caused by the action cited at that line, and that the numerical suffix of a B or D expression is identical with the number expressed by its line number.

Now it is immediately apparent that we cannot be certain that numerical proportionality holds. For this holds where for any line number j, **X(j)** : **Y(j)** represents the same ratio as the ratio of the **X** value to the **Y** value for any other line number. But now suppose that n > **m**. Then there is a blameworthiness designated by **B(n)** such that it is not true that **B(n)** : **D(n)** = **B(i)** : **D(i)**. This is because there is no distressing action designated by **D(n)**. Or where n < **m** it will be true again that the ratio **B(m)** : **D(m)** is not equal to **B(i)** : **D(i)** for there is no blameworthiness

designated by **B(m)**. But we cannot be sure that **n=m** and so we can't be sure that neither the relation **n > m** nor **n < m** holds. And we cannot be sure that **n=m** for the reason that the compilation of a maximal set in this context does not appear to be a task that we can ever be sure we have completed. For it is always logically possible that there be some situation which we have not taken into account which would yield an action of a degree of unpleasantness or a blameworthiness of a certain degree of badness such that some **D** or **B** value could be interposed between the objects at a number position and its successor in the **X** or **Y** set hence increasing the **m** or **n** value for that maximal set (where **n** or **m** designates the number assigned to its largest member.) Hence we can never be sure that we have arrived at the correct estimation of the value of **n, m,** or that **n=m**, and so the relation required for numerical proportionality is not one the obtaining of which we we can be sure we have achieved.

There is a more vitiating difficulty however. Two rankings may be numerically proportional to one another because objects within them are not simply rankable, but because quantifying numerical expressions are available in virtue of which a numerical ranking may be produced. (Although this doesn't mean, of course, that both sets of objects can be ranked in virtue of these quantifying numerical expressions in a single scale.)

But where numerical expressions for qualities or degrees of objects are not available then it cannot be that the relations on which numerical proportionality depends hold because these relations are arithmetical. For example, for it to be the case that **A1 : B1 = A2 : B2** it must be that **A1.B2 = A2.B1.** But multiplication is a function the arguments of which must be numbers. However for there to be numerical expressions in terms of which rankings of blameworthinesses and distressing actions might be made there must be units of blameworthiness and distress, which patently there are not. Therefore we must conclude that numerical proportionality cannot hold between rankings of blameworthinesses and distressing actions.

In place of numerical proportionality I shall propose a set of criteria which may give some definition to what is required for an 'intuitive fit' to obtain between a retribution and a blameworthiness or immorality. I shall call these the maxims of crude proportionality and express them initially in terms of what they require of the relation between a retribution and a blameworthiness. These maxims are only to be thought of as applying to a range of blameworthinesses and distressing actions. I shall first explain roughly what it is in this context for a relation to hold for a certain range of values and then will state the maxims.

In the next section we will turn to the question of
their substantiation.

Now if we represent blameworthinesses by the
expressions **b1...bn** and distressing actions by **d1...dm,**
assigning a **b** or **d** representation as the distressing
action or blameworthiness in question is assigned a
particular cardinal position in its maximal set in
accordance with the '>' relation (the suffix of its
representation being identical with the cardinal
ranking position occupied by the object so
represented,) then it will be true necessarily that in
virtue of uniform substitution for **k, i** if **bk > bi** then
dk > di within a certain range. Let me explain what it
is for this relation to hold within a certain range. As
has already been noted the maximal sets of
blameworthinesses and distressing actions may not be
equinumerous and so it may not be that **m=n.** Suppose in
fact that **m < n.** Then where we substitute **n** for **k** in
the above conditional then if **bn > bi** it must be that
n > i if the relation is to hold. But it is not true
that **dn > di** since, ex hypothesi, there is no **dn.**
However provided that we substitute for **k** no value
greater than **m** we will avoid this difficulty. Hence we
may say that the relation 'if **bk > bi** then **dk > di**'
holds where the values of **k** range between 2 and **m.**
Obviously **k** cannot take the value 1 because this would
mean that the substitution for **i** would have to be a

number which does not form part of a representation of a blameworthiness or distressing action. So the relation holds where 1 < **k** < **m+1** on the understanding that **k** must be a natural number.

Where a distressing action corresponds to a blameworthiness or a blameworthiness to a distressing action iff the representation of the pair of objects has the same numerical expression as suffix the maxims of crude proportionality are as follows:

1. where some blameworthiness **ba** is much/far/a lot greater than some other **bc**, then the distressing action corresponding to **ba**, call it **da**, is much greater than the distressing action which corresponds to the latter blameworthiness **bc**, call it **dc**. Moreover since it appears to be entailed by 'a is much greater than b' that **b** is much smaller/less than **a** we have as an entailment of condition 1 being met that 1a is met, namely, the condition that where **bc** is much less than **ba** then **dc** is much less than **da**.

2. where some blameworthiness **ba** is a little greater than some other **bc** then the corresponding distressing action **da** is a little greater than **dc**. Since again it appears to follow from the fact that a is a little greater than b that **b** is a little smaller/less than **a** we may say again that where condition 2 is met so is 2a, namely, that where **bc** is a little smaller/less than **ba** then **dc** is a little smaller/less than **da**.

3. where some blameworthiness **ba** is quite a lot/a
fair bit/ considerably greater than some other **bc** then
the corresponding distressing action **da** is quite a lot/
etc... greater than **dc**. Again since it appears to be
entailed by '**a** is quite a lot greater than **b**' that **b** is
quite a lot smaller/less than **a** we may say that where
condition 3 is met then so also is condition 3a,
namely, where **bc** is quite a lot smaller/less than **ba**
then **dc** is quite a lot smaller/less than **da**.

(iii) Truth conditions for the maxims of crude
proportionality.

Let us consider the first condition or maxim of crude
proportionality. Suppose we try to substantiate the
(doubtless true) claim that the blameworthiness of some
fully morally responsible agent in the commission of an
act of rape and murder is much greater than that of one
who double parks in a built-up area with equal moral
responsibility. Suppose the correct representations of
these two blameworthinesses in the ranking of the
maximal set of blameworthinesses are **b75** and **b10**. Now I
think much of the work of justifying the claims we must
justify if we are to show that condition 1 has been met
is done by claims in which the self-effacing word
'most' figures prominently. Clearly this is a word
which must be given clear criteria for its application

if we are to get anywhere. First, however, by way of
introduction let us consider for the moment a few
examples of the use of the phrase 'a lot,' though not
here figuring, as in the maxims, as an adverbial
modifier.

(i) A lot of children in the class have measles.

(ii) Many (a lot) of the residents of Blackheath are
well-to-do.

(iii) George has lost a lot of weight. His weight has
dropped from 18 to 12 stone(s.)

Now apparently there is no straightforward
arithmetical qualification which must be met if we are
to use the phrase 'a lot' appropriately. In (i) the
class may comprise 40 pupils while only 12 have
measles. This represents only 30% of the class. On the
other hand it may be that in (ii) 70% of the residents
(of employable age) of Blackheath are well-to-do. Of
course it may be that if this were the case it would be
more informative to say that most of or the majority of
the residents were well-to-do but it would nonetheless
be entailed by this that many or a lot of the residents
were well-to-do also. Finally, in (iii) six stones is a
great weight loss, but it represents a weight loss of
40%. So 'a lot of x are y' is not substantiated by some
percentage of x being y in any straightforward way.

However it does seem that there is a different thread
of consistency running between the three cases. In the

first case 12 out of a class of 40 is a lot of children to have just one reason for absence. In other words the implicit comparison is with other cases of numbers of absentees and we find that in most of such cases absenteeism is due to a wider variety of causes. We should say that in most cases fewer absences as a percentage are attributable to one cause. In (ii) in most areas the percentage of well-to-do residences is not so high. In (iii) most cases of weight loss are less than that of George in the example. So if we were to take the phrase schema 'a lot of x's are y' as an example it is not that the percentage 100x/y by itself determines whether the use of the phrase 'a lot' is appropriate but rather that a certain comparison of 100x/y with the percentages in other like cases, e.g. weight loss, well-to-do residency in a district, etc., is possible and yields certain results. It is as a result of making this comparison that it can be seen whether a certain relation holds, viz. in the case under discussion, whether most like cases are governed by the same sort of percentage. And it is in this relation perceivable in the comparison to like cases that the use of the word 'most' figures. In other words 'a lot of x's are y' is true with appropriate substitutions where in most cases in which x's are y the percentage given by the formula 100x/y is not so high as it is in the case under consideration. But to

say this does not at first sight appear to be a step

forward since 'most' seems a word no less prone to

vagueness than terms like 'a lot' or 'quite a lot'.

   The word 'most' appears to function in a manner

similar to the phrase 'the majority' in one of the

latter's senses. In one of its senses 'the majority'

means anything over a half. Hence 50.0001 % represents

a majority. But where there is such a majority it would

not be appropriate to use the word 'most'. It is

because just over 50% constitutes a majority that

epithets are introduced to characterize further just

what kind of majority is at issue. We speak of the

'overwhelming' and the 'great' majorities; 90%

constitutes an overwhelming majority perhaps while 80%

constitutes the great majority. Now perhaps it is true

that most x are y iff the great majority of x are y.

Since it is entailed by 'the overwhelming majority of x

are y' that 'the great majority of x are y' but

probably not vice versa it is entailed by the former

that most x are y, but not vice versa.

   If 'most x are y' is true in the same conditions as

'the great majority of x are y' then certain conditions

clearly make 'most x are y' true and others clearly

make it false. Clearly if 50% of x are y then it is

false and if 90% of x are y it is true. But there will

be other percentages where it is not clear whether the

claim is true or false. For example, if 68% of x are y

is it true that the great majority of x are y? Hence
there will be cases where we are not sure whether 'most
x are y' is true and so cases in which we will not be
sure that the claim using the phrase 'a lot' is true,
if this latter sort of claim is dependent on the truth
of a claim involving the use of the word 'most'.

Consider for a moment the limitations inherent in the
use of words like 'most'. In the previous paragraph
cases have been spoken of where we are not sure whether
the claim involving use of the word 'most' is true or
false, for example, where 68% of x are y we are unsure
whether it is true that most x are y. It is important
to note that our uncertainty concerning the truth or
falsehood of certain propositions under certain
circumstances, such as those using the word 'most', is
ineradicable. It is not that we could remove this doubt
under certain conditions. We cannot stand in a better
light and see the meaning of the word 'most' any
better. Hence it is not that the proposition in
question is true or false in these cases where we
cannot decide, only we just can't find out about it;
rather in view of the meaning of the proposition itself
it is neither true nor false in such cases. But it
follows from this that the proposition is not true
under such circumstances.

This sort of reflection might lead us to think that
we could at least determine the range of cases in which

the proposition (with suitable substitution instances)
"most x's are y" is not true by simply conjoining two
ranges of cases, the first in which the proposition was
false, the second in which it was neither true nor
false. For example, it might be thought that if "most
x's are y" is false wherever the percentage of x's
which are y, denoted by p, is such that $60 > p \gg 0$
(where $x \gg y$ means x is greater than or equal to y)
and if it is neither true nor false where $70 \gg p \gg 60$
then the proposition is not true where $70 \gg p \gg 0$.
But this is to misrepresent the matter. For it is not
possible to determine the ranges of percentage value
for p such that the proposition "most x's are y' is
true or false or neither true nor false. This is
because there are no exact boundaries to groups of
cases which determine these different truth values for
the proposition in question. For instance, because
there are no exact boundaries to the group of cases,
i.e. the values of p, which make the claim "most x's
are y" neither true nor false there will be no exact
boundaries for the adjacent groups which determine the
other truth values, that is, true and false.

To see that no exact boundaries are available we need
only reflect on the fact that where we do have an exact
boundary we are able to speak of cases different to an
infinitesimal degree which determine different truth
values. But the latter is not possible and so there

cannot be these exact boundaries. Let me explain what I
mean by way of an example. Suppose that "most x's are
y" is true where the percentage of x's which are y is p
such that $100 > p \gg 70$. Then, given the existence of a
range of cases where the proposition is neither true
nor false, it would have to be the case that where
69.999% of x's are y the claim is neither true nor
false. The fact that we cannot have very slightly
different percentages determining different truth
values gives evidence that there are no exact
boundaries to these ranges of percentage values
determining different truth values. Now I have spoken
of clear cases, that is cases in which a proposition is
obviously true or false. However we could extend this
idea of a clear case to include cases which are
obviously neither true nor false, that is, which fall
clearly within the range of percentage values for which
the proposition ' most x's are y' is neither true nor
false. I take p = 65% to be one such. Thus we are left
with a picture of ranges of percentage values with
inexact boundaries the core values of which determine
clear cases of the truth value associated with that
range. There is always, so to speak, a lacuna between
the clear cases within different truth-value-
determining ranges.

The upshot of this is that we cannot propose (without
it being a mere stipulation of course) a range of

percentages with exact boundaries such that where the percentage of **x**'s which are **y** falls within those boundaries the claim that most **x**'s are **y** is true. Thus we cannot give a clear indication of the truth conditions of a proposition of the form 'most x are y'.

So we cannot truly claim that "most **x**'s are **y**" is not true or not false within certain precise limits for the percentage p of **x**'s which are **y**. However, I don't think we need be disheartened by this. Provided that we can identify situations in which "most **x**'s are **y**" is clearly true, clearly false or clearly neither true nor false, that is, as I have called them, clear cases of truth, falsehood etc. for the proposition, then some claims that a certain relation of proportionality holds, namely those which depend for their truth on the truth of claims involving use of the word 'most' in situations which are clear cases of truth, falsehood etc. for the proposition, will be decidable in the following sense: we will be able to determine whether the relation obtains (which will be where the proposition involving 'most' is clearly true) or does not obtain (which will be where the proposition using the word 'most' is either clearly not true or clearly false.)

The maxims of crude proportionality, i.e. the three conditions so far proposed, are in fact inadequate for our needs on two counts. Firstly, although it is

necessarily true that if if **bk** > **bi** then **dk** > **di** then
if **dk** > **di** then **bk** > **bi**, where **m** > **k, n** > **k,** simply in
virtue of the fact that the same numerical expressions
are substituted for 'k' and 'i', it is not obviously
the case that necessarily if if **bk** is much greater than
bi then **dk** is much greater than **di** then if **dk** is much
greater than **di** then **bk** is much greater than **bi**. And
yet the crude proportionality we are proposing surely
requires that the conditional 'if **dk** is much greater
than **di** then **bk** is much greater than **bi**' should hold as
well as the conditional ' if **bk** is much greater than **bi**
then **dk** is much greater than **di**'. Likewise with the
second and third maxims the converse conditional must
hold also. Hence the maxims of crude proportionality
must be supplemented to take account of this.

Secondly, it must be made explicit that the maxims
hold for a certain range of values for the number
suffixes of b's and d's. Obviously there may be some
proportionality beyond simple proportionality between
blameworthinesses and distressing actions even if their
maximal sets are not equinumerous. We might regard the
condition of crude proportionality as being met where
if two series of blameworthinesses and distressing
actions constituting maximal sets are respectively
numbered from 1 to **n** and 1 to **m** then crude
proportionality holds for all number values of b's and
d's up to whichever of **n** and **m** is the smaller.

Taking these two deficiencies into account we get the following revised maxims of crude proportionality:

for all number suffixes of **b**'s and **d**'s up to the number of members of that maximal set of blameworthinesses or distressing actions which is smaller-

1. if **bk** is much greater than **bi** then **dk** is much greater than **di** and if **dk** is much greater than **di** then **bk** is much greater than **bi**.

2. if **bk** is quite a lot greater than **bi** then **dk** is quite a lot greater than **di** and if **dk** is quite a lot greater than **di** then **bk** is quite a lot greater than **bi**.

3. if **bk** is a little greater than **bi** then **dk** is a little greater than **di** and if **dk** is a little greater than **di** then **bk** is a little greater than **bi**.

As before these conditions will have certain corollaries in terms of corresponding relations. For example, the relation 'is much greater than' is paired with the relation 'is much smaller/less than' such that if **bk** is much greater than **bi** then it follows that **bi** is much smaller/less than **bk**.

So let us now take our earlier example of the claim that the blameworthiness of an agent who rapes and murders with full moral responsibility is much greater than one who double parks. How would we justify this claim? One way we might do this would be to cite the large number of intervening different degrees of

blameworthiness between the two blameworthinesses at issue. If, according to the earlier supposition, we could represent the former blameworthiness by **b75** in a maximal set of blameworthinesses in respect of moral gravity and the latter by **b10** it could be pointed out that 64 **b** values lie between the two and that 64 is, in an as yet unclarified sense, a lot of blameworthinesses. But the claim that there are a lot of blameworthinesses between the two has to be disambiguated. Consider for a moment the claim that **bk** is much greater than **bi** where this is backed up by the further claim that there are a lot of blameworthinesses of differing degree ranked between **bk** and **bi**. Now the justifying claim may be taken in at least two ways: firstly, it may mean that between **bk, bi** there is an unusually large number of intervening blameworthinesses for the particular degree of difference between the two. The claim would then be set against a putative backdrop in which in most cases where one blameworthiness is much greater than another there aren't so many intervening blameworthinesses. Alternatively, the justifying claim may just mean that a lot of the members of the maximal set of blameworthinesses are interposable, i.e. may be placed in their appropriate ranking positions between **bk** and **bi** where what this amounts to is that in most cases the number of blameworthinesses (that is, degrees of

blameworthiness represented by differing **b**-expressions)
interposable between any two **b** values of the
appropriate maximal set is smaller than it is in this
case. So we might suppose a difference in sense between
the two claims in which the phrase ' a lot' figures is
parallelled by the different claims in justification of
them, i.e. the different claims in which our word
'most' plays a key role. On the former interpretation
of the claim that there are a lot of **b** values between
bk and **bi**, call it sense 1, (nb. sense 1 is that there
are a lot of interposable degrees of blameworthiness
for whatever relation holds between them,) where the
claim is true it is not necessarily the case that a lot
of the members of the maximal set are interposable
between **bk** and **bi** on its second interpretation, call
this sense 2 of the claim; it is just that the claim on
sense 1 is true where by comparison with most other
cases in which a certain relation holds between two
blameworthinesses the number of interposable differing
degrees of blameworthiness between the two
blameworthinesses in the present case is greater. Of
course where it is always the case that if **bk** is much
greater than **bi** there are a lot of interposable degrees
of blameworthiness in the second sense between the two
we might still endorse the claim in the first sense on
the grounds that not only are a lot of the members of
the maximal set of blameworthinesses interposable

between **bk** and **bi** but in most cases where there is such
a degree difference between the two the number of
interposable blameworthinesses is less.

I shall suppose here that it is the truth of the
claim in sense 2 which is required to justify the claim
that **bk** is much greater than **bi**. At the very least it
would seem that the claim in sense 2 must be true if **bk**
is to be much greater than **bi**, and the truth of the
claim in sense 1 would not guarantee its truth in sense
2 as we have noted. Perhaps where the claim is true in
sense 2 and in sense 1 it is true that **bk** is very much
as opposed to simply much greater than **bi**. Thus the
truth of the claim in sense 1 looks more like a
necessary condition of **bk** being very much greater than
bi, rather than a necessary condition of its being
simply much greater than **bi**; nor is the truth of the
claim in sense 1 a sufficient condition of the truth of
the claim that **bk** is much greater than **bi**. Again, as we
have seen, it is consistent with the truth of the claim
in sense 1 that it is not the case that a lot of the
members of the maximal set of blameworthinesses may be
interposed between **bk** and **bi** in sense 2.

So we might try to justify the claim that **b75** is much
worse/much greater than **b10** by showing that in the
second sense above a lot of the blameworthinesses of
the restricted maximal set of blameworthinesses, i.e.
that a lot of the b values, are interposable between

b75 and **b10** and so the set made up of the
blameworthinesses which appear in the ordering of
blameworthinesses between **b75** and **b10** is larger than
most interposable sets.

Two important points are to be noted here. Firstly we
must use a restricted maximal set in the context of
this discussion. The reason for this is that the claim
that there are a lot of blameworthinesses between two
given blameworthinesses is surely not about how many
examples of each differing degree of blameworthiness
there are but rather about how many discernibly
different degrees of blameworthiness stand between the
two given blameworthinesses. A restricted maximal set
it will be remembered has only one blameworthiness at
each different degree of badness of blameworthiness.

Secondly it should be noted that there will be more
interposable sets of some sizes than of others. For
example, the interposable set for **bn** and **b1** will have
n-2 members, but no other interposable set will have
that many members. On the other hand the interposable
set with no members will have many exemplars: the
interposable set for **b1** and **b2**, for **b2** and **b3**, etc., up
to the set for **bn-1** and **bn**. In other words there will

be n-1 interposable sets with zero members but only one
interposable set with n-2 members. This means that in
general there will be more interposable sets of a
smaller size than a larger simply in virtue of the

92

combinations of **b** values available for smaller-sized interposable sets as opposed to larger ones.

The important point to be noted in all this for us is that the number of interposable sets smaller than some specified interposable set would be larger than the number of different sizes of interposable set smaller than that specified set. Thus when it comes to an assessment of the number of interposable sets smaller or larger than some given interposable set we have an option concerning which way we calculate such a figure. I shall choose to calculate the size of the comparison group of interposable sets by reference to the number of different sizes of interposable set, rather than the number of distinct interposable sets, whether of different size or not. This has two virtues: firstly it simplifies our calculation; secondly, it does some justice to the intuition that, if indeed **b75** is much greater than **b10**, this is because the size of the 'gap' between the two **b** values is larger than that between most of such values. But of course there is an ambiguity in the use of the word 'most' here. It could be that the comparison is between the interposable set under scrutiny and the number of other sized sets where sets are distinct in virtue of size alone or in virtue of which values determine that interposable set.

It ought to be pointed out here that there is some justification for rejecting the simpler view I take that

we count interposable sets as one if they are of the same size. The point of saying that George has lost a lot of weight when his weight drops from 18 to 12 stone(s) is in part to introduce an implicit comparison group in which most weight losses are not only smaller but if say we were to count the weight losses of one amount and compare it with the number of weight losses of George's kind there would be more of the former than the latter. In other words our method of counting would be sensitive to the number of identically-sized interposable sets. For those who think it essential to the use of the word 'most' that we take into account not only the number of differently sized interposable sets counted as distinct by size alone criteria for the maxims of crude proportionality may still be given. The calculation of whether such claims in justifying the maxims are true or false will be complicated by the different principle of counting from the one we use but this will not affect the manner in which one would determine whether crude proportionality holds. This does not gainsay, however, that introduction of a different principle for counting interposable sets of a certain size would yield different results concerning whether or not crude proportionality holds. I hope that it will become clear to the reader in the course of the following how this different counting principle could

be used in determining whether crude proportionality
obtains.

In view of the fact that the 'greater than' relation,
and indeed the other crude proportionality relations,
are to be examined by reference to numbers of
interposable sets where all sets of a certain size are
to count as one set phrases such as 'most interposable
sets are larger/smaller than the interposable set **X'**
are to be interpreted as justified by the fact that
most sizes of interposable set are larger/smaller than
X.

With all this in mind let us call the number of
members of a set the cardinal number of that set. Now
according to the criteria advanced earlier for the
truth of propositions in which the phrase 'a lot'
occurs, it is true that a lot of blameworthinesses in
the maximal set are interposable between **b75** and **b10** in
sense 2 iff most sets of blameworthinesses interposable
between any given pair of different blameworthinesses
from the maximal set (represented by **b** and a number in
each case and counted as distinct in virtue of size
alone) are smaller than the set of blameworthinesses
interposable between **b75** and **b10**. But a set **S1** is
smaller than some other **S2** where the cardinal number of
S1 is smaller than that of **S2**. Therefore the
proposition ' a lot of blameworthinesses in the maximal
set of blameworthinesses are interposable between **b75**

and **b10'** is true iff the cardinal number possessed by
most sets of blameworthinesses interposable between
pairs of blameworthinesses, where such sets are
individuated for our purposes by reference to their
cardinal number, is smaller than that of the set of
blameworthinesses interposable between **b75** and **b10**. If
we say that two sets are cardinally distinct where they
have different cardinal numbers then the proposition is
true iff the cardinal number of most of the number of
(cardinally distinct) sets of blameworthinesses
interposable between blameworthinesses from the maximal
set in question is smaller than 64.

Now we have seen that although we cannot precisely
delimit the range of percentages of **x**'s which are **y**
which makes the proposition 'most **x**'s are **y**' true,
false or neither true nor false, there are certain
clear cases, that is, cases of percentages which make
the proposition clearly true or false or neither true
nor false. Let us assume that if 80% of **x**'s are **y** then
it is clearly true that most **x**'s are **y**. But then if 80%
of the (cardinally distinct) sets of blameworthinesses
interposable between pairs of different
blameworthinesses have a cardinal number smaller than
that of the set interposable between **b75** and **b10** then
most (cardinally distinct) sets interposable between
pairs of blameworthinesses are smaller than the set
interposable between **b75** and **b10**. Hence if 80% of the

interposable sets between blameworthinesses have a cardinal number less than 64 most (cardinally distinct) interposable sets are smaller than the set between **b75** and **b10** and so it is true that **b75** is much worse/much greater than **b10**.(Henceforth the specification that talk of most sets is of cardinally distinct ones will be omitted.)

Let us suppose, as before, that the (restricted) maximal set of blameworthinesses has a cardinal number **n**, i.e. it has **n** members. In this case the largest possible interposable set (as we have been calling the set of all blameworthinesses interposable between a pair of different blameworthinesses) will be the set interposable between blameworthinesses **bn** and **b1**, which will have the cardinal number **n-2**. The smallest interposable set will be the null set, which of course will have the cardinal number zero, since it has zero members; this set might be obtained, for example, from taking the interposable set for the blameworthinesses **b25** and **b26**.

Now the interposable sets of blameworthinesses can be listed. However, we are only interested in interposable sets distinct in virtue of their size, i.e. their cardinal number. Let us list the interposable sets in a specific order, starting with the null set, which will occupy the first place or position 1 and finish with the set with cardinal number **n-2**, at position number

n-1. Hence there will be **n-1** different-sized

interposable sets, which we may see simply by referring

to the position number of the last interposable set in

the ordering, which is **n-1**. (It must be borne in mind

that these sets are in a sense representative. For

example, the set with cardinal number 2 which would

occupy position 3 in the ranking is not the

interposable set for **b4** and **b7**, nor that for **b36** and

b39 or **bx** and **bx+3**. In view of this it might be best to

think of them as the representative sets for

interposable sets with certain cardinal numbers.)

 Now it is being proposed that a sufficient but not

necessary condition of **b75** being much worse/much

greater than **b10** is that 80% of the interposable sets

have a cardinal number less than 64. In other words **b75**

is much worse/much greater than **b10** if 80% of the

interposable sets have a cardinal number less than 64.

But this means that **b75** is much greater than **b10** if 80%

of the **n-1** interposable sets have a cardinal number

less than 64. Now we know how many sets have a cardinal

number less than 64. We simply refer to the line number

at which the interposable set with cardinal number 63

appears and because of the way we have ordered the sets

this line number tells us how many sets have cardinal

number less than 64. But the line number of a set with

cardinal number **x** is just **x+1** (basically because

position number 1 is occupied by the set with cardinal

number 0.) Hence there are 64 different-sized sets with cardinal number less than 64. If 80% of the **n-1** interposable sets have a cardinal number less than 64 80% of **n-1** must equal the number of those sets less than 64, that is, must equal 64. But then $80(n-1)/100 = 64$, so $n-1 = 6400/80 = 80$ and $n = 81$. So the number of different-sized sets interposable between pairs of blameworthinesses is n-1, i.e. 80. Also n is the number of members of the maximal set of blameworthinesses and so if **b75** is much worse than **b10** on the grounds that 80% of the interposable sets are smaller than the interposable set for **b75** and **b10** then the number of blameworthinesses in the maximal set is 81. In other words where we have a range of blameworthinesses which is maximal, from **b1** to **b81**, it is true that **b75** in this range is much worse than **b10**.

Of course even if we have ascertained the numerical value of **n** it may turn out that the percentage of sets smaller than the interposable set for some pair of blameworthinesses under consideration is not a clear case, that is, the percentage will neither verify that nor falsify the claim that the relation of being much greater than holds between these two blameworthinesses.

A further problem concerns the numerical value of **n**. In the absence of a procedure for determining that a set is maximal it is always possible that we have incorrectly determined the value for **n**. But where we

have the wrong value for **n** we will be unable to

calculate the correct percentage of interposable sets

smaller than the interposable set for some pair of

blameworthinesses under consideration. Clearly the

numerical value ascertained for **n** will affect the truth

value of claims about whether one blameworthiness is

much worse than another. For suppose we wish to

determine whether the blameworthiness represented by **bx**

is much worse/much greater than **by**. As we have seen

this will be true where most interposable sets for

blameworthinesses are smaller than the set interposable

between **bx** and **by**. Now the cardinal number of the

interposable set between **bx** and **by**, where **x** > **y**, will

be **x-y-1**. But as we have noted this set with cardinal

number **x-y-1** will appear in our ordering of

(representative) interposable sets by cardinal number

at line number **x-y-1+1**, or, **x-y**. Now suppose that there

are **n** blameworthinesses in the maximal set. Then since

there must be **x-y-1** sets before the set at position

line number **x-y** then **x-y-1** sets out of **n** represents the

number of sets smaller than the interposable set

between **bx** and **by**. But where **bx** is much worse than **by**

x-y-1 sets out of **n** must represent most of the sets in

n. In other words the percentage by reference to which

the truth of the claim that **bx** is much worse than **by** is

to be determined is **100(x-y-1)/n**. In this way it can be

seen that the failure to ascertain the correct value

for the denominator **n** will undermine the procedure for determining the truth of a claim to the effect that some blameworthiness **bx** is much worse than another **by**.

Moreover the problem of ascertaining the size of a maximal set will affect our assessment of claims to the effect that one distressing action is much greater in respect of the distress it causes a target individual than another, since our assessment of such claims would proceed entirely analogously to the assessment of claims about some blameworthiness being much greater than some other. It will be true that distressing action **dx** is much greater than distressing action **dy** (i.e. **dx** is much more distressing than **dy**) iff **x-y-1** sets out of **m**, where **m** = the cardinal number of the maximal set of distressing actions, represents most of the sets of **m**. So, entirely analogously, the truth of '**dx** is much greater than **dy**' is to be determined by reference to the percentage **100(x-y-1)/m**. So again it can be seen that failure to determine the correct numerical value for **m** will undermine the assessment of the truth value of claims concerning distressing actions and the 'much greater than' relation.

Let us hold over these problems concerning the unobtainability of truth-values for certain claims in which the phrase 'a lot' figures, and the ascertaining of the correct numerical value for **n** or **m**. First I want to look at how one might justify claims to the effect

that either of the other two maxims of crude proportionality concerning pairs of blameworthinesses and distressing actions obtains.

How might we justify the claim that **bx** is a little worse/little greater than **by** (or that **dx** is a little worse/little greater than **dy**?) Let us call the set of blameworthinesses interposable between the blameworthinesses represented by **bx** and **by** Sxy. In terms of this I suggest that **bx** is a little worse than **by** (where we substitute different numerical expressions for **x** and **y** but not numerical expressions for the same numbers,) iff most interposable sets of blameworthinesses between blameworthinesses of the maximal set thereof, hereafter just interposable sets, are larger than **Sxy**.

We determine whether most interposable sets are larger than **Sxy** as follows: **Sxy** has **x-y-1** members and therefore has a cardinal number of **x-y-1**. The set with this cardinal number appears at position **x-y** in the ordering. But because of the way we order the interposable sets by size we know that **(n-1)-(x-y)** sets have cardinal numbers greater than **x-y-1** and hence **(n-1)-(x-y)** sets are greater than **Sxy**. Thus we need to know whether **(n-1)-(x-y)** sets out of **n-1** constitutes most of the **n-1** interposable sets; hence we need to determine what percentage of the sets this is. But this means we have to refer to the percentage

$100[(n-1)-(x-y)]/n-1$. Were this percentage 80, for example, it would be the case that most interposable sets are larger than **Sxy** and hence that **bx** was a little larger than **by**.

Finally it might be proposed that **bx** is quite a lot/considerably/a fair bit greater or worse than **by** iff (i) **bx** is worse/greater than **by** and (ii) it is neither the case that **bx** is much worse/greater than **by**, nor is it the case that **bx** is a little worse/greater than **by**. But this means that **bx** is quite a lot worse than **by** iff (i) **bx** is worse than **by** and (ii) it is not the case that most interposable sets are smaller than **Sxy** and it is not the case that most interposable sets are larger than **Sxy**. But this choice of conditions for the truth of the proposition appears to pose a difficulty. It might be claimed that where either of the propositions 'most interposable sets are smaller than **Sxy**' or 'most interposable sets are larger than **Sxy**' has no truth value and **bx** is worse than **by** the conditions are met for the truth of the proposition '**bx** is quite a lot worse than **by**' and hence we are forced to conclude that **bx** is quite a lot worse than **by**.

The difficulty might be put as follows: suppose the proposition 'most interposable sets are larger than **Sxy**' fails of a truth value where 70% of the interposable sets are larger than **Sxy**. Let us suppose that at the point of 70% it is not clear that the

proposition falls into the camp of the truth value

'true' nor clear that it falls in the camp of the truth

value 'neither true nor false'. (It might be objected

here that if the proposition isn't clearly true or

clearly false it must be neither true nor false, that

is, it must fall into the camp of the third truth value

suggested 'neither true nor false'. But although it may

in fact be true that in the case under discussion '70%

of **x** are **y**' is a situation which substantiates the

claim that most **x** are **y** is neither true nor false if

this is meant to show that the proposition 'most **x** are

y' in fact has a truth value in the third camp the

argument is invalid. We cannot infer from the fact that

a claim is neither clearly true nor clearly false that

it is clearly neither true nor false or in the camp of

our third truth value. This is because it is consistent

with the claim being neither clearly true nor clearly

false that it is not clear that the third truth value

applies either. The point is that an ambiguity in

'neither true nor false' is being exploited. If this

phrase means 'is neither clearly true nor clearly false'

then of course the argument is valid. But it is not

this interpretation of the phrase which is being used

in the conclusion of the argument. Rather it is that

'neither true nor false' means 'clearly neither true nor

false.' This latter may be doubted of course since it is

logically possible that in the case of 70% of **x** being **y**

that we are in one of the grey areas between truth

values.) According to an earlier proposal the absence

of a clear truth value means that there is no truth

value for the proposition '70% of the interposable sets

are larger than **Sxy**' under these circumstances. But it

is also true that under these circumstances, i.e. in

which 70% of the interposable sets are larger than **Sxy**,

something slightly less than 30% of the interposable

sets will be smaller than **Sxy**. But in that case the

proposition 'most interposable sets are smaller than

Sxy' is clearly false and hence we may infer that it is

not the case that most interposable sets are smaller

than **Sxy**. But then if we may infer from the fact that

the proposition 'most interposable sets are larger than

Sxy' has no truth value that it is not the case that

most interposable sets are larger than **Sxy** (call this

the factual inference) condition (ii) has been met.

Hence where **bx** is worse than **by** and 'most interposable

sets are larger than **Sxy**' has no truth value then

provided that the factual inference is sound it would

seem that we are forced to accept, on the strength of

the criteria proposed, that **bx** is quite a lot worse

than **by**. Moreover, we are forced to accept this in just

the same sort of way if the proposition ' most

interposable sets are smaller than **Sxy**' is the one

which has no truth value. If we suppose again that 70%

of the interposable sets are smaller than **Sxy** it is

false that most interposable sets are larger than **Sxy** and so it is not the case that most interposable sets are larger than **Sxy**. So once again if the factual inference is sound condition (ii) is met where the proposition 'most interposable sets are smaller than **Sxy**' fails of a truth value.

The reason it ought not to be the case that condition (ii) is met either by a case in which **bx** is worse than **by** and the proposition 'most interposable sets are smaller than **Sxy**' has no truth value, or a case in which **bx** is worse than **by** and the proposition 'most interposable sets are larger than **Sxy**' has no truth value is that one would expect the condition to be met where both propositions are false, not where one or the other has no truth value.

So let us look at the factual inference more closely. We are supposed to infer from the proposition "'most interposable sets are smaller/larger than **Sxy**' has no truth value" the proposition 'it is not the case that most interposable sets are smaller/larger than **Sxy**.'

Let us designate the proposition 'most interposable sets are smaller/larger than **Sxy**' by the letter **p**. So we need to consider whether we may infer from '"p" has no truth value' that 'it is not the case that p'. It is often claimed that 'snow is white ' is true iff snow is white. But then surely if we accept this 'snow is white' is true iff it is the case that snow is white.

Hence we may schematize as 'p' is true iff it is the
case that p. But then if p is the case then 'p' is
true. Hence we have the contrapositive in the form of
the conditional 'if it is not the case that 'p' is true
then it is not the case that p (is the case.)' But now
if 'p' has no truth value then it is not the case that
'p' is true, and so if 'p' has no truth value we may
infer that it is not the case that p. So the factual
inference appears to be a sound one.

  Now on the assumption that if the proposition 'most
interposable sets are smaller than Sxy' has no truth
value the proposition 'most interposable sets are
larger than Sxy' is false and vice versa then where
either of the propositions has no truth value it is
neither the case that most interposable sets are larger
than Sxy nor that most interposable sets are smaller
than Sxy. But this means that in such a case, because
it will also be the case that bx is worse than by, bx
will be quite a lot worse than by according to our
criteria. I think we have a choice in this situation.
We may either accept that where one of the constituent
propositions of condition (ii) has no truth value then
the claim using the phrase 'a fair bit greater than'
will be true. Perhaps it is reasonable to affirm that
 bx is a fair bit worse or greater than by where, say,
it is not clearly true that it is a lot greater and yet
it is clearly more than a little greater.

Alternatively, we may rule out cases in which substantiating claims for relations have no truth value by rephrasing the maxims in terms not of what is the case but rather in terms of the truth value itself of certain claims. We may stipulate the conditions under which **bx** is quite a lot worse than **by** in terms of the truth values of the propositions 'most interposable sets are smaller than **Sxy**' and 'most interposable sets are larger than **Sxy**' themselves. Also a further modification of the maxims suggests itself. So far the maxims have been set out as conditionals to the effect that if such-and-such a relation holds between two objects it holds between two corresponding objects and the converse of this. However, it will be more convenient for our purposes to avoid certain difficulties concerning the truth conditions for conditionals and with this end in view we shall re-express the maxims as requiring that one and only one of the maxims, where these will be conjunctional in form, be true of some pair of pairs of objects. In effect, the maxims of crude proportionality will be satisfied or, as we will say, two sets of rankings will be crudely proportional, where for each pair of corresponding pairs from the two rankings, one and only one of the maxims will hold true of it. Roughly speaking this means that if we have a pair of blameworthinesses **bx** and **by,** and their corresponding

distress-causing actions **dx** and **dy**, where such
representations are drawn from the appropriate maximal
sets of blameworthinesses and distresses respectively,
if the two sets of objects are to stand in the relation
of crude proportionality it must be true that **(bx,by)**
and **(dx,dy)** stand in one of the relations specified by
the maxims. This means that where **x** and **y** take
different numerical value substitutions and **x>y** either
bx is much greater than **by** and **dx** is much greater than
dy, or **bx** is a little greater than **by** and **dx** is a
little greater than **dy** or **bx** is a fair bit greater than
by and **dx** is a fair bit greater than **dy**. In this way it
may be seen that the maxims are going to require one of
three conjunctions to be true for each pair of pairs
from the maximal sets under consideration. Moreover it
may be anticipated that rankings of objects which are
crudely proportional are such that, with certain
exceptions to be stipulated, for each pair of pairs of
objects like **(bx,by)** and **(dx,dy)** one of the
conjunctional maxims will be true of it.

 With the preceding observations concerning a relation
which holds within a certain range incorporated into
our maxims we have for the particular case of crude
proportionality between blameworthiness and distresses
the following:

 where **m** and **n** are cardinal numbers of the maximal sets
of distressing actions and blameworthinesses

respectively, and the numerical values substituted for x and y are non-identical, x>y, and x and y are greater than zero and less than or equal to the smaller of m and n, then for any pair of pairs of values from the two sets of the form (bx,by) and (dx,dy) either

(i) 100(x-y-1)/n-1 is a percentage which clearly verifies the proposition 'most interposable sets of blameworthinesses are smaller than the interposable set for bx and by', call this latter the interposable set bx/by, and 100(x-y-1)/m-1 is a percentage which clearly verifies the proposition 'most interposable sets of distressing actions are smaller than the interposable set dx/dy, or,

(ii) 100[(n-1)-(x-y)]/n-1 is a percentage which clearly verifies the proposition 'most interposable sets of blameworthinesses are larger than the interposable set bx/by' and 100[(m-1)-(x-y)]/m-1 is a percentage which clearly verifies the proposition 'most interposable sets of distressing actions are larger than the interposable set dx/dy', or,

(iii) 1. 100(x-y-1)/n-1 and 100(x-y-1)/m-1 are percentages which clearly substantiate the claim that the proposition 'most interposable sets of blameworthinesses and distressing actions are smaller than the interposable sets bx/by and dx/dy respectively', is either false or neither true nor false and 2. 100[(n-1)-(x-y)]/n-1 and

$100[(m-1)-(x-y)]/m-1$ are percentages which clearly
substantiate the claim that the proposition 'most
interposable sets of blameworthinesses and distressing
actions are larger than the interposable sets **bx/by** and
dx/dy respectively' is either false or neither true nor
false.

(iv) Difficulties with the maxims of crude
proportionality.

Two problems raised earlier must now be faced: the
problem posed by cases in which the propositions
relevant to the substantiation of claims about crude
proportionality take no determinate truth value and the
difficulties involved in determining the size of a
maximal set.

Now suppose **bx** and **by** are such that $100(x-y-1)/n-1$
yields a percentage such that the proposition in maxim
(i) takes no truth value. Clearly if $100(x-y-1)/n-1$
yields a percentage of 50 or less it makes the
proposition false so we must assume it takes a value
>50 in order to yield no truth value. But then
obviously that percentage won't verify the proposition
in maxim (ii) either, viz. that most interposable sets
are larger than the set **bx/by**. If more than 50% of
interposable sets of blameworthinesses are smaller than
the set **bx/by** less than 50% of such sets are larger and

this certainly won't verify the proposition that most of such sets are larger. Lastly, since, ex hypothesi, one of the propositions to be substantiated in (iii)1 is precisely the proposition which takes no truth value in maxim (i), viz. that most interposable sets of blameworthinesses are smaller than the interposable set **bx/by**, (iii)1 doesn't obtain either. So clearly the possibility of cases in which certain propositions substantiating the maxims take no truth value raises the problem of cases in which none of the maxims is true of the pairs of values under consideration. And the reason that this is a problem for us is that it is inevitable that, even in the case of maximal sets which satisfy the maxims whenever determinate truth values for related justifying propositions are available, there will be cases where the maxims are not met.

The second difficulty concerning the size of a maximal set, stems, it will be recalled, from the fact that we cannot obtain deductive certainty that we have the correct numerical value for **n** or **m**. The reason for this is that it is always logically possible that for two blameworthinesses in a ranking, for example, **bx** and **b(x+1)**, a further blameworthiness **by** may be found such that in respect of degree of magnitude of blameworthiness it is the case that **bx<by<b(x+1)**.

It will be remembered that our maxims of crude proportionality are to be conceived of as having

application within a certain range of values as possible substitution instances for **x** and **y** in blameworthiness and distressing action representations of the form **bx,by,dx,dy** etc.. The effect of this restriction is that if we substitute certain numerical values for **x** and **y** such that either the suffix of either **b** or **d** representation is greater than the size, i.e. cardinal number, of the smaller of the two maximal sets under consideration the failure of any of the maxims to be true for such values does not show that the relationship of crude proportionality does not obtain between the maximal sets in question.

This restriction may not seem very important in cases in which the maximal sets are very different in size for then crude proportionality will not obtain with or without such a restriction. To see this imagine a case in which $n=2m$. But then for anything like the size of maximal set of distressing actions **dm** will be much greater than **d1** (i.e. the first action will be considered much more distressing than the second.) This is because most interposable sets of distressing actions are going to be smaller than the interposable set **d1/dm**. Ex hypothesi there will be no set larger than this set. So, apart from this set, all other interposable sets of distressing actions will be smaller, that is, out of **m-1** possible different sized interposable sets of distressing actions **m-2** of them

will be smaller than the interposable set in question. This means that the percentage of sets of distressing actions smaller than the interposable set **d1/dm** will be given by the fraction **100(m-2)/(m-1)**. But where **m > 4** this fraction will yield a percentage of 75 or greater. Hence where **m > 4** we have, I think, a clear case in which most interposable sets of distressing actions are smaller than the interposable set **d1/dm**. But clearly our maximal sets will have more than four members so **dm** will be much greater than **d1**.

However although for any size of set where the cardinal number of that set is greater than four the largest member of that set will be much larger than the smallest member, (because 75% or more of the interposable sets formed from members of the set in question will be smaller than the interposable set formed from the smallest and largest members thereof,) it is not the case that where **n=2m bm** is going to be much larger than **b1**, even if the set of blameworthinesses represented by these **b** values is of considerable size. Ex hypothesi **bn** is the greatest **b** value and **bn = b2m**. But then the largest interposable set of blameworthinesses is of cardinal number **2m-2**. But if, as before, we number the interposable sets of blameworthinesses by the numbers 1,2,3...in ascending order of size so that the smallest interposable set is at position 1 with size 0 (that is with 0 members) then

the largest set is at position 2m-2+1 i.e. at 2m-1.
Hence there are 2m-1 different sized interposable sets.
Now the interposable set **b1/bm** has cardinal number **m-2**
and since sets of this size appear at position **m-1** in
our ordering there are **m-2** interposable sets of
blameworthinesses smaller than the (representative) set
at this position. Hence there are **m-2** sets smaller than
the interposable set **b1/bm**. But then the percentage of
interposable sets of blameworthinesses smaller than the
interposable set **b1/bm** is given by the fraction 100(m-
2)/(2m-1). But (m-2)/(2m-1) = (m-2)/2(m-½) so it can be
seen that the percentage will be the product of 100 and
a fraction the denominator of which will always be more
than twice the numerator. (This is because the
denominator is twice a certain figure, call it **x**, where
x itself is larger than the numerator.) But in that
case the percentage of interposable sets smaller than
the interposable set **b1/bm** will be the product of 100
and a fraction which will be smaller than a half (which
it must be if the denominator is always more than twice
the numerator.) Hence the percentage of interposable
sets of blameworthinesses smaller than the interposable
set **b1/bm** will always be less than 50. But then it
cannot be the case where n=2m that most interposable sets of
blameworthinesses are smaller than the interposable set
b1/bm and so it can't be the case that **bm** is much
greater than **b1**. So crude proportionality cannot hold

between the pairs **b1,bm** and **d1,dm** where **n=2m**. But even
if we had our restriction on taking numerical values
for **x** and **y** in the representations **bx, by, dx,** and **dy**
it would not affect this result (i.e. that crude
proportionality does not obtain) because we have only
considered numerical value substitutions for **x** and **y** up
to the value **m** and **m** is the size of the smaller of the
two maximal sets.

The upshot of all this is that we can never be
absolutely certain that crude proportionality holds
between two maximal sets of objects since we cannot be
certain of their respective sizes. However the fact
that for certain pairs of values from both sets, in
this context **b** and **d** values, the proposition that most
interposable sets are smaller/larger than the
interposable set of the two **b** values or **d** values will
have no truth value is perhaps not as serious a
difficulty as it might seem. I suggest that we accept
that it has been shown that crude proportionality
obtains between pairs of sets (ignoring now the
reservation expressed earlier concerning the estimation
of maximal set size,) where for any pair of pairs of
values from each set (or what is believed to be the
maximal set,) one of the maxims of crude
proportionality is true, with the exception of cases in
which justificatory claims of the kind discussed above
take no determinate truth value. Thus we shall ignore

cases in which values are chosen such that we face the difficulty of justificatory claims for the maxims which take no determinate truth value. I shall not try to offer much of a defence for this beyond the following suggestion: for any pairs of values, call them **bx** and **by** though they need not be blameworthinesses, where corroborating propositions have no truth value it is still the case that both **bx** and **by** can individually be shown to stand or not to stand in certain relations vis-a-vis other **b** values and so it can be shown that these other relations are duplicated by their counterparts in the other maximal set. So it would be possible, for example, to show that **bx** stood in a certain relation to some **b** value **bz** and that this relation was or was not duplicated by the corresponding **d** values **dx** and **dz**. In other words we can show that proportionality obtains or does not for the pairs of maximal sets in question by showing that their members stand or do not stand in a web of relations the confirmation or disconfirmation of the individual strands of which can be made for the most part without entering the problematic area of truth value lacunae, as we might call them.

Nor should we suppose that the device of restricting the range of values to be considered so that no values larger than the lesser of **n** or **m** be taken is without use. Clearly the utility of such a restriction is to be

found in cases unlike the one above in which n=2m. It is likely that maximal sets of blameworthinesses and distressing actions will be very large, at the least. But we should not, I submit, wish to rule that crude proportionality does not obtain in the case of maximal sets paired members of which for the most part stand in the appropriate relations to their identically numbered counterparts in the opposite set (with the exception of cases in which truth value lacunae arise) simply on the grounds that no identically numbered counterparts in the opposite set are available in some cases. If two maximal sets have a million and a million and one members respectively, and the criteria of crude proportionality are met except for the case of **b1000001**, which of course has no **d** counterpart, we should not want this quibble to result in the sets being, as far as our criteria are concerned, not crudely proportional.

(v) Crude proportionality and the conditions of retributive equivalence.

It is important to note that in a ranking of objects from a maximal set we use representations of the objects with numerical suffices identical to the numerical expressions which indicate the relative ranking of the object in question where ranking

positions are expressed by a number sequence beginning with 1 and in which each successive ranking position is designated by that number which follows that of the representation of the object immediately preceding it in the ranking. Hence we get a numerical sequence for the relative rankings of objects governed by the successor function. In this way we avoid certain problems. Firstly we avoid having cases in which two maximal sets are not crudely proportional according to our criteria just because the sequences which indicate the relative rankings of the members of the different maximal sets start at different numbers. For example, we might start the ranking of distressing actions at 1 and the ranking of blameworthinesses at 500 so that the first member of the latter set would be represented by the expression b500. Secondly, by using a sequence ordered by the successor function we avoid the circumstance in which pairs of maximal sets are not crudely proportional according to the criteria suggested owing to the fact that the number suffixes of the representations of the members of the different sets are not 'aligned'. The following would be an example of this lack of alignment between two maximal sets:

d1   i2

d3   i4

```
d5                    i6
d(2x-1)               i(2x)
```

Now **ix**, where this designates an immoral action,(or
bx where this designates a blameworthiness,) is MCR
equivalent to **dy** iff the numerical value substituted
for **x** is identical to the numerical value substituted
for **y**. But clearly the fact that some **ix** or **bx** is MCR
equivalent to some **dy** does not entail that the maximal
sets from which these representations or expressions
are drawn are crudely proportional. For example, in the
case already discussed of maximal sets of
blameworthinesses and distressing actions, where the
sets have cardinal numbers of **n** and **m** respectively and
n=2m, although the object represented by **bm** is MCR
equivalent to that represented by **dm** the two sets have
been shown not to meet the criteria for crude
proportionality.

I shall now stipulate under what circumstances we
shall regard two objects, either a blameworthiness or
an immoral action on the one hand, and an action
causing distress on the other, as retributively
equivalent, that is, under what circumstances a
necessary condition of the latter counting as a
retribution for the former has been met. Also I shall
try to show how this bears on the original example
which prompted this long excursus into the nature of

proportionality. It will be recalled that it was suggested that we might have two scales of objects properly represented such that although all corresponding pairs were MCR equivalent (this is necessarily true anyway,) the two scales might still be unacceptable for the purpose of setting retributive penalties since a certain 'intuitive fit' might not obtain between individual objects in the corresponding pairs. No analysis was given of what this intuitive fit required but I shall now propose one interpretation of the requirement of 'intuitive fit', viz. it is not enough in setting retributive penalties for blameworthinesses or wrongs that we adhere to the requirement of SP that the greater the wrong or blameworthiness the greater the penalty; something more is required and this something is that if one wrong/blameworthiness is worse, to a certain extent, than some other the penalty for this object should be more distressing to that same extent than the penalty for the other. The 'intuitive fit' requirement is to be interpreted, then, as the requirement that the two scales of objects, blameworthinesses or immoral actions on the one hand, distresses on the other, should be crudely proportional in the precise sense that they meet the maxims of crude proportionality.

Let us say that retributive equivalence holds between any action causing a certain distress designated by **dx**

and ranked in respect of the distress it causes to its target and any immoral action or blameworthiness designated by the expression **ix** or **bx** respectively and ranked in respect of how wrong or morally grave it is iff 1. **dx** and **ix** or **bx** are rankings within the maximal sets of distresses and immoral actions or blameworthinesses respectively, 2. **dx** and **ix** or **bx** are MCR equivalent and 3. the maximal sets from which the expressions **dx** and **ix** or **bx** are drawn are crudely proportional.

Now what exactly does the third condition for retributive equivalence add? Its significance is simply this - if **dx** and **ix** or **bx** are MCR equivalent then it may nonetheless be the case that some other distress, call it **dy**, from the maximal set is much worse than any distress **dx** and yet the blameworthiness or immoral action **by** or **iy** which is MCR equivalent to **dy** is not much worse than **bx** or **ix**. It is for this reason that occupation of the same position in a maximal set is not a sufficient condition of retributive equivalence. This is part of the reason why the earlier example which I shall repeat below was considered unsatisfactory. We

have

i1   r1

i2   r2

i3   r3

i4   **r4**

i5

i6

It will be noted that we use **r** rather than **d**
representations here. This is of no importance: **r** terms
designate retributions only in the weak sense of
actions causing distress to certain targets. Clearly if
retributive equivalence is a necessary condition of
retribution for something then ex hypothesi the **r** terms
in the example do not in fact designate retributions
proper at all. It may be remembered **r1** is a fine of
£10, **r4** is hanging, **i1** is a first parking offence, **i4** a
fourth parking offence. Clearly **r4** and **i4** are not MCR
equivalent and so retributive equivalence cannot hold
between them. However the original objection to the
claim that **r4** and **i4** are retributively equivalent was
that there was a lack of proportionality in some sense
between them. One way of giving evidence for this lack
of proportionality would be as follows: while **r4** is
very much greater than **r1** (in respect of distress to
its retributive target of course) it is not the case

that i4 is very much greater than i1. However we cannot

say on the strength of this that the pairs r1,r4 and

i1,i4 aren't crudely proportional on the grounds that

they don't satisfy the requirements of crude

proportionality, that is, on the grounds that they

don't make one of the maxims true and fail to

constitute one of the exceptions to the applicability

of the maxims. To begin with our maxims are only for

what I have dubbed 'crude' proportionality which does

not take into account the relation 'very much greater

than' (though a way has been suggested by which such a

relation could be incorporated.) But more obviously the

pairs are not designated by expressions drawn from

their placements in maximal sets. Yet I submit that our

criteria of crude proportionality do capture part of

the objection made earlier to the example. For like the

earlier objection our objection, involving reference to

the criteria of crude proportionality, is based on the

claim that certain relationships between each term of

the pairs (r1,r4, and i1,i4) and the other members of

the same set, that is, here, the set of retributions

for an r term, blameworthinesses or immoral actions for

i or b terms, where such relations take the forms 'is

much greater than' 'is considerably less than' etc., do

not obtain. In this way the essence of the earlier

objection to the supposed equivalence in the example

along the lines that it is not proportional or that

there is no 'intuitive fit' between **r4** and **i4** is captured by the concept of crude proportionality as it has been elucidated.

Moreover the example serves to show how in theory it can be determined whether a distressing action imposed for an immoral action or blameworthiness is appropriate in so far as it satisfies the condition of retributive equivalence. Firstly, we have to re-express terms like **i4** and **r4** so that they are the appropriate expressions given their placement in rankings of objects from maximal sets in respect of certain properties, namely, in respect of amount of distress, degree of wrong, and gravity of blameworthiness. Once this has been done it can be ascertained whether the new expressions are such that MCR equivalence holds between the objects designated thereby. If it does not hold we need look no further, retributive equivalence obviously does not hold. If it does hold we have to go on to determine whether the maximal sets from which the expressions subsequent to ranking are drawn, are crudely proportional. If we were to imagine that when re-expressed in the form of their maximal set expressions **r4** and **i4** above were MCR equivalent it could still be denied that they were retributively equivalent because the maximal sets would surely not be crudely proportional under such circumstances. This again shows that our conception of crude proportionality may

provide a fundamental criticism of the claim that retributive equivalence holds between **r4** and **i4** in the example and vindicates our explication of this concept as a way of making sense of the earlier objection that **i4** and **r4** were not proportional in some sense.

Now it will be recalled that our initial aim was to find an equivalence of some kind between two properties of things, how wrong an action was and how distressing was its response. By analogy we will extend this to include a relation between the properties of how bad some blameworthiness is or was and how distressing is its response action. I suggest that in view of the way in which we have defined retributive equivalence, viz. as a relation holding between objects ranked in respect of wrongness, badness and distress, there is a perceivable sense in which the equivalence between the objects which are retributively equivalent is between those properties. However, in order to be rigorous, it must be shown that where, say, **bx** and **dx** are retributively equivalent the properties of degree of moral gravity and degree of unpleasantness which characterize **bx** and **dx** respectively are themselves equivalent in some sense. It will be recalled that an earlier suggestion was that the properties be equivalent in respect of rankability. Yet in the same way that it has been seen that equivalence of actions or actions and blameworthinesses in respect of

rankability does not yield a useful equivalence, and
that even MCR equivalence between such things is
deficient as it may lack 'intuitive fit', it may be
appreciated that equivalence between the appropriate
properties in respect of rankability alone would be
unsatisfactory. This suggests that what we are looking
for is some equivalence between properties, MCR
equivalence, where the objects which are MCR equivalent
take their rankings from sets which are crudely
proportional.

Now it turns out that where **bx** and **dx** are
retributively equivalent we can point to a relation
which holds between the properties of badness and
distressingness which characterize these two objects.
But care is required. Consider the set from which **bx** is
taken, the maximal set of blameworthinesses. Now it is
tempting to think that each object in the restricted
maximal set has a property which would go into the
restricted maximal set of the properties of badness to
a certain degree of things. But it cannot be assumed
that such a set would be maximal in respect of having
for content of properties each shade of moral badness.
The reason for this is that there may be things which
take moral hues which no blameworthiness takes. Put
another way I am saying that it may be that there is
some property of being bad to a certain degree where
this degree could not characterize a blameworthiness.

But then this degree would be absent from the maximal set of properties of degree of moral badness where such a set was constructed simply from the properties of badness attributable to the members of the maximal set of blameworthinesses. In view of this we must be careful to note that if we take each property of specific degree of badness from the members of the maximal set of blameworthinesses we will be able to construct the maximal set of properties of moral badness in respect of degree of badness of content where such properties apply to blameworthinesses. Likewise if we take each property of causing a certain amount of distress from the members of the maximal set of distressing actions we get another maximal set of properties of being distressing to a specified degree in respect of content of being distressing to a certain degree where such properties apply to actions. Now if we call these two new maximal sets B and D respectively it can be shown that if **bx** and **dx** are retributively equivalent then the property of badness which characterizes **bx**, call it **pbx**, and the property of being distressing to a certain extent which characterizes **dx**, call it **pdx**, are MCR equivalent where their representations are taken from their rankings in B and D and B and D are crudely proportional. The converse of this can also be shown.

I shall not burden the reader with the proof of this.
How the proof goes should be fairly obvious. One point
may suffice. If we can pair each member of the maximal
set from which **bx** is drawn with a unique and different
member of **B**, and vice versa, and we may do the same for
the maximal set a member of which is **dx** and the set **D**,
then the first two sets are equinumerous and so are the
second two. But this means that the calculations
required to show that **bx** and **dx** are retributively
equivalent are duplicated where we try to show that the
analogous relation for **pbx** and **pdx** holds.

In the manner outlined it is therefore possible to
show that there is a sense in which our retributive
equivalence between objects is also a sign that an
analogous equivalence holds between certain properties
of those objects. This relation between the properties
is one of rankability, albeit of a specific kind: it is
an equivalence of rankability where the properties are
ranked in maximal sets of properties of being morally
bad or wrong and being distressing to certain degrees,
where such properties characterize, respectively,
blameworthinesses or wrongs and actions. Moreover these
maximal sets must be crudely proportional.

So an equivalence has been found between a
blameworthiness or immoral action on the one hand (or
properties thereof,) and a distress on the other (or
property thereof,) which (i) meets our earlier

criterion of SP and (ii) meets the intuitive
requirement that at least a coarse-grained
proportionality exist between certain pairings of
objects (or properties thereof) within the sets of
blameworthinesses or wrongs on the one hand and
distresses on the other, (or sets of certain properties
thereof) respectively. This relation of retributive
equivalence, as I have called it, is to be understood
as a necessary condition of a distress being
retribution for an immoral action or blameworthiness.
Thus if a distress is to be retribution for some
blameworthiness, say, both must be ranked vis-a-vis
others in maximal sets of the relevant kind. Secondly,
these maximal sets, provided that they are restricted
in the sense defined earlier, have to be crudely
proportional. (They must be restricted so that
interposable set sizes are calculated in a certain
way.) As we have seen it is in the nature of a
calculation of whether crude proportionality obtains
that we be able to ascertain the size of maximal sets
and rank their elements. Thirdly, the fact that **bx** and
dx are retributively equivalent doesn't mean that we
can perform **dx**. **Dx** might represent an action which is
essentially temporally specified, where the temporal
specification is, for example, 'in the reign of
Tutankhamen.' Lastly, retributive state punishment was
defined as state punishment which is an act of

retribution. I shall not have more to say here with regard to the definition of retribution or the proper characterization of state punishment but clearly criteria for the latter may rule out the possibility of some retributions being state punishments and vice versa. Moreover since retributive equivalence is a relation between retributions and immoralities rather than punishments and offences there is some latitude in what is to count as a system of retributive punishment. I shall say that a system of punishment is retributive provided that punishments for the range of actions considered to be offences are such that the greater the immorality of the offence the greater the punishment for it. This of course means that it is not necessary for a punishment system to make all immoralities punishable offences for it to be a retributive system. But we have delayed enough the investigation of certain retributivist justifications of retribution and it is to these that we must now turn.

Part II Retributivism

Chapter 3 Retributivist justification

(i) Retributivism as a logical doctrine.

Before we can consider what sort of a justification
of retribution might be offered by various
retributivist theses we must consider a general
objection to this strategy: it has been held that
retributivism is a logical and not a moral doctrine.

One series of arguments for this point of view may be
found in Quinton's famous paper 'On Punishment'. He
begins by considering an account of retributivism
characterized by some or all of the following features:

(i)   punishment is self-justifying

(ii)   punishment is only justified by guilt

(iii) it is necessary that a man be guilty if he is to
       be punished

(iv)   there are circumstances in which the infliction
       of suffering is a good thing in itself.

Now apparently these conditions taken together can be
dismissed immediately. If (i)-(iv) characterize the
retributive doctrine of punishment this doctrine cannot
be moral because it cannot even pass muster as a

doctrine; (i)-(iv) are inconsistent. If (i) obtains then punishment is not justified by anything other than itself. But guilt is not punishment and so if (i) is true then (ii) cannot be.

But this is too quick a dismissal. The fact that guilt and punishment are not one and the same thing does not show that (i) and (ii) are inconsistent. This is because a reasonable interpretation of (i) will allow punishment to be justified by things which are not equivalent to punishment. What (i) means is that we are not to look to the value of any consequences of punishment in the justification thereof but we may nonetheless look to the value of features of the act of punishment itself, that is, things which we might say are involved in the act. ('Involvement' may sound like too imprecise a notion to do the work here but I use it merely to draw attention to a distinction which can be grasped not really to direct one toward that distinction.) For example, the infliction of suffering is, let us suppose, a part of or involved in the act of punishment while the moral regeneration of the punishee is not. But in that case the value of infliction of suffering on the punishee might figure in the justification of punishment and yet punishment still be self-justifying. Something is self-justifying if it is to be justified by features of it which do not include its consequences.

But this does not appear to help much either, even if we were to suppose that the idea of a distinction between act and consequences is not problematic. For though not a consequence of punishment neither is guilt thought of as a part of it; rather it is thought of as a precondition. So punishment can't be self-justifying and yet it be true that if it is justified it be justified by guilt. So (i) and (ii) appear inconsistent.

Yet this cannot be Quinton's argument against retributivism as a moral doctrine since he wishes to show that retributivism is a doctrine, albeit not a moral one. It is a bizarre logical doctrine that is itself inconsistent.

One way in which to secure the consistency of (i) and (ii) would be to reinterpret (ii) as the claim that punishment of someone is only justified if the punishee is guilty. In other words we can infer from the fact that someone's punishment is justified that they are guilty. Of course as it stands (ii) doesn't amount to the claim that punishment is ever justified, rather that if it is, such cases of justified punishment involve punishment of someone guilty. Moreover (i) could be cautiously interpreted as the claim that if punishment can be justified then it is not to be justified in virtue of the value of its consequences. However the cautious interpretation of (i) is not

necessary to show that retributivism could be a moral
doctrine for we can put together (i),(ii) and (iv) to
get: an act of punishment is justified only if the
punishee is guilty and the punishment constitutes a
good thing in itself. Admittedly this is not to include
condition (iii) but even if we were to interpret (iii)
as a conceptual condition of punishment, as Quinton
intends, we still have a moral and logical doctrine,
rather than simply a logical one. On the other hand if
we interpret (iii) as a moral claim then it appears to
have the same content as (ii) and so cannot generate an
inconsistency. It could be said that this version of
retributivism is somewhat vague because the
circumstances in which punishment is a good in itself
are not specified but it is surely nonetheless a moral
claim about the justification of punishment. In fact if
we add the not unnatural assumption that it is
precisely when a person is guilty that there is
intrinsic value in their suffering we get a version of
retributivism sometimes referred to as intrinsic
retributivism, viz. that punishment is justified when
an individual is guilty because the suffering of the
guilty is a good in itself or has intrinsic value. In
conjunction with this if we take condition (iii) as the
claim that it is a necessary condition of punishment of
someone that that individual be guilty we get a version
of retributivism which claims that all punishment (of

the guilty) is justified. Of course this is implausible
for it seems unlikely that any punishment whatsoever of
a guilty individual is good in itself. Is the electric
chair for minor tax evasion a good in itself? Secondly
even if it is good in itself why aren't other goods and
evils to be taken into account to measure the force of
the justificatory claim? But the point is that it is
surely a moral claim, among other things.

So we don't appear to be forced by the content of the
claims alone to reject retributivism in this form as a
moral doctrine. What then of what we might call logical
retributivism, the claim that it is a necessary
condition of punishment that it be of the guilty (and
so taking condition (iii) above now as the sole
condition.) This means that it must be logically
impossible to punish someone who is not guilty. In the
face of certain intuitions that this is evidently false
it is clear that some strong argument will be needed
for such a view. I shall consider two arguments offered
by Quinton at this point.

The first argument draws on a supposed analogy
between locutions concerning punishing and promising.
If I say 'I promise to x' where it is not in my power
to x then I have not in fact promised to do x. One
reason for this is the familiar maxim 'ought implies
can'. If this is true then the contrapositive of the
maxim, with 'have a duty to' replacing 'ought,' yields

something like 'can't implies not having a duty to'.
But then if I say 'I promise to x' where I can't x I
have no duty or obligation to x. But then if I have no
obligation to x I can't have promised to x, since a
promise would issue in certain obligations of mine, at
least of a prima facie kind. So it is perhaps
reasonable to agree that I haven't promised to x under
the circumstances. Moreover this would argue for the
absurdity of such phrases as ' I promise to do x which
is not in my power.' But the supposition which does the
important work in the analogy is that someone else may
say correctly of me under such circumstances where it
is not in my power to do x, that I promised to x, even
though it was not in my power to x. But I think that
one must be careful if one is not to be misled by this.
If someone says of me under such circumstances 'he
promised to x' then surely for the claim to be true it
must mean something like 'he said he promised to x'
(where this is equivalent to "he used the words
'promise to x.'") Quinton's claim here is surely based
on an ambivalence concerning the conditions for
promising. If we decide that I did make a promise under
circumstances where I promise something beyond my power
then the report of another that I have made one is
true. But surely if we decide that I did not make a
promise to do x in these special circumstances then a
report that I did is false, if it is understood that

the report 'he promised to **x**' does not simply mean something about words I have used. The stringency of consistency between the two cases operates provided that we come down squarely in favour of a promise either being made or not.

Now the analogy is supposed to run like this: just as we may say truly of me in the third person 'he promised to **x**' even where it was not in my power to **x**, we may say 'he punished Smith for **y**' truly even though Smith didn't do **y**. And this supposedly accounts for cases where we seem happy to say 'Evans was hanged (=executed by hanging) for a crime he didn't commit, because we now know that Christie did it.'

But the parallel isn't convincing. To begin with, if I am unable to promise to **x** where it is not in my power to do so then the judge saying 'I am going to punish you, Evans, for such-and-such a crime' is a case in which the judge may say truly of himself 'I oughtn't to feel bad about this because I'm not actually punishing Evans'. But I am not sure that if I am advised of this fact I am happy to say that although the judge rightly thinks that he hasn't punished Evans I may say truly that the judge punished Evans or simply that Evans was punished. This just flies in the face of simple consistency, the sort of everyday consistency we are not inclined to balk.

Secondly, if the first person utterance is the one
that is invalidated by the absence of certain
background conditions, this is more serious (and hence
worth making something of) in the case of promising
than it is in that of punishing. Very few people
actually do the punishing, and so very few people can
say of themselves that they are not in fact punishing
but are doing something else. The rest of us will just
say 'the poor old devil was punished, and he didn't
actually commit the crime'. In other words the effect
of this logical constraint in the context of punishment
is marginal. In most cases, it would seem, it is quite
proper to talk of punishing people for things they
haven't done. So Quinton surely conceded too much by
this, certainly too much to make for retributivism a
significant claim about the logic of punishment.
Moreover if the logical constraint on punishment is
only partial we do not seem to gain much in the way of
obviating difficulties for consequentialist accounts
like the apparent justification of punishing the
innocent. We could not say that punishment of the
innocent was logically impossible and hence that
consequentialist justifications of punishment cannot
require the punishment of the innocent under any
circumstances.

Lastly, 'I promise to do something I cannot do' is
absurd as is 'I am punishing you for something you have

not done', but in the latter case it might not be the absence of guilt which gives rise to the absurdity, but rather assumptions about who utters the absurdity and as part of which role in which institution. In the latter case the absurdity might equally stem from the fact that if I am in a position to punish then a background of institutions of which I am a part is presupposed. If I am the judge, for example, and say 'I am punishing you for something you have not done' it is the fact that I am not convinced of your guilt, and hence as part of the judicial process under certain circumstances should not reach the conclusion that you are liable to punishment, which makes the claim absurd. Punishment is part of the judicial process and is a rule-governed procedure. This is one of the reasons why I don't simply punish someone I harm in retaliation for a wrong done me. Where I am the judge and I don't think the relevant crime was committed it may well be that the rules haven't been followed. If one thinks of punishment as defined by these institutional rules then there is a definite contradiction involved in my words as judge, a contradiction based not on guilt as a necessary condition of punishment, but on the belief of certain key individuals of that guilt as a necessary condition thereof.

The second argument purporting to show that guilt is a necessary condition of punishment incorporates the

distinction between a rule which attaches punishment to actions of a certain kind and the application of a rule to a particular case. Punishment is then thought of as the infliction of suffering attached by these rules to certain kinds of action, guilt the condition of a person to whom such a rule applies.

Now Quinton concedes that guilt will not be a logically sufficient condition of punishment. This will be so whether we understand guilt as the condition of having performed an action proscribed by a rule or as the condition of having been found by a court to have broken such a rule. In the first case one may commit an action proscribed by rule and not be caught, let alone punished; in the second, sentencing is a separate procedure from finding guilty and the latter may take place without the former, and the former may well be waived at the discretion of the court in certain cases. At best, if the rule specifies a certain punishment for the rule infraction in question guilt is a sufficient condition of liability for punishment, not punishment.

It is more likely, however, that guilt, on some interpretation, is a necessary condition of punishment. Clearly on the former interpretation of guilt as the condition of having performed an action proscribed by a rule it will not be. I may be punished for something which I did not in fact do even though the court has found me guilty. But if we specify that punishment is

the infliction of suffering on someone specified by a
rule for one who is found by the court to have broken
that rule then obviously the second interpretation of
guilt will be a necessary condition of punishment. By
building the condition into the concept of punishment
we can achieve this result. This shows that
retributivism qua the claim that guilt is a necessary
condition of the concept of punishment is indeed a true
logical claim. However, alternative interpretations of
guilt will make the (logical) claim false. For example,
it is not part of the concept of punishment that it be
of an individual who has done something morally wrong.
Moreover retributivism could just as well be thought of
as the moral claim that it is a necessary condition of
the justification of punishment that it be for a moral
wrong as of the logical (and false) claim that it is a
necessary condition of proper use of the word
'punishment' that the individual punished be one who
has done something wrong.

So even if there is a sense in which retributivism
can be thought of as a logical doctrine it does not
appear that we are forced to conclude that it is not
also, on other interpretations, a moral doctrine.

Kaufman suggests a way in which we might try to
relate the logical and moral doctrines of
retributivism. The idea is that since only some of the
cases in which we deliberately inflict suffering are

morally justified, those in which moral guilt is

present, and because of the importance of this group of

actions, we use a word for them, namely 'punishments'.

In this way if the principle of retributivism is: guilt

is a necessary condition of punishment we can interpret

this as (i) moral guilt is a precondition of justified

infliction of suffering and (ii) justified infliction

of suffering will be called 'punishment'. It is because

(i) is accepted, that is, the moral doctrine, that

acceptance of (ii) commits us to the belief that moral

guilt is a necessary condition of the concept of

punishment. Of course, as it stands, this argument

could at most show that moral guilt is a necessary

condition of the concept of punishment. However it is

dubious that the argument achieves even that. If we

should change our minds about what kinds of deliberate

infliction of suffering were justified and no longer

regarded moral guilt as a precondition of it being

justified, it seems that only our moral beliefs would

have changed, not our concept of punishment.

'Punishment' would still only be used for justified

inflictions of suffering. It is not part of the concept

that something is justified that it is for moral guilt.

If we consider the somewhat tedious example of the word

'bachelor', for example, and define it as 'unmarried

male of marriageable age' then it seems quite clear

that although all bachelors will have heads 'not being

headless' is nonetheless not part of the concept of 'bachelor'. One way of accounting for this would be to say that although anything answering to the description in the definiens will have a head, 'not being headless' is not logically entailed by any of the concepts in that definiens alone. By analogous reasoning it is not entailed by the definiens 'justified deliberate infliction of suffering' alone that moral guilt obtains in the case of anyone punished. Moreover, even if the argument were sound, it is dubious that moral or legal guilt is a precondition of the justified deliberate infliction of suffering. If an armed robber eager to humiliate his victims tells me that I must slap the cashier in the bank across the face or he will shoot them I may be justified in acquiescing to his wishes to prevent a greater evil. So the Kaufman argument does not appear to strengthen the claim that the retributivist doctrine qua logical doctrine (in the form of the claim that the concept of punishment necessarily involves moral guilt) is true.

Some have attributed something like a logical doctrine of retributivism to Mabbot. Certainly in his justly famous paper on punishment of 1939 he makes claims about the logic of certain concepts. One of these is that if a headmaster decides to have rules to govern the behaviour of pupils at his school, a decision which he would take on the basis of

utilitarian considerations, it follows necessarily that the headmaster must punish those who infringe these rules. There are at least two ways of taking this. Either he is claiming that if the headmaster does not punish someone who has broken a rule then in fact no rule exists and no rule has been broken, or, he is maintaining that where a rule is broken there is a duty to punish and the rule-breaker is liable for punishment. If we don't have punishment, according to Mabbot, what we have is not a system of rules but mere exhortation or advice. Firstly the contention itself ✓ seems implausible. For one thing games have rules and not all games have penalties or punishments. Also we might imagine a ruler of a populace known by that ruler to be sheep-like. He might set up rules knowing that no punishment system need be set up because the rules would be followed. Furthermore the populace might be told that they had to or must follow these rules. Surely the force of the 'must' here is not dependent on being able to give specific content to an 'or else'; most moral rules are without sanction for infraction. Secondly, this point that rules without penalties aren't rules, if true, would not entail that rule infraction of some kind is a necessary condition of the concept of punishment. If punishment, either in the form of liability for it or actual infliction of it, is a necessary condition of the existence of a rule, it

doesn't follow that a necessary condition of punishment
is that a rule has been broken. If true the claim is
simply that whenever we have rules we have some kind of
punishment, not that whenever we have punishment we
have some kind of rules, and specifically infractions
of those rules. This is to mistake a necessary
condition about rules for a necessary condition about
punishment. However, I am far from claiming that Mabbot
makes this mistake; it is simply that this comment of
his, though it does amount to a logical claim of some
kind, certainly will not stretch to the claim that a
necessary condition of the concept of punishment is
that it be for a rule infraction.

Of course there may be some truth to the claim that
you can't have punishments without rules. But of course
to be interesting this must be more than the logically
true claim that a legal institution of punishment
necessarily involves rules. If it is legal it must
involve legal rules. It seems plausible that punishment
must be for a rule infraction in the sense that it is
part of the concept of punishment (though not in the
family context) that it be for a rule infraction. But
then if a punishment is to be just it is necessarily
true that it be for a rule infraction. Yet this does
not entail that only one thing, a rule infraction, can
justify punishment. If retributivism be construed as
the central claim that only the breaking of a rule can

justify punishment this is logically independent of the claim of retributivism as a logical doctrine, viz. that it is part of the concept of punishment that it be for a rule infraction. For the logical claim is surely consistent with it being true that no punishment is ever justified, in which case if we interpret retributivism qua moral doctrine as the claim that legal rule infraction alone justifies punishment then the logical doctrine may be true while the moral doctrine is false.

In fact Mabbot does offer us a fully fledged moral doctrine of retributivism which depends on what is by now a familiar distinction between the justification of a practice and the justification of an action which is part of that practice. I shall not discuss this distinction however as it more properly forms part of that group of 'compromise' or 'hybrid' theories of the justification of punishment which aim at an amalgamation of elements from both retributivist and utilitarian theories.

In concluding this section I should like to say something about the familiar distinction of forward- and backward-looking as applied to the justification of punishment. It has been proposed that this distinction sits squarely on the shoulders of the retributivist-utilitarian justification distinction; while the former kind of justification is backward-looking the latter is

forward-looking. Now if we interpret retributivism as the claim that the justification of punishment is to be found in the guilt of the offender/wrongdoer alone then this would appear to be a backward-looking justification because what justifies punishment happens before the punishment takes place. But clearly this is in fact a relationship which may not hold under conceivable circumstances in which criminal or moral desert is taken to be the sole justification of punishment. For example, I may know that x is going to commit a crime at some point in the future but also know that after commission of the crime punishment of x will be impossible. In such a case I may surely justify punishment in the present on the grounds of criminal or moral desert, provided of course that punishment now does not render future commission of the crime impossible. Secondly even if such a scenario can be dismissed as fanciful it seems ill-advised to constrain retributivist justification to be backward-looking. For if what justifies punishment must be in the past any effect of punishment is excluded from its justification, nor can it be relevant to it. But this is surely bizarre. Suppose that punishment of any kind except capital punishment or life incarceration always had the effect of making the criminal return to society hell-bent on creating anarchy. Then according to retributivism the imposition of punishment on many

criminals who, let us suppose, ought not to be sentenced to death or life imprisonment, where such action might induce anarchy, would still be justified, and this simply because desert is the only consideration relevant to the justification of punishment. To dismiss the effects of punishment is to ignore the fact that punishment itself can only be justified against a background assumption that it will not lead to the worst possible mayhem. But to accept the necessity of some such set of background assumptions is surely to accept that the effects of punishment can be relevant to the matter of its justification. And this leads us to be sceptical of a solely backward-looking justification of punishment.

Thirdly, to justify a course of action in terms of its justice need not commit us to a backward-looking justification where this latter is understood to mean that the source of justificatory value is prior to the punishment. If I hand out sweets to a group of children and give them one each I may justify this as being fair, fair that is because I have brought about a certain kind of distribution under certain circumstances. The justification of my action would look to the future, that is, its consequences, and these consequences are fair. Thus it is that one of the most important modern retributivist accounts of Morris seeks a justification of punishment in a consequence of

that punishment, which consequence itself is appraised as just. Indeed many of the recent retributivisms rely heavily on the value of consequences of punishment to explicate the source of justification for it.

In view of this it seems reasonable to opt for an interpretation of retributivism as the doctrine that if punishment is to be justified that punishment must be of an offender or wrongdoer for an offence or wrong, in other words that criminal or moral desert are necessary but not sufficient conditions for a justification of punishment. But this is not to make desert a necessary condition of the concept of punishment of course.

So there appear to be good reasons for rejecting the idea that retributivism is only a logical doctrine. One might reject it outright were it not perhaps for the sense of legal guilt outlined earlier in which such guilt is a sign that certain legal procedures have been followed. Such guilt on the part of the punishee may well be a necessary condition of state punishment as we understand it. Also we may reject the constraint that a retributivist justification must be backward-looking, in the sense that the only considerations relevant to its justification concern events or state of affairs occurring or obtaining previous to the imposition of the punishment. But then if we abandon this distinction how are we to draw a line between retributivist and non-retributivist justifications?

One suggestion as to how this is to be done has been
made by Michael. Using certain distinctions drawn by
Alvin Goldman (A Theory of Human Action. 1970, Chapters 1 and 2,) he suggests that the
crucial difference between the two theories of justification is that while
utilitarian justifications depend on relating
punishment to certain causal consequences like
deterrence, retributivist justifications are
underpinned by relating punishment to certain non-
causal consequences. Perhaps we can make clearer the
idea of a non-causal consequence by an example. Morris
offers a justification of punishment in terms of the
restoration of an equilibrium of benefits and burdens.
Now the non-causal consequence of punishment here is
putatively the restoring of the equilibrium of benefits
and burdens. This is not a causal consequence because
it begins simultaneously with the punishment, which the
restoration of the equilibrium does not. Also, if some
x is a non-causal consequence of punishment we may
answer the question of how x comes about by using the
word 'by', that is, punishment is by-related to x. So
in the Morris case the restoring of the equilibrium is
achieved by punishment.

I think one has reason to be sceptical of this
distinction however. Michael considers the objection
that we can find the appropriate non-causal consequence
in connexion with the utilitarian justification of
punishment in respect of deterrence. Thus punishment

may effect a lower crime rate but simultaneous with the punishment is the event of the lowering of the crime rate. Now this event of the lowering of the crime rate is by-related to punishment in the sense that it is achieved by punishment. If the distinction is to hold it must be that that which is by-related to punishment in a retributivist justification does the work of justification while it is the causal consequence and not the object by- related to punishment which does the justificatory work in a utilitarian justification. But there seems little ground for supposing that this distinction is at work. It seems arbitrary to say that what justifies punishment on Morris's account is the restoring of the equilibrium, which is simultaneous with punishment, and not the effect of punishment, i.e. the equilibrium effected, while on the utilitarian account in respect of deterrence, say, it is the effect of a lower crime rate and not the by-related event 'the lowering of the crime rate' which is the morally desirable thing which justifies the practice. The event which is by-related to punishment is equivalent in meaning to 'the causing of such-and-such' where the 'such-and-such' is replaced by the state of affairs which is caused by the punishment. For example, in the Morris case 'the restoring of the equilibrium' is equivalent to something like 'the causing of the state of affairs in which the equilibrium obtains'. But then

what is there specifically in this by-related event

which can have justificatory force distinct from the

state of affairs the causing of which constitutes what

the event is about? In both the Morris and the

utilitarian account it is surely the state of affairs

which is causally resultant which does the work of

justification. So it does not appear that this

distinction will enable us to distinguish between a

retributivist and a utilitarian justification if we

regard Morris's as a retributivist account.

Perhaps the previous claim that punishment is

justified only where the punishee is guilty of a crime

or is morally guilty (that is, has done something wrong

for which he is morally responsible,) could serve to

distinguish retributivist and utilitarian

justifications. Thus in a retributivist justification

of punishment, it might be thought, it is essential to

the operation of the justification that the punishee be

guilty in some sense. Thus it might be claimed that

deterrence can't be retributivist because it may

sometimes justify the punishment of one who is not

guilty, either criminally or morally. But the problem

with this is that it is the sort of condition of which

something we might want to call a retributivist account

might fall foul. For example, if retributivist

justification must make law-breaking a necessary

condition of punishment being justified then an account

which justifies punishment in respect of its capacity
to reconnect the punishee with certain values will
doubtless fail to be retributivist on the grounds that
severance from the values in question may occur via
acts other than infractions of the law. Or conversely,
it is questionable whether on an account framed in
terms of the results of legal infractions, as Morris's
account happens to be, justification must take account
of moral guilt. Perhaps legal rules which fail to
mirror moral ones can be infringed without breaking a
moral rule (as would happen for example where there
was no moral duty to obey the law in question,) in
which case a Morris-style justification might be made
out for the punishment of one who had done nothing
wrong. Clearly, at the least, we need to make the
retributivist condition on the justification of
punishment disjunctional in form, i.e. moral or
criminal guilt must be necessary if the justification
is to operate. But I do not wish to put any weight on
this putative distinguishing feature. In the final
analysis we may want to call an account of
justification retributivist even though it fails to
satisfy such a condition. It may well be that there is
no interesting purpose served now by making a
distinction. Moreover it seems clear that in some
outlandish circumstances punishment of the non-guilty
would be justified, in which case we would have to

think of retributivist justification which required
guilt of the punishee as applying rather to most normal
cases. But if we accept this then it is no longer clear
that the retributivist injunction that we may only be
justified in punishing the guilty is actually superior
to a utilitarian account, where this latter does allow
the possibility of justification of the punishment of
the non-guilty. An alternative would be to accept the
retributivist thesis about the concept of punishment as
requiring guilt in the punishee in which case an
account of justification which would only justify a
practice to which guilt was integral might have
advantages over an account which would justify the
deliberate infliction of suffering on the non-guilty in
some cases. The former account would be designed to
justify only punishment or at least infliction of
suffering on the guilty. Whether or not this would be
an advantage is too large a question to be considered
here but it is worthwhile noting that a restriction on
the concept of retributivist justification may be a
disadvantage if we do not adopt retributivism as a
thesis about the concept of punishment as well.

(ii) Some Minor Retributivisms.

As a preface to a more detailed discussion of certain
modern views regarded as retributivist it is of
interest to have a look at certain ideas about
retributivism which have had some currency. One general
view associated with retributivism for many years was
that the punishment of the guilty is a good in itself
or is of intrinsic value. This is sometimes thought to
be underpinned by the view that happiness and
unhappiness ought ideally to be in proportion to moral
goodness and moral badness. One difficulty with this
underpinning is that moral goodness and badness are
surely moral appraisals of character and hence it would
be at least prima facie consistent with application of
such a principle that the bad man who had never broken
a law ought to be punished while the individual who
leads an otherwise unblemished life should not be
punished for some minor law infraction. Moreover it
seems to make the amount of punishment depend on such
factors as natural cheerfulness. The naturally gloomy
individual may be able to forgo punishment simply
because he has been unhappy all his life; perhaps in

extreme cases he ought to be rewarded rather than
punished and indeed many atrocities are committed by
those who appear to have led dreadfully miserable
lives. So it is at least arguable that such a cosmic
principle of proportionality could well fail to deliver
results which a retributivist might want, e.g.
variation of severity of punishment in accordance with
blameworthiness.

More importantly such a view does not appear to rule
out utilitarian considerations since it is apparently
not incoherent to suppose that something may have both
intrinsic and instrumental value. It may be that it is
good in itself that the guilty should suffer but also
that criminals and would-be criminals be deterred. But
insofar as the idea that the suffering of the guilty is
intrinsically good offers no guidance as to size of
penalty the introduction of a plurality of values into
the justification of punishment seems to bring us back
to the old problem of the lack of proportionality which
may be the result of a deterrence-based punishment
system. Moreover the fact that there is value in the
suffering of the guilty does not preclude the
possibility that such a pluralist justification of
punishment may open the door to the occasional
justification of punishment of the innocent. The fact
that this intrinsic retributivism does not rule out the
value of other punitive effects like deterrence means

that it is a form of retributivism which does not block
the potential justification of punishment of the
innocent, and hence it is not congruent with earlier
retributivist tradition.

Certain doubts arise in connexion with the idea that
the suffering of the guilty has intrinsic value. If
pressed on the question of why something like pleasure
is thought to be of value we may reply that it is
something which motivates action or is an end of
action. But to substantiate the claim that it has
intrinsic value we may fall back on the idea that we
pursue it not in order to get something else. Now it is
apparently plausible that, even if in some
circumstances we seek pleasure for reasons other than
our wanting just pleasure or the absence of pain, in
some cases pleasure just is what we want and not
something else which it facilitates. However, it seems
dubious that we could say with the same confidence that
sometimes the pain of the guilty is sought for its own
sake, even though often it is sought for other reasons,
e.g. because it deters or reforms or whatever. One
reason for the view that pleasure has intrinsic value
might be the supposition that it is somehow an analytic
truth that we seek pleasure and everything else as a
means to it. For suppose that I want **x** in order to get
y, where y appears to be that which I want for itself.
Whatever **y** is it always appears possible to say that I

want the pleasure involved in satisfying my desire for
y. It is because if we have a desire for x we have a
desire to satisfy our desire for x, and the
satisfaction of a desire brings pleasure, that whatever
else we want it can be said that we really want
pleasure, since if some y is putatively the thing of
intrinsic value to us the claim that we want the
satisfaction of our desire for it will always be true.

The first point to make here is that certainly on one
sense of desire it doesn't follow from the fact that we
desire something that we desire the satisfaction of
that desire. Suppose I want an apple and an orange but
I can't have both. I prefer to have the orange. In such
a case although I desire in some sense to have the
apple I also desire not to have it because having it
will result in my not having that which I prefer, the
orange. So in a sense it isn't true that I have a
desire for the satisfaction of my desire for an apple.
Now let us introduce a small piece of special
terminology to clarify this a little. We may say that
although I have a desire for the apple what I want is
the orange. Then the analytic claim can be reformulated
as ' if I want y then I want my want for y to be
satisfied'. But then surely it does turn out that
pleasure is something which I always want. So although
I may sometimes not want a desire for the satisfaction
of a desire to be satisfied it appears true nonetheless

that the object of a certain kind of desire, which

desire we have called a want, is pleasure. But because

of the sense for 'want' we are using it turns out that

our strongest desire is always for pleasure and that is

true whatever it is that we claim to want most.

This may be one source of the idea that pleasure is

of intrinsic value, viz. that it is universally desired

most strongly. Of course one must beware of inferring

from the fact that something is desired that it is

desirable. In this respect we need not conclude that

pleasure is the only thing which has intrinsic value

even if it turns out that it is the only thing which is

wanted. Nonetheless it might be hazarded that the fact

that something was universally desired for itself would

present a good starting point in an argument to show

that that thing was of intrinsic value.

However general criteria for even the plausibility of

a claim that something is of intrinsic value are

somewhat difficult to get at. Rashdall (Theory of Good and Evil, I,

pp. 287-290,) claims that only states of conscious beings can be of intrinsic value,

but then the suffering of the guilty will satisfy this

criterion. He adds to this, however, the claim that the

pain of punishment is an evil and therefore can't be an

end in itself. In the absence of a criterion for

something to be an end in itself it is of course

difficult to assess this sort of claim. Presumably it

follows from the fact that something is an evil that

that something is not of value. Hence if punishment is an evil and an end in itself is something of intrinsic value it is necessarily true that punishment cannot be an end in itself. But of course we may object to the idea that punishment is an evil. If by 'punishment' we understand the 'infliction of suffering on the guilty' then it is precisely whether or not this is an evil which is at issue.

One suggestion of possible relevance here may be borrowed from Moore's idea of organic unities (Principia Ethica, 1903, pp. 27-31.) If an end in itself is equivalent to something which has intrinsic value then it might be argued that the suffering of the guilty could be an end in itself because the suffering of the guilty was less of an evil than the non-suffering of the guilty. Suppose this doctrine has some validity, would it help us? Clearly if we suppose the suffering of the guilty to be one source of value among others to be considered in the justification of punishment if deterrent and other effects of some punishment remain constant whether we punish or not then the principle of organic unities would provide us with a reason for punishing. But it is still not clear that, notwithstanding this, it could help to show that the suffering of the guilty has intrinsic value. If we take the example of organic unity cited by Ewing of the pain aroused in sympathy by one who contemplates the suffering of others, then the

idea is that situation comprising the pain of the
sympathiser and the suffering of others is better than
the suffering of the others alone and the indifference
of another. Now sympathy is presumably an instrumental
good, since it is an attitude which generally issues in
valuable actions. Pain is, let us suppose, an evil when
taken in isolation. And the suffering of others is
also, necessarily, an evil since it is just a case of
pain taken in isolation. Now we are trying to show that
the suffering of the guilty can be something good so we
have to admit the possibility that something evil taken
alone, like pain, can be good under certain
circumstances. So we can't infer from the fact that
pain is an evil that pain aroused in sympathy is an
evil. Nonetheless it is clear that if we are to accept
the example as one in which two evils X and Y are less
evil than one evil X we have to assume that the pain
aroused in sympathy is an evil. But then if we do this
it begins to look as if X and Y can only be less of an
evil than X on its own if the instrumental good of
sympathy outweighs the evil of the pain to which it
gives rise with some remainder. But there is no
especial reason for thinking this. Moreover the obvious
way round the problem defeats the point of introducing
the example. We can say surely that the pain of the
sympathiser is part of the attitude of sympathy which
has been accepted as an instrumental good, in which

case there is no reason to bring in an assumption that the pain of the sympathiser as an evil is outweighed by the good of sympathy. It is simply swallowed up by it. But then of course this move actually deprives us of the second evil which, supposedly in conjunction with the first, yields less of an evil than the first on its own.

It seems that once we accept that it is better to have a world with sympathy in it than one without it it is better to have the pain of the sympathiser and the suffering of others than not. But this rationale is only acceptable because of an assumption which remains tacit in the example. It could be that in the circumscribed situation of suffering and sympathy the sympathy does not issue in anything of value, only the evil of pain (as we are supposing it to be here.) But then on its own merits this situation is worse than the one without the sympathy. It is only because it is better to have the circumscribed situation and a world in which the sympathetic attitudes function in a characteristic way than one in which there is the suffering of others only that we are tempted to think that the circumscribed situation of sympathy and suffering is better than its comparison situation of suffering alone in the example.

So either the example is not one of two evils being better than one or it is simply interpretable as an

example in which two situations each involving some
evil is better than one which involves a certain evil.

How does this bear on the idea of the suffering of
the guilty being of intrinsic value? As the previous
discussion shows it is not by adding another evil that
we improve a situation but by adding another situation
which contains an evil outweighed by a certain good.
Nonetheless this idea could perhaps still be used to
show that the suffering of the guilty improves matters
and is an intrinsic good. The modification of the
organic unities example makes this possible. For on the
original example two evils add up to something better
than one, but then to bring this to bear on the case of
the suffering of the guilty the suffering of the guilty
has to be conceded to be two evils. This can be done of
course: suffering is an evil and so is guilt. But just
as with our example of suffering plus sympathy if we
regard suffering and guilt as evils their coming
together will only be of more value than guilt alone if
the good effects of the suffering of the guilty
outweigh the bad aspect of mere suffering. But for one
thing we have no way of knowing in the abstract whether
this will be true; it certainly seems less likely the
more suffering is imposed. And secondly this doesn't
make the claim to intrinsic value look any more
plausible because suffering of the guilty looks as if
it promotes some end, i.e. has instrumental value. But

then if guilt and suffering in themselves are of negative value it appears that the only positive value is instrumental in kind.

So although this version of retributivism has not been shown to be false it is not compelling in view of the fact that it strikes one as implausible that the suffering of the guilty is a good in itself. The principle of organic unities as I have modified it only makes it look more likely that the suffering of the guilty is at best an instrumental good, while the fact that the view is consistent with the introduction of other utilitarian values in the justification of punishment only serves to diminish the force of whatever this intrinsic value is by comparison with much clearer traditional values of deterrence and reformation. Finally it is only of marginal interest in our present project which will be to look for a retributivist account of justification of punishment which actually justifies something like the retribution we have already outlined in preceding chapters. Clearly intrinsic retributivism offers no justification as it stands for the kind of differences in a penalty scheme which I have taken to be the hallmark of retribution and this deficiency is not satisfactorily made good by the idea of a proportioning of happiness and unhappiness to moral goodness and badness, a point noted at the outset of this section.

Walker offers a different account of the retributivist motivation (Why Punish? 1991; pp. 83-7.) Strictly he proposes both a psychological account of the feelings of a retributivist and a logical account of what is taken to be the justification for a retributivist position. The psychological account is that the retributivist motivation for punishing an individual is that he has broken a rule; if there is a rule that we punish **x** in circumstances **c**, and this is why we punish **x** in **c**, that is, we punish without reference to any other considerations (for example whether there is good reason for such a rule, whether the rule is just etc.,) then our motivation for punishing is retributivist.

The logical account of retributivism, as Walker puts it, is that an individual deserves to be punished if he has broken some rule r, and another rule prescribes a certain punishment for infraction of that rule. What legitimizes punishment is that there is a rule which prescribes it.

Now in respect of his psychological account Walker might be thought to be facing a chicken-and-egg situation - it might be argued that it is just as plausible that because we feel people ought to be punished in certain circumstances that we have laws prescribing such punishment rather than it being the case that we have the feelings we do because we have certain kinds of laws. What concerns us here however is the logical

account. Now of course it is a truism in one sense of
what it is to legitimize something that a law to the
effect that we punish for infraction of some rule r
legitimizes the punishment of someone who has infringed
r. This is the sense in which to legitimize something
is simply to make it lawful. However it is far from
being a truism that because there is a law that someone
who infringes r be punished to a certain extent we are
justified in punishing one who has infringed r (to that
extent.) Even if I am a guard in a concentration camp
and the killing of Jews has become something which I am
legally required to do the fact that I am justified in
killing a Jew is surely a contingent matter, and this
contingency is not dependent on legality at all. If I
have to kill someone because I'll face a firing squad
if I don't this hardly shows that it is the putative
lawfulness of what I do that justifies my doing it. And
it seems quite clear that if I can forgo executing a
Jew without incurring terrible consequences I am
morally required to do so. It is far from obvious that
one is always morally required to follow an unjust law.

It might be objected that where laws are made via a
democratic process they must in virtue of this fact be
just. Hence in a democracy we are morally required to
obey all laws. But this is to suppose that laws are
made as a result of a procedure which is one of pure
procedural justice, that is, a procedure which is such

that whatever results from it is ipso facto just. But it is plainly false that such a democratic procedure is one of pure procedural justice. More likely parliamentarily enacted law is an example of imperfect procedural justice. It seems that we may criticize a law for being racist even though it has passed through parliament, and even if we were to suppose that any law which has gone through parliament is in some sense the will of the people surely we can make sense of the idea of being unjust to oneself. In any case the idea that a law which I abhor is the result of my will is just a farce, if we are to understand what is the result of my will in any normal way.

A different interpretation of the Walker point is that where a system of rules has some justification, let us suppose a utilitarian one, then we may nonetheless justify a particular application of a rule on the grounds that it is required by a rule of the system. But as I have already suggested there seems to be some contradiction at work in the idea of a system of rules which has a utilitarian justification individual rules of which have some other justification. If it follows from the fact that the system is justified by utility that individual rules are so justified then to say that the individual application is justified in virtue of its being a rule of the system is simply shorthand for saying it is

justified by utility. So the legality itself of an action does not entail that it is justified, and where the legal system is justified the fact that an individual application is in accordance with a law seems to signify that its justification has a basis in whatever justifies the system of law itself. In other words where lawfulness of an action can justify it it is not likely that mere lawfulness itself does the work of justification but something else which justifies the legal system. So if retributivist justification is of the kind Walker proposes it is a non-starter. Moreover if the justification of a punishment is simply that there is a rule prescribing punishment it need not be that the justification requires proportionality between offence and penalty in order to operate. The law itself might prescribe extremely disproportionate punishments. Of course it might be objected that such a system of punishment could not be justified on utilitarian grounds but to suppose that the system justified by utility would adhere to the principle of proportionality, itself a matter of contention, again meets the objection that it is not in virtue of there being a rule which prescribes proportional punishment that proportional punishment is justified. So although Walker's suggestion has the virtues of clarity and absence of metaphor which he boasts for it it does not

appear to account for the singular power of the

retributivist pull in our thinking on these matters.

A distinction has been drawn by Mackie between

positive and negative retributivism (Persons and Values, 1985, pp. 207-8.) The latter

position is that a person must not be punished unless

he is guilty, or, in other words, guilt of the punishee

is a necessary condition for his punishment to be just.

Now if retributive punishment is defined as state

punishment which is also retribution it follows that

any justification of retributive punishment is eo ipso

justification of punishment of the morally guilty.

However, such a justification would not necessarily be

a form of negative retributivism in the sense given.

Firstly, we can imagine a justification of a practice

of punishment which did not show that that practice was

just. There is a clear distinction between showing that

an action is justified and showing that it is just. For

example, it might under certain bizarre circumstances

be justified to punish an innocent man, but no one

would suppose that this was just, either viewed

independently of the justification of the action or in

virtue of its being a justified action. So a

justification of retributive punishment might fail to

show that such punishment was just. But then it would

also fail to show that punishment of the morally guilty

was just. Secondly, we can imagine a justification of

punishment of the morally guilty which failed to make

moral guilt a necessary condition of the justification of that punishment. For example, if it could be shown that all sinners ought to be punished, and that the class of agents of morally wrong actions was a proper subset of the class of sinners, then it would have been shown that retributive punishment was justified, but also that there was scope for the justification of a practice which punished individuals who were not morally guilty. So in the search for a justification of retributive punishment success does not guarantee that we have found an account which satisfies the conditions of negative retributivism.

Positive retributivism on the other hand, taken as the view that 1. the guilty ought to be punished or, to retain symmetry with negative retributivism, as the view that 2. justice requires that the guilty be punished, is again a criterion of an account which may not be met simply by providing a justification of retributive punishment. For if retributive punishment is justified, or a certain kind of proportionate punishment is justified, then again it may not have been shown that this system of punishment was just or meets condition 2, although it would be entailed that we had an account the justificatory scope of which included punishment of acts performed by the guilty. In other words such a justificatory account would include justification of the punishment of the guilty, and even

if it included justification of the punishment of the non-guilty (however unlikely this might be) this of itself would not debar it from being a positive retributivism on 1 or 2.

So the account we are seeking may fail to be either a version of negative retributivism or positive retributivism.

Another view which has been assimilated to the retributive position is Lord Denning's dictum that the ultimate justification of any punishment is not that it is a deterrent but that it is the emphatic denunciation by the community of a crime (198 H.L.Deb. 5 July. 576, 577, 596. 1956.) I shall not discuss this at length, because the objections to this view are fairly obvious and shall in any case be treated in greater depth when we come to look at the view put forward by Feinberg concerning the expressive function of punishment. Benn has pointed out that denunciation needn't involve the deliberate infliction of suffering and it is this which is to be justified in punishment. Of course if denunciation simpliciter were justified, and it could be shown that punishment was a way of denouncing a crime, then the fact that denunciation didn't entail infliction of suffering would be beside the point. But this appears highly implausible, for if denunciation can be achieved in various ways it would seem that we are morally obliged to choose that way which causes the least suffering. But then it is hardly

likely that any way of achieving denunciation is equally justified and, moreover, it seems prima facie unlikely that punishment would be the most cost-effective way of achieving this denunciation. A further difficulty concerns the point behind retributive equivalence that there be some proportion between punishment and moral gravity of offence. Even if it could be shown that greater punishment constitutes a greater denunciation, it does not appear to be part of the logic of denunciation, where to denounce means to speak out or rail against, that we are required to speak out more strongly the greater the immorality of that against which we are speaking out. And clearly if we concentrate on the feature of denunciation as a way of speaking out punishment doesn't appear to be a form of denunciation let alone entailed by it. So Lord Denning's view does not appear to provide the key to a justification of retributive punishment as it is being understood here.

Cottingham gives us a good round up of the different kinds of accounts which have been proposed and in large part accepted under the retributivist banner. I shall briefly consider a few of these views. Firstly there is the repayment view. This has the virtue of being in line with the meaning of the Latin 're-tribuo' meaning to pay back. Now as Cottingham notes there is plenty of precedent for the belief that one who breaks the law

owes a debt to society and that suffering can somehow

pay for sins. For all that, however, it is unclear why

these beliefs are held. If we understand a debt in

terms of a duty to give something of value then we are

supposed to understand the commission of a crime as

giving rise to a duty to give something of a certain

value to society. One way we might make sense of this

would be to say that in committing a crime the criminal

takes from society something of a certain value and

hence owes something of equivalent value. But what does

he take? Perhaps he detracts from the stability of the

social order. At least it is fairly clear that this

order has a value. But then the obvious coin of

repayment would be to give something to the stability

of the social order, perhaps by doing some special

community service. But the problem is that we are not

seeking to justify community service but rather the

infliction of suffering. And it is difficult to see how

suffering on the part of the criminal has value

equivalent to that lost to the social order. Moreover

we want an account which shows that more suffering of

the criminal is of more value. For if we assume,

perhaps not unreasonably, that the worse a crime is the

more destabilizing it is and hence the more of value

which is lost, we would need to show that, in order to

justify greater punishment for greater wrong, more pain

for the criminal was more valuable and hence consistent

with repayment of the debt incurred by wrongdoing. But unless we think specifically of the victim it seems quite a contingent matter that the criminal's greater pain should be of greater value. And if it is the value to society of that suffering it seems quite clear that something other than pain would give society something of compensating value.

A different view considered by Cottingham is the annulment view of punishment. I shall not consider the Hegelian account of this view but instead shall concentrate on two recent accounts, one by Cooper, an explicit attempt to make sense of the Hegelian idea that punishment annuls a crime, another by Hampton which, though not ostensibly based on Hegel, does offer an attractive account of punishment part of which deals with features of punishment in virtue of which it annuls crime.

I shall begin then with Cooper's ingenious account of how we might begin to interpret the Hegelian idea that punishment annuls a crime. The rough idea is that the commission of a crime involves an implicit denial of a right, the right of the victim of the crime, by the criminal. However it is because we punish that we show that there is indeed a right violated by the criminal and so we show, by punishing, that the denial of the right is false. Punishment shows that there is a right because it is a necessary and sufficient condition of

the existence of a right under certain circumstances. It is necessary because if there were putative rights and we didn't punish for violations of them then people would be correct in thinking that these were just token rights, not real rights at all; punishment is sufficient for the existence of rights since, given the existence of some rule, a statute for example, that there is a right, our punishing completes a conventional procedure which, when carried out demonstrates that there is indeed a right. Since punishment demonstrates the existence of a right the claim implicit in the criminal's action that there is no right is shown to be false or invalid.

It is important to remember that Cooper is only providing a preliminary account of how one might interpret Hegel's remarks, but since I think he does succeed in offering an interpretation which is not only free from the apparent incoherence in the idea of annulling a crime or wrong, as opposed to say a marriage, but also is quite a respectable attempt to justify punishment, I shall take it as an account worthy of consideration on its own merits.

Now the obvious problem of course is the recurring one that if we take the account as a justification of retributive punishment, and not just punishment simpliciter, then it is not clear how the account would show that only punishment proportional to moral

wrongness of offence would be justified. If we suppose that the existence of some class of rights **R** depends on punishment for infractions of members of **R**, or at least some of those infractions, then any punishment would seem to fit the bill, even a grossly disproportionate one. Against this it might be objected that certain punishments aren't consistent with the existence of the corresponding right. For example if our punishment for murder is a fine of 10p then people don't really have more than a token right not to be murdered. But of course to say that some punishments don't establish the existence of a corresponding right is not to argue for proportionality. For, it might be said, provided that we give the criminal good reason not to violate a certain right, then we have established the existence of that right. But if one has good reason not to violate a right just because its violation involves an unpleasantness which is a function of the probability of being punished and the punishment itself, and this unpleasantness is greater than that of not performing the crime, then giving people good reasons not to infringe laws is consistent with non-proportional punishments. It is presumably consistent with the death penalty for infractions almost impossible to detect for example.

But perhaps even this can be countered. It might be objected that if we endorse the claim that justice

requires that the criminal only forfeits a right he has
violated then non-proportional punishment of the kind
in the example above will fail to respect or protect
the rights of the criminal. But then if the idea behind
the claim that punishment for the violation of a right
is a necessary condition of its existence is that
punishment is needed to manifest the society's desire
to protect the right, we might beef up the claim of
necessary connexion to say that a certain kind of
punishment is necessary to the existence of a certain
complex of rights.

Two points are relevant here. Firstly, we have
amended the idea that punishment for a right violation
is a necessary condition of the existence of that
right. This claim as we have seen is not enough on its
own to deliver a justification of retributive
punishment. What we have now is something like the idea
that a certain system of punishment is a necessary
condition of a certain institution of rights. But if we
are to justify an institution in which a criminal only
forfeits certain rights, which ones being governed by
reciprocity, the question now is what will justify such
an institution. But now it can be seen that the whole
question of what justifies retributive punishment is
left substantially untouched. If we suppose a certain
distribution of rights after commission of a crime is
morally required of us and such a distribution requires

retributive punishment the key question is of course why we are morally required to distribute rights in such a way. Doubtless once we suppose that the criminal only forfeits certain rights, and that those which he has forfeited make punishment morally permissible, and further, that punishment is a logical requirement of the existence of the rights which have been violated, if we can justify the existence of the violated rights, and can justify the distribution of rights to the criminal after the crime, then we have justified retributive punishment. But this leaves open the justification both of the distribution of violated rights and the distribution of retained rights by the criminal after the crime. But surely it is no exaggeration to think that this is the main part of the problem of the justification of retributive punishment. So ingenious as Cooper's account is, and the account may well be the right interpretation of Hegel, it leaves the justification problem more or less untouched.

Another way in which we might understand the idea that punishment annuls a crime is provided by Hampton (A New Theory of Retribution. 1991; Chapter 11.) According to Hampton if we suppose that individuals have value certain kinds of treatment of them are appropriate. If we fail to treat someone in the manner suited to their value by treating them in a manner appropriate for one of lower value than they are we

demean them. Now she says that a person wrongs another iff she treats him in a way that is objectively demeaning, hence we wrong someone if we treat them too poorly given their value. She suggests that in order to flesh out this proposal we follow Kant in the assumption that people are of equal, objective, intrinsic value. Now if someone wrongs another he has, apparently, sent a message about his relative value vis-a-vis his victim. This message is that he is of greater value than his victim. But of course if we assume that people are of equal value this message is a falsehood. We can deny the truth of this message and send the correct message about the equal value of victim and offender by punishing the offender or wrongdoer. How is this to happen? Punishment involves a defeat or mastery of the wrongdoer equivalent to the mastery which the wrongdoer had over the victim and since the punishment is inflicted in the name of the victim it is as if the victim were to achieve that equal mastery over the wrongdoer. It is because the victim has this equal mastery that the superiority implicit in the wrongdoing is denied. In this way the correct relative value of the wrongdoer and victim is symbolized by the act of punishment and so punishment annuls insofar as it corrects the false claim to superior value implicit in the original wrongdoing.

Even if this is the correct sort of account to explain and perhaps justify retribution there are many questions of detail which remain unanswered. To begin with the justification and motivation of the retributivist is supposedly one of setting the moral record straight. By punishing we assert or reaffirm moral truth. But one might wonder whether the assertion of moral truth is sufficient to justify a practice like punishment. For Hampton admits that in many cases the assertion of superiority implicit in the wrongdoing may be believed neither by the public at large nor the victim. I would go further to suggest that it need not be believed even by the wrongdoer. But in that case why do we need to go to all the expense of punishment to affirm something which was never in doubt? Moreover can we justify the infliction of pain or the 'humbling of the will' involved in punishment by this somewhat academic observation, rather like if we spent a lot of money promulgating and encouraging belief in the claim that there is some red in the Union Jack? Of course there are important effects of this observation to bear in mind, like the way that the manner of the assertion may deter others from committing crime, and its supposed content may reform others, but Hampton is explicit that this sort of thing is not part of a retributivist justification.

Secondly, there seems to be some problem with the idea that the wrongdoing is sending a message. If the wrongdoer believes that he can inflict pain on another whenever he feels like it because that other is of less value than he we would expect that wrongdoer to feel justified in his action. But surely one can do wrong without feeling that one is justified in it at all. In fact one's conscience may be telling one quite the opposite all along. But if we may suppose an absence of this feeling we might equally suppose the absence of the belief on the part of the wrongdoer that there is a difference of relative value between wrongdoer and victim. Indeed the wrongdoer may believe that their equality of value is a reason for harming the victim if he believes that his own value amounts to nil. But the point is that Hampton does not want to say that the message implicit in the wrongdoing is in any way dependent on any beliefs which the wrongdoer has. But then it seems that the message involved in the wrongdoing is not dependent on an intention to send that message. Why should a wrongdoer intend to send a message he does not believe unless he wishes to deceive others? Of course I can send a message without intending to, as when I put out flags to dry on a line and send a message 'England expects..' but in the case of an action involving no conventional symbols how are

we to determine the message except by reference to intentions, conscious or otherwise, of the agent?

So it is not clear how any message can be sent, let alone one of the required form. And even that required form seems to pose problems. Certainly the idea of mastery can make sense in the context of punishment and wrongdoing. But even this is not felicitous in certain contexts, e.g. embezzlement. Who is mastered if I defraud the Inland Revenue? Indeed in such cases accounts like those of Morris which emphasize the justificatory role of enjoying the goods of social cooperation look more comfortable than that of Hampton. But even if we accept that there is a kind of mastery here how does the idea of a relative value fit into this? If I master someone it does not appear to be entailed by this that I am of greater value than my servant, or that I have sent this message, especially if the value which is relevant is the assumed value of equality (which apparently cannot be contradicted by any action anyway.) More to the point, if I and my neighbour are at one another's throats but neither can get mastery of the other, or periods of mastery alternate, again it seems that this may have significance with regard to ascriptions of power between us but not with regard to ascriptions of value. If the content of the message concerns an evaluation then equivalence of mastery does not force itself upon

us as the kind of symbolism which will deliver that message. No doubt this is partly due to the fact that we are used to regarding differences of power as irrelevant to the kind of worth which is being assumed to attach intrinsically to the individuals involved.

It would seem that there are technical difficulties involved in the idea that the infliction of defeat is morally permissible and hence that it involves no demeaning of the wrongdoer. If the wrongdoer's equal value stays the same throughout prima facie it would seem to argue for equal treatment of wrongdoer and victim at the hands of the state. For even if the wrongdoer has claimed unequal value treating him unequally is only appropriate if that claim is true, which it is not. If we, qua the state, wish to underline the evaluation that wrongdoer and victim are equal, equal treatment seems to recommend itself. Rather it would seem that we have to suppose that as a result of wrongdoing the wrongdoer is of less value and hence the differential treatment of him relative to the victim is appropriate. And this runs counter to the initial assumption about value. This cannot change, according to Hampton, unless perhaps some key feature is lost, for example the capacity to reason, and in this context there is no reason to suppose the relevant change.

Perhaps the problem is the result of the intervention by the state, and many difficulties would be circumvented if victim and wrongdoer could settle things among themselves. If the victim wants to assert that they are of equal value he might reason as follows: you have treated me as if I am of low value; I have a moral duty to treat you in accordance with your value; since we are of equal value I must also treat you as being of the same low value; I owe you treatment with this same evaluation implicit in it. In this way by adhering to an account in which treatment is morally permissible or justified if in accordance with an evaluation which reflects the standard evaluation to be adopted we might get into the neighbourhood of a justification of something like the treatment inflicted by the state but only when inflicted by the individual. But it is to be noted here that if we start out with a standard assumption about equal high value of individuals, rather than just equal value, something like the response proposed for the victim above would be ruled out as just as wrong as the treatment he received at the hands of the initial wrongdoer. If poor treatment is wrong then so is poor equal treatment.

This concludes our treatment of a selection from the range of minor retributivisms. I want to consider now in greater detail some of the most important accounts

of recent years aimed at the justification of
punishment.

(iii) Some recent retributive theories

(1) Morris on benefits and burdens.

I want to consider our account of retributive state punishment in the light of Morris's account of the fairness of punishment. I shall now sketch one part of his enormously rich and inspiring paper 'Persons and Punishment' and discuss certain features of it.

Imagine a group of people and rules which apply to all members of the group. Compliance with these rules provides all with the benefits of non-interference with what each person values. However these mutual benefits require the assumption of a burden in the form of the exercize of self-restraint by these people over inclinations which would create, if given way to, substantial risk of interference with others in certain ways, namely interference with what others value. Where a person does not exercize self-restraint though he might have he voluntarily puts down a burden which others have assumed and so gets the advantages of the system, in the form of non-interference with what he values, without assuming the voluntary burden which

others have assumed. He gets the advantages without all the disadvantages. But this constitutes an unfair advantage; punishment removes this unfair advantage.

Now the first key point in the justification of punishment as fair, call this F1, is Morris's idea that where there is an equal distribution of benefits and burdens it is only fair that there be some mechanism which minimizes the likelihood of the occurrence of a maldistribution, i.e. a non-equal distribution of benefits and burdens.

The second respect in which punishment is fair, call this F2, is that it erases the unfair advantage of one who voluntarily renounces the burden of self-restraint yet continues to enjoy the benefits accruing from the exercize of self-restraint by others.

Let us consider F1. The first difficulty concerns the disambiguation of the concept of an equal distribution of benefits and burdens. Let us distinguish between an active and a resultant distribution. Suppose that three individuals **A, B** and **C** have 1,2, and 3 respectively, though we need not question here what exactly these numerical expressions stand for. Now if I give to **A,B** and **C** respectively 3,2 and 1 then my active distribution is of 3,2 and 1. However the resultant distribution is 4,4,4 to **A,B** and C. What this means is that although the active distribution was unequal the resultant distribution was equal.

Thus an equal distribution of benefits and burdens may be one in which the active distribution, but not the resultant, is equal or one in which the resultant but not the active distribution is equal.

In line with the distinction between an active and resultant distribution an equal distribution of benefits and burdens occurs where either :

1. an assignment of benefits and burdens is such that the benefits assigned which are enjoyed and the burdens assigned which are borne are equal though it may not be that the total resultant distribution of benefits and burdens is such that benefits enjoyed and burdens borne are equal.

2. the total resultant distribution is such that benefits enjoyed and burdens borne are equal even though some particular assignment may be such that the benefits enjoyed and the burdens borne as a result of that assignment alone are not equal.

Next we have to get clearer about what it is for the benefits enjoyed and the burdens borne to be equal. Clearly, B is not necessarily an equal burden for x and y just because B represents the same task. For example where x and y have the job of carrying the same number of heavy pails of water up a hill all day long the fact that x is stronger than y may have a bearing on whether the burden of the job is equal for them. Thus we might say that although x and y bear the same burden B, B

does not constitute the same burden for **x** and **y**.
Moreover sameness of burden doesn't seem to be a
necessary condition of equality of burden, quite apart
from its not being a sufficient condition. X and **y**
might be performing quite different jobs, yet surely
the burden which each job constitutes might be the
same.

This raises the question of what it is for **B1** and **B2**
to constitute equal burdens for **x** and **y**. It is not easy
to decide which factors are relevant here, let alone to
give an exhaustive list of such relevant factors. For
example, is it relevant to the matter of burden in the
case in which **x** and **y** are carrying water that **y** enjoys
his work and **x** does not? If we leave aside the rather
vexed question of what makes something a burden certain
factors appear to be relevant to how much of a burden
something is. Effort appears to be relevant in many
cases. If we suppose that all other factors relevant to
degree of burden are equal then if **x** has to expend less
effort in the bearing of burden **B1** than **y** has to in the
bearing of burden **B2** then **B1** constitutes less of a
burden for **x** than **B2** does for **y**.

Is unpleasantness a factor? Take the example of a
chronic illness. Such an illness may be a burden. But
then bearing such a burden will involve no effort
expenditure at all if it is having the illness which
constitutes our bearing of the burden; we have not

usually expended any effort in having an illness. But then we should not assume that the unpleasantness of a burden figures in an assessment of what it constitutes simply because the more unpleasant something is the more effort is expended in putting up with it. Of course size of the burden may be affected by effort expenditure as a result of unpleasantness as well; the unpleasantness of affliction may force us to exert ourselves in various ways.

So it appears that there may be cases in which the burden constituted by something with an unpleasant feature is not necessarily drained by the effort coping with such unpleasantness draws from us.

This is certainly not exhaustive of the factors by reference to which size of burden is to be established. But leaving the problem of this incompleteness aside how are we to determine relative (constitutive) size of burden even by reference to just the two factors mentioned?

Problems occur where both factors are involved in the assessment of a burden. For example, x and y may bear the burdens B1 and B2 respectively, and it may be that B1>B2 in respect of effort required in bearing the burden, while it may also be the case that B2>B1 in respect of unpleasantness. Here there may be obvious cases in which we could determine the relative sizes of the burden, e.g. where B1 required only slightly more

effort than B2 while B2 was very much more unpleasant
than B1 and they were equal in all other respects
relevant to determination of size of burden. In such a
case the burden of B2 would be greater than that of B1;
but there will be cases in which no obvious
determination is available.

Secondly problems may occur where one burden in which
effort is the only relevant factor for assessment is to
be compared with another in which the factor of
unpleasantness alone is relevant. If B1 requires a lot
of effort on the part of x but is not unpleasant and B2
is very unpleasant for y but requires no effort how are
we to compare the two?

Thirdly we have the problem of comparing sets of
burdens. For example if B1>B2 and B4>B3 and B1,B3 are
the burdens of x, B2,B4 the burdens of y problems may
arise in determining whether the burdens of one are
greater than the other.

Now we need to interpret what it is for x and y to
bear equal burdens. I shall call factors relevant to
determination of the size of a burden b factors. Then
where B1 and B2 are equal in respect of unpleasantness
and effort required to perform them (where they are
actions,) and they are equal in respect of all other b
factors then B1 and B2 constitute equal burdens for x
and y. But even this falls short of a clear criterion

in the absence of an exhaustive list of the things
which are **b** factors.

Again when it comes to the question of determining
whether two sets of burdens borne are equal only
sufficient conditions are available, even if we suppose
we have exhaustive knowledge of the **b** factors. If we
suppose that the size of burden which a set of burdens
constitutes is to be found by thinking of the
individual burdens in a set as comprising one overall
burden it would be reasonable to measure this overall
burden by reference once again to **b** factors, only this
time the effort **b** factor of the overall burden would be
the sum of all the individual effort expenditures,
while the unpleasantness relevant here would be the
unpleasantness of the individual burdens taken overall.
Finally the other **b** factors for the overall burden
would have to be assessed. Such other factors would be
applicable to the overall burden because this would be
treated as a burden just like one of the elements of
the set of burdens. But this means of course that the
determination of whether two sets of burdens were equal
would be subject to the same vagueness as to the nature
of other **b** factors and to problems concerning the trade
off between the unpleasantness and effort
characteristics of burdens. The only clear case of
equality between sets of burdens would then be the same
as that for a pair of individual burdens: a sufficient

condition of two sets of burdens being equal would be equality between the two overall burdens comprising the individual burdens of the respective sets based on equality between the two overall burdens in respect of unpleasantness and effort and the remaining **b** factors.

It is worth noting here that no explanation has been given of how the unpleasantness of the overall burden comprising the set of individual sub-burdens is related to the unpleasantness of those individual burdens. For example, it might be thought that where we have two equinumerous sets of burdens **b1** and **b2** and there is a one-one correspondence between the members of the two sets such that each member of one set is mapped onto a unique member of the other to which it is equal according to the **b** factors then the overall burden comprising the burdens of **b1** is equal to the overall burden comprising the members of **b2**. But it has not been shown that where the individual burdens of two sets are equally unpleasant in the way explained above they are equally unpleasant taken overall. It is not logically impossible that individual members of two sets which are equally unpleasant individually in the way sketched fail to be equally unpleasant taken as a whole-the unpleasant whole may well be more than the sum of the unpleasant parts. But then in a case in which the individua of the sets are equal with their correspondents with regard to all **b** factors and in

respect of all **b** factors apart from unpleasantness the overall burdens are equal it still remains to be seen whether they are equal in respect of the factor of unpleasantness overall. So it remains uncertain even how this overall unpleasantness is to be assessed.

Perhaps more clear criteria are available concerning equality between two benefits enjoyed. Reasonably, benefit **b1** to **x** is equal to benefit **b2** to **y** where **b1** is of equal benefit to **x** as **b2** is to **y**. I shall regard something as being of benefit to some individual where on balance what he or she gets from it is of greater value than what he or she loses by it. Another way of putting this is that something is of benefit where it is a greater source of value than disvalue. Thus, reasonably, two things are of equal benefit to two individuals where they are the source of equal net value (net=balance of value over disvalue) to the two individuals in question. Hence a group of benefits **A** enjoyed by **x** would be equal to another group **B** enjoyed by **y** where the net value derived by **x** from **A** was equal to that derived by **y** from **B**.

Ignoring the question of how we ascertain whether two things are of equal net value to two individuals and treating the foregoing criteria for equality of benefit as necessary and sufficient we have: **x** and **y** enjoy equal benefits and burdens if

1. the overall burdens of **x** and **y** constituted respectively by the burdens of **x** and **y** are equal in respect of all **b** factors including the factors of effort and unpleasantness and

2. the overall benefits enjoyed by **x** and **y** constituted by the individual benefits they enjoy are of equal net value. (This remains a sufficient condition of equality of benefits and burdens for the reason that the first condition above is a sufficient and not necessary condition, and so both conditions taken together become sufficient rather than necessary and sufficient.)

Now with these two conditions in place it must be remembered that an equal distribution of benefits and burdens is either active or resultant, in which case the above conditions apply either to what is actively distributed or to the total distribution all things considered. However, given the way we are assessing a burden this introduces an ambiguity over how the burdens of an active distribution are to be assessed. For consider the b factor of unpleasantness. If certain burdens are distributed to individuals who have an existing base of burdens already in place the unpleasantness of any additional burdens will doubtless be coloured by their being conjoined with the existing burdens. This means that the measure of an overall burden in a resultant distribution may be assessed

straightforwardly in terms of its **b** factors, i.e. the
conjoined **b** factors of the individual burdens being
distributed and those already distributed. However the
b factors of the overall burden of an active
distribution, that is, the conjoined **b** factors of the
burdens making up the active distribution, might be
assessed either in the light of existing burdens or
not. The difficulty here is introduced by the fact that
burdens already borne and subsequently introduced will
affect one another via their **b** factors. The
distribution involved here is not just the simple case
in which, say, Mandy already has two sweets and is
given three more, so that we can speak of an active
distribution of three sweets and a resultant
distribution of five; how many sweets Mandy is
distributed by the active distribution is not affected
by the number of sweets she already has. This
difficulty is all the more important because in
practice the important distributions are not going to
be of the sweets-to-Mandy variety but more akin to the
benefits-and-burdens kind. I shall call the sweets-to-
Mandy kind of active distribution an independent
active, the other a concomitant active distribution. In
practice it is the second kind of distribution which is
important. If **x** is already labouring under considerable
burdens what matters from the point of view of fairness
is almost certainly going to be what extra burdens will

mean to him given his existing position. However we should not attribute too much significance to this distinction because the real predicament of the already burdened individual will be amply registered by the resultant distribution in any case.

So much then for the attempted clarification of what is meant by an equal distribution of benefits and burdens. Let us consider F1 in the light of this.

Is F1 plausible?

If, as would seem reasonable, we understand the equal distribution as a resultant one, one consequence of our interpretation of equal benefit is that F1 appears to impose too strict an egalitarianism. For if the members of a society derive equal net value from their spheres of non-interference the sacrifice of liberty involved in the application of a mechanism to ensure the occurrence of no inequalities in net value derived from such spheres would presumably be thought unacceptable. For one thing if my fellow makes more of his freedom than I do it would seem harsh in the extreme to endorse some mechanism which would try to remove some of the extra value he gets from his freedom to make us equal. This is not to say that if **A** and **B** are equally free and **A** makes a lot more money than **B** it might not be fair to have some mechanism of taxation which aims at least to lessen the inequality between the two. Rather the problem of the interpretation of equal benefit as equal

net value derived is that it is too sweeping in scope.
I might make more of my freedom because I have good
luck and a more cheerful disposition than the next
fellow but it would seem unfair, if we assume that my
fellow and I started off getting the same value from
our freedom, say, to restrict my sphere of non-
interference somewhat so that equality of value from
freedom should be maintained between me and my gloomy
fellow.

The alternative here is to interpret equality of
benefit of sphere of non-interference as formal
equality in respect of rights. Thus if people have the
same rights which define their freedom they share the
benefits of these spheres equally. But the problem with
this is that equality of rights can amount to no more
than token equality of liberty. For example, if anyone
can vote provided that they own a certain acreage of
land, but not everyone has the financial capacity to
acquire this land, the fact that everyone has the same
right to vote does not yield the same freedom to vote.
Equality of rights will not translate into equality of
freedom since whether or not I am free to do x may
depend on matters other than an obligation on others
not to interfere with my doing x. Thus if we construe
equal spheres of liberty in terms of equal rights to
certain things liberty may be very unequal and the idea

that it is only fair to maintain this inequality of liberty loses its prima facie plausibility.

We can perhaps circumvent this problem if we turn away from ideas of rights and benefits and think instead of freedoms. We may suppose that there is an equal distribution of freedoms. But this is of little help if, in the absence of the two interpretations of equality of freedom offered, we do not have in mind what equality of freedom is. Certainly, in fairness to Morris, the interpretation of equal benefit as equal net value derived from something may seem inappropriate since we may speak of A and B sharing the benefit of health equally and yet maintain consistently that A but not B utilizes this health. But is there an alternative to the interpretation of equal freedom as equal rights?

Reasonably, there are certain prerequisites to the enjoyment of a right. In order that I enjoy the right to vote it can't be that I am housebound and that there is no provision for those physically unable to reach the polling station. Thus we might suppose an initial situation in which individuals have equal rights and also have a right to the necessary prerequisites to the enjoyment of those rights. Since one of those rights would be the right to the prerequisites to the enjoyment of the rights in question it would appear that one had a right to the prerequisites necessary for the enjoyment of the right to the prerequisites

necessary for the enjoyment of the rights. But this would not entail that one had those prerequisites (but rather a right to the prerequisites of such prerequisites) and so equality of rights might again amount to no more than the formal equality of rights consistent with a cynical scenario of considerable inequality of freedom. Rather to get the scenario of equal freedom of a substantive kind off the ground we need to suppose not merely the right to the prerequisites but also the enjoyment by all of that particular right to such things, viz. the possession of the prerequisites.

So let us suppose that we have equality of rights and equal and full enjoyment of the right to the necessary prerequisites for the enjoyment of those rights. Can we suppose on the basis of this that we will have something like an equal distribution of benefits and burdens? It would seem not. For one thing if we are to interpret equality of benefit from spheres of non-interference as equality of freedom where this equality is substantive and not cynically formal, equal rights will not deliver equal freedom even now, since no account has been taken of how comprehensive these rights are. A and B may be unequally free if much of their activity is conducted in a milieu where no rights prescribe rules of behaviour; if most of life is a jungle it may be true that the strong have more freedom

than the weak. All we can say perhaps is that, with

respect to the sphere of non-interference defined by

the rights, the individuals have equal freedom.

Let us suppose then a distribution of an equal set of

rights including the right to the prerequisites of the

enjoyment of those rights and which latter right is

enjoyed. Is it plausible that it is only fair to have a

mechanism to maintain this distribution?

Suppose it is required by fairness that we continue

to abide by the rules which specify the rights in the

above situation. Then it might be argued that it is

only fair to have a mechanism to maintain this sort of

situation. Now there are certainly difficulties with

this approach to a justification of F1. For one thing

even if it is fair to have a certain state of affairs

it does not appear to follow that it is fair to have a

mechanism which maintains this state of affairs tout

court. For, prima facie, it is far from obvious that

any state mechanism which minimized the likelihood of a

deviation from a particular situation must itself be

fair. It is possible that a mechanism securing this

result would involve torturing people for very minor

offences. But, it could be argued, such treatment would

be unfair to the individual guilty of the minor offence

and certainly of a kind not envisaged within the

justificatory schema of punishment which Morris is

proposing. F1 is questionable even if we can establish

that a continuance of obedience to the rules of the
initial situation of equality is required by fairness.

Let us consider an attempt to show that something
like the situation described requires people to
continue to obey the rules or laws as a matter of
fairness. One such account is provided by Rawls in his
mercifully concise paper 'Legal Obligation and the Duty
of Fair Play.' The basic idea is that where a just
scheme of social cooperation is in place which is
mutually beneficial and dependent on the efforts of
nearly everyone then one who has accepted the benefits
of the scheme is bound by a duty of fair play to do his
part and not to take advantage of the free benefit by
not cooperating. In other words one has a duty not to
be a free rider.

Rawls suggests that the scheme of social cooperation
is just where everyone has an equal right to the most
extensive liberty compatible with a like liberty for
all and, secondly, inequalities are to everyone's
advantage, and provided by offices which are open to
all. So we may ask two questions: 1. does the duty of
fair play as explicated by Rawls show that we are
required by fairness to continue to obey the laws and
2. is the principle itself acceptable?

Consider the principle itself first. Rawls claims
that since the benefits of social cooperation are the
result of everyone's effort then prior to some

understanding of how they are to be shared they belong to no one. But how does this show that the free rider has no entitlement to the benefits of social cooperation? Firstly, if the benefits of social cooperation are produced by everyone the free rider scenario is ruled out from the beginning. On the other hand if the benefits are produced by nearly everyone the fact that the product belongs to no one does not appear to rule out an agreement that the product should be shared by all. Surely the point is that even in the absence of an agreement, and certain preexisting claims, like ownership of means of production, the obvious conclusion is that the product belongs to all those who have contributed to it. It is this which rules out the free rider's entitlement, not the fact that no one owns the product of social cooperation.

Perhaps it can be shown that, with the principles Rawls gives us for a just constitution or scheme of social cooperation, the free rider is in breach of one of these principles. The first principle requires an equal right to the maximum liberty compatible with a similar liberty for all. Now in commission of an offence, for example, it is not obvious that the free rider has gained a right to a greater liberty than others. The fact that we do x does not show that we have a right to do x. But, it might be argued along the lines suggested earlier by Cooper, if we do not punish

the free rider, the right violated by him does not

exist. But then, so the argument might go, there is an

inequality of right to liberty since the victim didn't

have the right not to be treated as he was by the

offender. Of course if we imagine a society without

punishment then it may be that there are no rights,

which would be consistent with equality. But presumably

we would be required to punish in order to have the

equal right to the maximum liberty. Punishment of

different rights violations secures that that to which

all have an equal right is a maximum liberty. But even

if we were to accept this claim about a necessary

condition of the existence of a right it does not seem

to establish that the free rider has done something

wrong or unfair. For if we do punish then the free

rider has not ┆ diminished someone else's right to liberty .   In

fact he diminishes his own. So we are left with the

problem that, if we don't punish we don't distribute

the right to a maximum liberty and so fail to adhere to

the first principle, while if we do punish, we achieve

an inequality of the right to liberty by taking away

certain rights of the offender. Thus whether or not we

punish, according to the Cooper suggestion, we don't

adhere to the first principle of justice.

So if we don't accept the Cooper suggestion about

rights and punishment we don't get an inequality and

hence we don't show that the free rider has done

something unfair; on the other hand if we do accept the suggestion in a context in which we punish then it seems that we are being unfair to the free rider; it looks as if the free rider is only doing something unfair if we don't punish him and we accept the Cooper idea, but even then the unfairness appears to be one to everyone including the offender. The free rider generates an unfair situation to everyone, himself included, by bringing about a situation in which there are not equal rights to maximum liberty, and this seems odd since it looks as if the free rider is primarily failing in an obligation to others than himself. In any case the Cooper suggestion appears to be false, so we have not shown that the free rider behaves unfairly in view of the first principle.

Nor is it clear that the free rider offends against the second principle. The reason for this is that the second principle concerns inequalities. But now suppose that in accordance with the second and first principles some people have considerably greater wealth than others. Commission of a crime may serve to produce an equality, or more likely a series of crimes committed by the poor against the wealthy might have this result. Now in so far as it has not been shown that mere commission of an offence offends against principle 1 we need ask only here whether it offends against principle 2. But now there is no inequality which has to be of a

certain kind to meet principle 2. So under the circumstances described the free rider or free riders have not done anything unfair.

Let us turn to the second question posed at the outset. Could it be shown on the basis of the supposed duty of fair play that one who committed a crime given Morris's distributional presuppositions would have done something unfair? Suppose then that there is an equal distribution of the benefits derived from spheres of non-interference in the sense outlined earlier, that is, individuals have sets of equal rights to certain things and one of the rights of which such sets are composed is the right to the necessary prerequisites of the enjoyment of such rights and, further, this right is enjoyed by these individuals. As before I shall suppose that for our purposes it makes little difference whether the right to the prerequisites which is enjoyed is a right to the necessaries for the enjoyment of all the rights in the set of rights with which each individual is endowed or just the right to the necessaries for the enjoyment of all rights other than itself. I do this because I think that any regress involved in the former assumption is, if infinite, nonetheless benign. The infinite regress is supposed to occur as follows: suppose we enjoy the right to the necessaries for the enjoyment of all the rights with which we are endowed; but then we have the necessaries

to enjoy all our rights; but one of those rights is the right to the necessaries to enjoy all our rights; hence if we have the necessaries to enjoy all our rights we have the necessaries to enjoy this one, viz. the right to the necessaries to enjoy all our rights; but this means that we have the necessaries to have the necessaries to enjoy all our rights; and one of those rights we may enjoy is the right to the necessaries for the enjoyment of all our rights..,etc..I regard this infinite regress as benign because to have the necessaries for having the necessaries for enjoying something will amount to having the necessaries to enjoy something. No matter how long the chain of provision of necessaries for necessaries is it will amount to no more than provision of what is necessary. So in effect we will have what is necessary to enjoy all our rights, even though we generate an infinite regression when we try to specify that we have what is necessary to enjoy a particular right, namely, the right to the necessaries for the enjoyment of all our rights.

But if we adopt the above interpretation of Morris's equality of benefits what are we to make of the equality of burdens? One possibility is that by this Morris simply means that we all play our part in restraining ourselves; there is an equality of burden in the sense that we all do what is necessary to

conform to certain rules, where this body of rules could be said to be a burden. Alternatively, and this seems more likely, Morris does mean to be understood as requiring a state of affairs in which it is no more difficult for some to do what is required to stay within the rules than for others. Take the first possibility. Now if the duty of fair play is to operate it must be that there exists a just scheme of social cooperation. But clearly Rawls's first principle of justice need not be satisfied by this first interpretation of Morris's equal distribution of benefits and burdens. The first principle requires that the freedom to which all have an equal right is maximal given that it must be compatible with a like freedom for others. Now the problem with our distribution of equal rights is that this distribution need not result in an equal right to such maximal freedom. Freedom may be limited by having too few or too many rights claims. Let me exemplify what I mean by this. Freedom, that is the freedom to which one has a right, may be less than it could be because a fairly parsimonious distribution of rights is made. To recite an earlier case, if there is a right not to be murdered but no other right to protection from physical injury then individuals are not given freedom, in the shape of a sphere of non-interference, resulting from others forgoing interference with people in ways which might lead to

bodily injury. The radius of my sphere of non-interference is smaller because rights do not define a sphere of of complete physical non-interference. Conversely, rights may limit freedom more than is necessary. Imagine a settler in the New World hundreds of miles from the next settler who was acting in accordance with a rule that he only had a right to farm land within one hundred metres of his house. Now the fact that individuals have equal rights doesn't mean that the freedom which is right-protected is maximal; it may fall short of this in either of the directions outlined above.

Nor is it clear that the Morris distribution on the first interpretation will be in conformity with Rawls's second principle either. The prerequisites to the enjoyment of the rights which have been distributed which are guaranteed as an initial assumption of the distribution may well not establish for their possessor anything like economic well-being. Certainly the distribution of rights may be very parsimonious-there may only be a right not to be killed intentionally by another. In this case where equal rights are only to one right, it does not appear that the assumption of the necessaries to the enjoyment of this right guarantee any economic well-being at all. Clearly the distribution of wealth in a society with so few rights may infringe the principle that inequalities of wealth

be to the advantage of all. Indeed one may envisage
scenarios in which inequalities of wealth consistent
with an equal distribution of rights are to no one's
advantage.

However things look more promising on the second
interpretation of Morris's initial situation. For if we
are required to ensure that it is not more difficult
for some to abide by the social rules than for others
certain kinds of economic inequality at least might be
thought to be eradicated. For example if it be supposed
that where one is starving or deprived of certain
necessities of existence because one cannot afford them
one is much harder pressed not to steal from one's
neighbour than if one had the money to buy such things
then it could be argued that equality of burden in the
scheme of social cooperation would rule out the
possibility of some people being in such a
disadvantaged situation. But of course this is false,
for the equality of burden of itself seems to have no
significance for the level of economic well-being of
those who bear it. Such equality is consistent with
everyone being very badly off indeed. But of course
this misses the point of the second principle. This
principle merely requires that, if there are
inequalities of wealth, for example, they should be to
the advantage of everyone. One way of interpreting this
is as the requirement that the inequality makes the

worst off better off than he would be under any equal
distribution which is feasible. However the principle
becomes more explicit in later writings of Rawls and we
may interpret the principle as requiring instead that
if there is inequality it should be such that the worst
off representative group is better off under it than he
would be under any other distribution. But then of
course it seems clear that we might effect equality of
burden a la Morris by means which did not have the
consequence that this rather stringent principle of
Rawls, understood narrowly here as an economic
constraint, would   be met. This conclusion seems
forced on us by the observation that many things other
than our economic status appear to have an effect on
how hard it is for us to obey the rules. Those of
abnormal psychology, for example, may find some rules
almost impossible to obey. Thus effecting equality of
burden might be achievable by educative means which
left such matters as that of the economic position of
the worst off untouched. Secondly, the presence of an
equality of burden does not seem to entail that the
first principle is met either. That it is, roughly
speaking, equally hard for everyone to obey the rules
does not seem to have much bearing on how comprehensive
the liberty to which individuals have a right is.
Indeed, given the number of factors which appear to be
relevant to how hard it is for someone to obey a

particular law, the process by which equality of burden is to be achieved suggests the necessity of a quite far-reaching invasion of individual liberties in order to 'iron out' the sort of differences between individuals which might stand in the way of such equalization of burden. On the second interpretation of what Morris means by equality of burden it is plausible that liberty to which one has a right will be far from maximal.

So the principle of fair play as Rawls conceives it appears neither entirely satisfactory in isolation from the problem of Morris's distribution of benefits and burdens, nor does it appear that the sort of situation Morris envisages is such that this principle can be brought to bear to show that one who breaks the rules and so gets the benefits of the scheme of cooperation without making the sacrifices which others have made is acting unfairly. So if it is true that it is only fair to have a mechanism for minimizing the likelihood of the occurrence of a deviation from an equal distribution of benefits and burdens we have not shown that such a mechanism is required by fairness as a means of reducing the likelihood of unfairness.

Another line of enquiry would concentrate on the ideas that it is only fair that I benefit from some scheme if I have contributed to it and that my contribution should proportionally determine my share

of the benefit. Thus if I am contributing equally to the scheme of social cooperation, where this contribution is construed narrowly as the generation of equal spheres of non-interference, because I share equally in the burden of that generation I deserve to get out of it a share of the spheres of non-interference equal with that of others. With this simple model in mind it becomes clear why Morris wants the initial assumption of an equal distribution of burdens of self-restraint. If it is not equally burdensome for individuals to generate the spheres of non-interference then those for whom it was more burdensome to obey the rules out to receive a larger share of the good in question, namely the spheres of non-interference. But how is this to be thought of? One way in which a sphere of non-interference can be larger than another would be if I could do certain things to others which they could not do to me. Perhaps I would be legally permitted every second Thursday of the month to go out and hit someone. Any retaliation would be unlawful, an infringement of my rights and intrusion in my sphere of non-interference. This would have the odd consequence that those who were naturally law-abiding would come off worst. Indeed it would constitute a means by which the thugs could legally have their way. Suppose that I wish to hit my neighbour because I don't like the look of him, or just because I am an

aggressive person. Now provided he doesn't entertain similar urges vis-a-vis me or someone else I may be able to claim, on the basis of the difficulty I have forgoing the opportunity to hit my neighbour, that I am entitled to a greater sphere of non-interference than my innocuous neighbour and so legitimize my hitting him. Of course, given that proportionality must be maintained relative to all individuals I might not be able to get my way. Considerations of proportionality governing relations between me and my neighbour, my neighbour and others, me and others, or others and others, would have a bearing on size of spheres of non-interference which could be distributed. Moreover how size of sphere and contribution in the form of burden are to be calculated are doubtless fraught with epistemological difficulties. But the point is that even these rather informal remarks about the consequences of applying the fairness model above, where contribution proportionally determines distribution, either seems to suggest that it is the wrong model or to point out the infelicity of making too much of the size of the burden constituted by obeying laws. Even if the model is wrong it does seem that introducing any idea about the size of burdens of obedience to laws not known to be unjust runs the risk of generating hyperbole about possible entitlements of those of us who really would like to break the law.

Alternatively we may restrict ourselves to the special case in which benefits and burdens are equally distributed, overlooking difficulties generated by an application of the fairness-in-contribution model (as we might refer to the above model of fairness,) and then ask whether it can be shown that it is unfair that some individual break the law on such a model. Is this so?

The idea must be that if someone breaks a law his receipt of the product of social cooperation is in some way disproportionate to his contribution. Now presumably his commission of a crime shows that he is not contributing equally, on the assumption that others are obeying the laws, and yet he is receiving the same proportion of the social product in the form of spheres of non-interference which others receive. Yet this is unfair and so it appears that the fairness-in-contribution model shows that the criminal behaves unfairly.

Two immediate problems seem to face us: the first based on a consistent application of the fairness model itself, the second arising from anxieties about the scope of the model.

The first problem arises if we start to apply the model to the distribution of things other than spheres of non-interference. Suppose in fact that although distribution is as per Morris's initial distribution

situation   wealth is very disproportionately
distributed. Thus some individual may well be treated
unfairly in respect of his income. What are we to say
if this individual commits a crime, especially if it is
a theft of money from one who receives more by the
distribution of wealth than is merited by his
contribution? But if, as seems reasonable, a situation
in which Morris's distribution was effected alongside a
disproportionate distribution of wealth was regarded as
unfair, if we are to accept F1 as plausible but also
believe in a redistribution of wealth to coincide with
the requirements of the fairness model, we appear to be
in difficulties since any redistribution of this kind
would involve interference with people's supposed
spheres of non-interference. And then we have lost
equality of benefits and burdens and so are not really
applying F1 anyway. One answer to this is that we are
not creating an inequality of spheres of non-
interference by redistributing wealth. If a rule is
introduced that everyone's income should come into line
with a certain principle then this rule in applying to
all affects equally everyone's sphere of non-
interference. Equality has been retained, while spheres
of non-interference have been redefined. So even if the
initial Morris situation is part of a larger situation
which by the fairness model is unfair we are apparently

not forced to abandon the requirement that equality of benefits and burdens a la Morris be maintained.

The other problem concerns the scope of the principle. Contribution is not always the measure of what an individual ought to receive. If someone falls sick and can no longer contribute to a country's GNP we feel required to continue to allow this individual some amount of the GNP, even though his proportional distributed share should, according to the model, be zero. However if someone chooses not to contribute any more the attitude is somewhat different; under some circumstances we feel such an individual should receive nothing. I refer to these as matters of scope. Other principles and considerations are brought into play in certain contexts in which we might try to apply the model of fairness-in-contribution. But if in some cases the principle is to be overruled we need some sort of account which shows why in general this may happen and an explanation of why it is that the criminal is not to be one who is to be kept at a level of equality in respect of the share of the social product which he receives. This is not to rule out the principle of fairness-in-contribution as an explanation of why the criminal does something unfair but rather to suggest that if the initial distribution proposed by Morris is to yield a model explaining why criminal behaviour is unfair then a general argument must be offered in which

a principle or principles more general in application than fairness-in-contribution must at least be adumbrated.

Of course even with this F1 is still open to the initial criticism that it is unlikely that fairness would require the implementation of quite literally any mechanism which minimized the likelihood of deviations from the original distribution of equality. It seems reasonable to suppose that fairness itself will impose constraints on which mechanisms may be used but mere maintenance of the distribution of equal benefits and burdens does not look promising in itself as the principle by which we may limit the range of satisfactory mechanisms. This is partly because the equality postulated is not equality at any given level, in which case a mechanism which maintained equality but reduced people to slaves would still not be ruled out by this kind of account. It seems in any case that a more general account of fairness than merely fairness-in-contribution is required to set limits to F1 in such a way as to render it plausible.

What are we to make of F2? (F2 is the claim that punishment erases the unfair advantage of one who voluntarily renounces the burden of self-restraint yet continues to enjoy all the benefits accruing from the exercize of self-restraint by others.)

Let us consider F2 in the light of two interpretations of punishment, one on which punishment is retributive in our sense, the other on which punishment is just some state mechanism for depriving individuals of certain rights.

So once again we are to imagine the backdrop of an equal distribution of benefits and burdens, where this is to be understood as a distribution of equal rights yielding equal spheres of non-interference and either equal burdens of self-restraint in obeying laws or simply a distribution in which each individual obeys the laws whether or not this calls for equal self-restraint on the part of each individual. It is to be borne in mind then that we have two interpretations of 'punishment' and two interpretations of the initial situation.

What then is the unfair advantage which punishment is to erase? On either interpretation of the original situation the criminal renounces the burden of obedience to the rules, or at least to some rules, and continues to enjoy the same sphere of non-interference as the others. So we might say that the burdens are no longer equal between the criminal and others while the benefits are equal. In order to erase the unfair advantage it is reasonable to suppose that we should give the criminal some extra burden while keeping the benefits the same as before. If we can quantify burdens

what we need is to compel the criminal to bear the

burden which others bear plus a burden equivalent to

that which he put down by commission of his crime. But

this strategy will not work. If we are to make the

criminal's obedience to the rules more onerous we might

try to do this by giving him more rules to obey where

the result is a restriction of his sphere of non-

interference. But if we do this then it is clear that

the benefits which he enjoys, where these are assessed

by reference to his sphere of non-interference will no

longer be equal with others. So what we will be doing

is reducing the benefits and increasing the burdens he

experiences, not keeping the former constant while

increasing the latter. Since we are understanding equal

benefits of spheres of non-interference in this context

as equal right-defined spheres of non-interference any

deprivation of rights aimed at increasing the

criminal's burden will move him from a position of

equality with regard to the benefits from the system.

Perhaps we could increase the burdens of the criminal

in some other way. We might do this by increasing his

incentive to break the laws in such a way that the

burden of self-restraint involved in his obedience to

rules would be greater. But whether we assume that

obedience to rules involves an equal burden or not, the

problem still seems to be that in all likelihood any

interference with a criminal's situation sufficient to

alter his motivation in certain respects will equally

constitute a change in the sphere of non-interference

which he enjoys and so, once more, we have altered the

base equality of benefits we wanted to hold constant.

So on either interpretation of the initial situation

the increase of a burden required to erase an unfair

advantage does not look possible if we are to keep

benefits constant. But perhaps we need not keep the

benefits constant. If when we erase the unfair

advantage attendant upon commission of a crime we

inevitably introduce an inequality into the benefits

received by the criminal we might try to increase the

criminal's burden and take into account that for the

duration of that increase there will be a corresponding

decrease in the benefits which the criminal receives.

But now we seem to face a problem of quantification of

benefits and burdens. If commission of a crime gives

the criminal an advantage in the form of his having

less of a burden than others we want to put him in a

situation in which he has something equivalent to 1.

the equal distribution of benefits and burdens of the

original situation plus 2. something equivalent to the

burden put down. Since, as we have seen, it is

reasonable to suppose that any interference with the

criminal's situation in order to give him 2 will have

an effect on 1, we are forced to try to establish for

the criminal a situation for him which is equivalent to

1 plus 2 but which does not involve his being in the situation described by 1. But this means we have to be able to quantify the benefits and burdens situation of an individual as a whole so that we can establish what situation is the equivalent of 1 plus 2 where neither the benefits and burdens of the criminal will be equal to the initial distribution.

Leaving this problem behind us it seems reasonable to suppose that the criminal's situation must be changed so that he has some increase in burden. If the criminal breaks a particular law, and he does this voluntarily, we may suppose that he puts down a burden of self-restraint in respect of this law which others are carrying. Now on the first interpretation of the original situation we are supposing that the burden of self-restraint in obeying the law is equal for individuals, even if the facts of the matter appear to be different. On the first interpretation of the initial situation a special significance is attributed to the idea of a burden of self-restraint, one different from the usual idea in which, in the absence of evidence to the contrary, it appears reasonable to suppose that the burden varies from individual to individual. But in that case we seem to have no way of knowing what the burden of self-restraint which is put down by the criminal is. We cannot take into account considerations which it is reasonable to suppose may

have some effect on how difficult it is for the criminal to obey the law he has broken. These matters are by assumption irrelevant. Nor does it seem to be entailed by the fiat that the burden of obeying the law in toto is equal for all citizens that the burden for obeying any particular law must be equal for each citizen, though this is perhaps the most obvious line to take. But the point is that if we are not to look into certain matters in order to assess the burden, except perhaps to assume that it is the same for all in respect of a particular law, how are we to assess how great is the unfair advantage which is gained by commission of the crime. It seems but a short step from the assumption that the burden for each of obeying a particular law is the same to the assumption that the burden is equal whatever the crime committed. But then we would have to have in view a justification of punishment qua eraser of unfair advantage which required the same imposition of extra burden, or rather its equivalent, for every offence. Clearly this cannot be the means to a justification of retributive punishment.

Alternatively we may suppose on the second interpretation of the original situation that although things have been contrived to establish an equality of burden, this has been effected with regard to a kind of self-restraint which is sensitive to features of the

world in the way that our normal sense of self-restraint is. Thus it is reasonable to suppose that we are operating with a sense of self-restraint such that it is more difficult, other things being equal, for a hungry man to restrain himself from stealing a loaf of bread than it is for a sated one.

Now on the second interpretation of the initial situation the burden put down by the commission of a crime depends on how hard it would have been for me to obey the law broken. But then clearly if we are to justify retributive punishment we must suppose that the more morally serious our rule infraction the greater the burden which we are putting down. Only in this way will erasure of unfair advantage be effected by more serious, i.e. onerous punishment, the more morally serious our rule infraction. But the problem is that there is really no reason to believe that the average individual, let alone all individuals, will find it more difficult to refrain from murder than minor tax evasion. Of course it might be claimed that we are unaware of any difficulty involved in refraining from more serious crimes against the person because of unconscious acculturation and indoctrination processes. But it is not clear that this argues for greater self-restraint being involved, unless by self-restraint we understand something which we may be unconsciously engaged in. But even if we accept this idea it is hard

to resist the thought that we are measuring self-restraint by moral seriousness of crime and not vice versa. Where self-restraint may be largely unconscious we no longer have even a tolerably clear idea of how to measure it.

So certainly we could not use the Morris model to justify retributive punishment. Nor, given the difficulties of quantifying benefits and burdens does it appear to generate a workable model for determining punishments in the form of deprivations of rights where these need not satisfy the criteria for retributions. However it has been proposed by others that we can modify the model which Morris uses and get something like a justification of a punitive process which is retributive.

Sher (Desert. 1987; pp. 69-90,) retains as the justification of punishment the idea that it equalizes a distribution of benefits and burdens. For him the benefits are those which accrue as a result of the moral restraints on the behaviour of others while the burdens are once more the exercize of restraint such that one's behaviour accords with those moral restraints. However when an individual does wrong he gains an extra benefit in terms of freedom from the (or some) moral constraints such that the greater the moral constraint which is evaded the greater the freedom obtained. Once this has occurred

'the natural way to restore a fair balance is to reduce the protection he ordinarily would have gained through moral restraints on the conduct of others. By treating the wrongdoer in what is ordinarily a forbidden way, we strip away part of the protection that moral restraints on our behaviour would ordinarily have afforded him. Thus we remove precisely the sort of advantage he has gained. Because the resulting disadvantage can be assessed in terms of its usual moral wrongness, it can be weighed on the same scale as the wrongdoer's unfair advantage.'

This is a difficult passage to understand. The idea seems to be that the extra benefit in the form of extra freedom is to be annulled by decreasing the amount of freedom of the wrongdoer. We do this by treating him in a way which we would normally be morally prohibited from doing. Moreover we can ensure that the wrongdoer loses the same freedom as he gained by assessing the amount of extra freedom the punisher gets by the act of punishment if we suppose him to be acting with the normal moral prohibitions in place, i.e. as if the punishee had done nothing wrong. Where the freedom of the punisher is equal to the freedom which was gained by the punishee then the freedom lost by the wrongdoer as a result of punishment is equal to the freedom he gained as a result of doing the wrong for which he is punished.

The first piece of the argument we need to look at

more closely is the idea that the greater the moral

wrong which we commit the greater the freedom from

moral restraint we obtain. Sher claims that

1.'a person who acts wrongly does gain a significant

measure of extra liberty: what he gains is freedom from

the demands of the prohibition he violates.'

Moreover,

2.'as the strength of the prohibition increases, so

too does the freedom from it which its violation

entails.'

Clearly in one sense 1 is false. The demands made

upon an individual by moral rules are the requirements

of that individual stipulated by the rules. But a

person who acts wrongly is not free from the

requirements; he simply does not do what is required of

him by one of the rules. So what exactly is the

wrongdoer free from? He might of course be free from

some compulsion to obey the rule which others

experience. But perhaps some who obey the rule which

the wrongdoer violates feel no such compulsion. But

what we are seeking is a freedom special to the rule

breaker and so freedom from such a compulsion is not

that freedom. Perhaps the point is that the wrongdoer

is free from the constraint of the prohibition while

those who obey it are not. But if constraint is not to

be interpreted as a requirement how are we to

understand it? A constraint may be something which forces us to obey and the wrongdoer will obviously be free from this. But how do we know that others aren't free from this? Clearly we can't infer from the fact that X doesn't break a rule that he was forced to obey it any more than we can infer from the fact that I eat an ice cream that I was forced to eat it. So again it is not clear that the freedom in question is special to the rule breaker. The same applies if we suppose that people obey moral or legal rules as a result of indoctrination. Now we can't infer from the failure of X to obey a rule that he hasn't been indoctrinated. At best we can infer that he hasn't (successfully) been indoctrinated to obey the rule he breaks. But then the same problem arises: how do we know that others who obey the rule aren't equally free from indoctrination to obey it? Some may obey it because they see the sense of it. Of course in the final analysis we could say that the wrongdoer is free from obedience to the rule he violates. But the word 'free' is odd here if it means nothing more than that the wrongdoer does not obey. I may not go for a walk but I am not as a result free from going for a walk.

Suppose we provide analyses of the phrases 'free to' and 'free from' and define phrases in which the latter occurs in terms of phrases in which the former occurs. Suppose A is free to x where P, which would normally

prevent **A** from **x**-ing, is absent. Then we may analyse **A** is free from **y** as **A** is free to do something **z** in the sense that a normal preventive to **A**'s doing **z**, namely **y**, is absent (though other preventives to **A**'s doing **z** may nonetheless be operant.)

But now suppose we make the claim that where **A** performs a more morally wrong action than **B** then **A** is more free from moral prohibitions than **B**. This opens up two difficulties: firstly, we need some way of analysing degrees of freedom from something; secondly, if we analyse '**A** is free from moral prohibitions' as '**A** is free to **x** in the sense that a normal preventive to **A**'s **x**-ing (where the normal preventive takes the form of moral prohibitions) is absent' then we still don't have much of an analysis until we know what to substitute for '**x**'. Now if we suppose that a possible substitution for '**x**' is 'doing what he wants' and that degrees of freedom may be analysed in terms of degrees to which a normal preventive is effective then we have something like the following: **A** is more free from moral prohibitions than **B** where **A** is more free to do as he wants than **B** in the sense that a normal preventive to **A** and **B** doing what they want, call this **P**, is less effectively operant on **A** than on **B**. If we cut this down slightly what we have is that **A** is more free from moral prohibitions than **B** if moral prohibitions are less preventive of **A** doing what he wants than they are of **B**

doing what he wants. Yet if we use this analysis it is quite clear that if **A** infringes a more grave moral prohibition than **B** this will not entail that **A** gains more freedom than **B** since it may nonetheless just not be the case that moral prohibitions interfere less with **A** doing what he wants than **B**. In fact doing less of what is morally prohibited may be an indication that moral prohibitions interfere less with what we want. So the proposed analysis fails to show that 2 is true and rather shows it to be false. So either it is not clear what this freedom from the demands of the moral prohibition means or the claim takes a sense which makes 2 false. But then in this respect Morris's account appears to be superior in that a clear reason can be given why there is more or less freedom upon commission of some wrongs or rule infractions than upon commission of others; the stronger the inclination the greater the burden of which one is free by giving way to the inclination.

Thirdly it is difficult to see why the fact that the punisher would derive the same freedom through punishment, if such punishment were morally prohibited, is an indication that the freedom the punishee loses as a result of being punished is equal to the extra freedom he (the punishee) gained. If we suppose this extra freedom is freedom from the moral prohibition violated then removal of this freedom would seem to

require an increase in the moral prohibitions which the violator has to abide by. Such prohibitions would concern performance of actions which others were allowed to perform. But to remove the freedom so that the wrongdoer has equal freedom with the others would require the capacity to quantify freedom. But there seems to be no way of doing this. Even if we were able to show that there was more freedom resulting from a great than from a little wrongdoing the subtraction of this extra freedom would require the capacity to determine something completely different, the amount of freedom lost by others as a result of being the targets of certain normally prohibited actions. But Sher offers no account of how we might do this. But it is this kind of calculation rather than just the calculation of what extra freedom the punisher might be getting under certain theoretical circumstances which is needed.

So although Sher's account offers promise of adapting the Morris account to demonstrate the fairness of retributive punishment it does not fulfil that promise. It is by linking the extra benefit obtained through wrongdoing to freedom which increases as the moral wrongness of the action increases that it is made a requirement of punishment that it increase according to the moral wrongness for which it is meted out. But this extra freedom, and how the right amount is to be

removed, and how punishment achieves this, is left
unexplained.

In conclusion it would seem that one important
deficiency with the idea of what fairness requires
given equality of distribution of Morris's benefits and
burdens is that this distribution takes no account of
how other things, e.g. economic well-being, are
distributed. At least Rawls sees that the duty of fair
play has to operate against the backdrop of a scheme of
just cooperation where the justice of such cooperation
is determined by factors other than difficulty in
obeying legal rules and how spheres of non-interference
are distributed. Moreover, the absence of a fuller
account of what fairness requires in the way of the
description of the initial situation is responsible for
the absence of constraints in the expression of F1 on
what mechanisms of social control are permissible. We
need to know what constraints fairness puts on the
mechanism for maintaining the distribution described at
the outset. However in so far as it is punishment that
Morris has in mind as this mechanism the deficiency is
less serious than it might be. In so far as Morris is
proposing a model only we might build into the
description of the initial situation whatever is
necessary to make it just but then of course it would
be clear that mechanisms other than punishment would be
necessary to maintain a just state of affairs. But this

might not be a problem, since it may be that given that other mechanisms were in place, punishment of a certain kind was the right mechanism for maintaining the distribution of the sort of benefits and burdens Morris has in mind. With this in view I do not wish to be too pessimistic with regard to F1. In fact it is F2 which seems to run into the more serious difficulties in which case it might be thought that what is wrong with the Morris account is the detail. This is why it is important to remain cautious in an assessment of the general model Morris proposes. Notwithstanding this, both the Morris and Sher accounts fail to argue for what I have called retributive punishment as the appropriate mechanism for maintaining a fair distribution of certain things, though this does not invalidate F2 where punishment is not construed as retributive punishment. Rather if we are dealing with this alternative interpretation of punishment the difficulties with F2 concern just how we are to assess the advantage gained by one who commits a crime and how this advantage is to be translated into something equivalent to what is lost by punishment.

(2) Retributive punishment and treatment as equals.

Let us turn now to an extremely interesting account produced by Reiman in his paper 'Justice, Civilization and the Death Penalty.' Reiman seeks to show, among other things, that the lex talionis is just. He makes a number of claims in support of this.

1. The lex talionis bears certain affinities to the Golden Rule (GR,) viz. 'Do unto others as you would have others do unto you.' Now 'if people were treated as they treated others [that is, if they were treated in accordance with the lex talionis,] then everyone would necessarily follow the Golden Rule (call this conditional GRL) because then people could only willingly act toward others as they were willing to have others act toward them.' But then since to follow the Golden Rule is to treat others as equals this means that if we employ the lex talionis in punishment then people will treat one another as equals.

The lex talionis, it will be remembered, is 'an eye for an eye, a tooth for a tooth, a life for a life.' But this means that if **A** injures **B** in a certain way in punishing **A** we must injure him in the same way he has injured **B**. Of course the standard objection to this runs that in some cases it will not be possible to do to **A** what **A** has done to **B**. It is contended that, for

example, if **A** rapes **B** it will not be possible to rape **A** (noting here that the same injury does not include a specification of who performs the injury.) But this is surely false. A man can be raped. Alternatively it might be claimed that if **A** embezzles funds from an institution **B** then the same action cannot be performed on **A**, for **A** is an individual, not an institution. However the lex applies to injury to sentient individuals, perhaps only human ones, and so it is not obvious that the embezzlement example counts as a case within the field of application of the principle in which the lex could not be applied; rather it is a case to which the principle is inapplicable.

A further problem with the lex is that it appears to apply to injuries rather than pains. Thus if **A** causes **B** pain but does him no injury (it being a necessary condition of injury that **B** is damaged in some way) then the lex would not require any duplicate pain to be administered to **A**.

This second difficulty at least shows that the conditional GRL, the antecedent of which is 'if people were treated as they treated others,' has an antecedent which is not the equivalent of 'if people were treated in accordance with the lex talionis.' Despite this I think we are justified in trying to cleave to the spirit of the lex talionis in which case we would interpret it as requiring that if **A** causes **B** to

experience a certain amount of pain then **A** must also be caused to experience equal pain.

The first problem with the lex - that it is inapplicable to cases like embezzlement - arises because institutions are not sentient individuals with bodies. But then all kinds of immoral actions will not involve injury or pain to sentient individuals. For example, we can imagine the unmourned dead Smith being dug up and mutilated by an evil exhumer. The deceased is neither injured nor pained. But then no response is required by way of the lex. Nor is it in keeping surely with the spirit of the lex that we do to the exhumer what he did to the deceased if we imagine the deceased to have been alive and to have experienced the injury and pain of the mutilation.

This sort of problem with the evil exhumer shows that even if we were to extend the lex so that instead of causing the same injury or pain it required equivalent injury or pain to be administered it still would not require any punishment of the exhumer. But then the lex does not require a punitive response to all immoral actions. Secondly in the case of the embezzler it is not clear what is meant by the injury to an institution. This may require us to calculate to what extent the institution has been damaged or to what extent the members of the institution have been injured or both.

Even if we content ourselves with the fact that in most cases of immoral action the lex will dictate a clear response involving the causing of pain or injury other difficulties remain. As it stands it is not clear what the lex requires in a case where **A** accidentally injures **B**. Does this mean that we are to cause **B** equivalent injury or that we are to cause him equivalent injury accidentally? And if we opt for the former how is the lex to account for the feeling that justice requires some punishment of the individual who accidentally kills a pedestrian while driving a car completely intoxicated, though not the punishment of death?

So the following problems are outstanding : i) what does the lex require in cases of moral wrong where no injury or pain is caused; ii) what does it require where the wrong is the result of damage to a non-sentient individual, even if some sentient individuals are pained or injured as a result? iii) what does it require where pain or injury is caused but not as a result of a morally wrong action?

Problem (iii) could be dealt with as follows: the lex dictates the just punishment for harms which are caused as the result of wrong actions. But this means that if **X** kills **Y** and fully intended to do so while **A** kills **B** but only meant to scare **B** it is just to punish them by administering the same pain or injury. So this way of

dealing with (iii,) though providing clear guidance,
falls far short of being morally satisfactory even
though it does exclude from the scope of the lex acts
which are not morally wrong.

Leaving these unsolved difficulties about the lex
behind we will think of it for the moment as requiring
for an immoral action a response causing equivalent
pain or injury.

Let us turn now to the Golden Rule. A certain
ambiguity arises in connexion with the meaning of the
phrase 'as you would have others do unto you'. This
could mean ' as you would like others to treat you' or
'as you regard it right for others to treat you'. Call
these interpretations 1 and 2 respectively. Suppose
then that the lex talionis is employed to determine how
people are to be punished. Does it follow then that
people are treated as they treated others?

It still appears that employment of the lex does not
guarantee that people are treated as they treated
others. For example, as we are understanding the lex at
the moment if A tries to scare B and accidentally
shoots him he will be punished by being killed; but
this is surely not to treat A as A has treated B. So if
people are treated in accordance with the lex this
still does not entail the antecedent of the conditional
GRL, i.e. that we are treating people as they have
treated others. Suppose instead that we do in fact

employ some principle which requires us to treat individuals as they treat others. Does it follow from this that everyone would necessarily follow the Golden Rule?

Firstly it is not clear what is involved in treating someone as he treats others. Do we treat the gangster as he treats his mother or as he treats his murder victims? So there is no clear line on what it is to treat someone as they treat others. Assume then that what is required is that we treat **A** as he has treated **B** if we are to treat **A** as he has treated others. Then clearly on the first interpretation of the Golden Rule it does not follow from the fact that we treat **A** as he has treated **B** that **A** is following the Golden Rule. Even if we suppose that **A** lives in a society in which all acts such as his are punished he might nonetheless not want to be punished, that is, treated as he has treated **B**. But in that case he does not treat others as he would like to be treated. But neither is it entailed by our treating **A** as he has treated **B** that **A** is following the rule on the second interpretation since **A** may fail to regard it as right that he be punished in the way that he has acted towards **B**. He may feel that he is a special case.

A different interpretation of the Golden Rule seems to be called for. Perhaps it means something like 'what you do to others you authorize their doing to you.' But

even on this interpretation it is still clear that **A** may fail to authorize his being treated like **B**.

Of course it might be objected that in a society in which all wrongdoers are punished such that they are treated as they have treated others then in choosing to kill **B A** has chosen to be killed. And since he has chosen this then he wanted this outcome. Therefore if he acts as he does in a society where people are treated as they treat others then **A** indeed wants to be treated in the punitive manner he is and so is following the Golden Rule on its first interpretation. Yet the assumption here is that if **A** chooses a certain outcome then he wants that outcome. This requires disambiguation. If at the point of a gun I have the option of handing over some money or having my brains blown out there is a sense in which I will want to hand over the money. Here 'I want a certain outcome' means that, of the available options, I prefer a certain option-outcome. Thus the first sense of 'I want **x**' is: of the available options which include **x**, I prefer **x**.' But clearly there is a further sense in which I may want neither to hand over the money nor to have my brains blown out. I may want to do neither. Call the option of doing neither **z**. Now **z** is a possible option, in so far as it might have been an option with which I was presented by the gunman, but it is not one with which, so we are supposing, I was presented. So in our

present terminology I shall regard z as a possible rather than an available option. So 'I want to x' may mean in its second sense 'of all possible options, I prefer x.'

Now it would seem that A has as available options being treated in any way in which he can treat others (where 'can' implies an available option.) But then the option he chooses will be the option he wants in the first sense given for 'I want x'. We may not infer from A's choice of some option O of those available that O was an option A wanted in the second sense of 'I want x', that is, that A preferred O out of all possible options. But then the strength of the inference from something like 'A lives in a society in which people are treated as they treat others', 'A treats B in manner m', and 'A is rational' of the conclusion 'A wants to be treated in manner m' is determined by which sense of 'wants' or 'would like' is appropriate for the verb in the principle GR. But it is surely plausible that in GR 'as you would like to be treated' means 'as you would prefer, given the range of possible outcomes, to be treated.' But if this is the way we are to interpret GR then the above inference is invalid. So in one sense of the Golden Rule if we treat people as they treat others then it doesn't follow that they are acting in accordance with it.

Again it might be argued that in killing B he, A, chooses the death of B and A over the survival of B and A. Hence what A wants is equal treatment for himself and B and both their deaths. But again this may be a want in the first sense of 'wanting' above. If we were to infer from this that A wanted to be treated equally with B it would only be as an 'available option' want of A's.

A final interpretation of the Golden Rule might be put as follows: the way an individual treats others it is right for him to be treated by those others. This is not the lex talionis. We can see this by remarking that interpreted in this way the Golden Rule has greater scope than the lex talionis, for while the latter prescribes certain conduct in the event of an individual acting immorally, the former prescribes conduct under these circumstances and circumstances in which an individual acts virtuously. Secondly as it has been interpreted the lex requires us to treat people who have performed morally wrong actions as they have treated others where like treatment is determined by injury or pain caused. The Golden Rule is more general than this since likeness of treatment could be construed other than as likeness or equivalence of injury or pain.

But if people are treated in accordance with the lex talionis as understood so far (that is, those who do

moral wrong are caused pain or injury equivalent to
that which is caused by the immoral action,) and we
assume that this principle is one of treating
individuals as they have treated others may we infer
that everyone follows the Golden Rule? Clearly not. The
fact that we treat others who have wronged in
accordance with the lex does not mean that everyone
follows the rule that it is right that one be treated
as one has treated others. Suppose again that A kills B
and is punished in accordance with the lex. At best
this would show that the punisher follows the Golden
Rule. But we cannot infer from the fact that A is
punished in accordance with the lex that A follows the
Golden Rule. The Golden Rule is now a moral precept the
following of which is evidenced by a conforming of
one's behaviour to it, that is, by treating people as
they treat others (including us.) But we may imagine
that B has not treated A as A has treated B. If he had
then the punishment of A would be the punishment of one
who had administered a just punishment for the
administering of that punishment - which would be
absurd. The whole point about punishing A in accordance
with the lex talionis is that A has not treated another
as he has been treated. So far from entailing the
Golden Rule punishment in accordance with the lex
surely entails that at least one person has failed to
follow the Golden Rule.

So it is far from clear that 1 is true.

2. Action in accordance with the Golden Rule implies treating others as your equals. ''Doing to another what you would not have that other do to you violates the equality of persons by asserting a right toward the other that that other does not possess toward you. Doing back to you what you did 'annuls' your violation by reasserting that the other has the same right toward you that you assert toward him.'' In this way punishment in accordance with the lex talionis restores him to equality with you.

I shall comment on two points here. Firstly, why is it that by violating the Golden Rule we assert a right? Secondly, why is it that doing the same back to the wrongdoer asserts the same right toward him?

According to Reiman acting according to GR implies treating other people as equals. Suppose we classify the different interpretations of GR as follows: GR1- treat others as you want to be treated; GR2- treat others as you regard it as right for others to treat you; GR3- treat others as you authorize others to treat you. But we still need an interpretation of what it is to treat others as equals. Consider the following interpretations: E1- to treat others the same (as oneself;) E2- to treat others with equal concern and respect; E3- to treat people as possessors of equal rights and with respect for those rights.

Now suppose one acts in accordance with GR1. But the fact that I treat others as I want to be treated doesn't mean that I treat other people the same as myself. Obviously I may fail to treat myself as I want to be treated and so I don't treat all people the same - rather I treat all except myself the same. Alternatively, where E1 is construed as the requirement that one treat others the same, but not necessarily the same as one treats oneself, then the fact that I treat others as I would like to be treated still doesn't mean that I treat others than myself the same. On Monday I may treat **X** abominably, on Tuesday I may treat **Y** with courtesy and sympathy. This shows, at least in one sense, that I needn't be treating others the same. And yet it is quite consistent with these different styles of treatment I adopt vis-a-vis others that I am nonetheless treating others as I wish to be treated. So the fact that I act in accordance with GR1 doesn't mean that I am acting in accordance with E1. Likewise if I can treat **X** and **Y** so differently while acting in accordance with GR1 it is clear that I needn't be acting in accordance with E2 either; under the envisaged circumstances it doesn't seem that I would be treating **X** and **Y** with equal concern and respect, at least, on one reasonable interpretation of what is meant by treatment with equal concern and respect. Lastly, if I act in accordance with GR1 my appalling

treatment of **X** on Monday may involve failure to respect
rights of his which I respect in **Y** on Tuesday (and
don't evince any lack of respect for on Monday either.)
So action in accordance with GR1 doesn't entail action
in accordance with E3.

Secondly, my action in accordance with GR2, if it
issues in the different treatment of **X** and **Y**, clearly
doesn't entail that I am acting in accordance with E1
on either of its interpretations; that it doesn't
entail my treating **X** and **Y** the same obviously entails
that it doesn't entail my treating **X,Y** and myself the
same. Also that I regard it as right to treat **X** and **Y**
so differently again seems to show that I am not acting
in accordance with E2. Thirdly, it can't show that I am
acting in accordance with E3. If I treat **X** abominably
and so regard it as right for him to treat me like
this, while I show that I regard it as only right for **Y**
to treat me well it is hardly obvious that I am
treating them as possessors of equal rights vis-a-vis
me, for one.

If I act in accordance with GR3, that is I treat
others as I authorize them to treat me, my different
treatment of **X** and **Y** might be thought to entail that I
authorize **X** to treat me badly and **Y** to treat me well.
But if we assume, for the sake of argument, that I can
act in accordance with GR3 at all, i.e. that I have the
legal power to confer legal rights on others by my

actions where there is no limit on the kind of legal

rights I can confer on others, then it is clear that I

don't treat **X** and **Y** the same, so I don't follow E1.

Alternatively, if it is moral authority which I am

supposed to confer on others in their treatment of me

then again it is clear that I don't treat **X** and **Y** the

same for I authorize their different treatments of me.

Now if I act in accordance with GR3 does it follow that

I am treating **X** and **Y** with equal concern and respect?

It is not that the action of following GR3 and my

subsequent treatment of **X** and **Y** entails that I am not

following E2. It simply leaves it an open question

whether I am following E2, i.e. treating others with

equal concern and respect, or not. Likewise my

conferring different legal or moral rights on X and Y

leaves it an open question whether or not I am

following E3.

Finally, it might be proposed that action in

accordance with GR by **A** implies that if **A** treats some

other, **B**, in a certain way then **B** has a right to treat

A in the same way. Suppose we interpret 'right' as

'legal right' here. But certainly if **A** treats **B**

terribly no legal right of **B** is necessarily established

that he may treat **A** badly. **B** doesn't have a legal right

to treat **A** badly either because **A** wants him to or

regards it as right that he do so; nor does **A**

necessarily have the power to authorize **B** to act in a

certain way towards him. For example, a girl of 10 doesn't have the legal power to give another the legal right to have intercourse with her; a fortiori her thinking it right does not confer such a legal right.

So if **A** treats **B** badly does this mean that **B** has a moral right to treat **A** badly (given that **A** acts in accordance with GR?) Again it is dubious that **B** has a moral right to treat **A** badly, e.g. to torture him, just because he wants to be tortured, or regards it as right that he be tortured. But then if **A** isn't empowered to confer on **B** the moral right to torture him, **A**, he cannot give **B** the moral authority to do this. Hence in the context of moral authority it is not even possible for **A** to act in accordance with GR3.

Thus it is not clear why by acting in accordance with GR we are treating others as equals. Moreover even if the following entailment were sound: 'if you act in accordance with GR then you treat others as equals,' then were we to try to infer from your not acting in accordance with GR that you were not treating others as equals we would be committing the fallacy of denying the antecedent. Rather to show that failure to act on GR means not treating people as equals what we need is the conditional: 'if you treat others as equals then you are acting in accordance with GR.'

So it is unclear why it is that in violating GR we are not treating others as equals and hence why it is

that some equality has to be restored in some way. And
how is it that violation of GR amounts to the assertion
of a right?

Suppose **A** kills **B** and in so doing violates GR. In that
case he either violates GR1, GR2 or GR3. Suppose
further that to assert a right is to claim a right. But
then **A**'s violation of GR hardly shows that he is
asserting a right. This is because by killing **B** and so
violating GR it may be that **A** is not willing to be
killed himself, or he doesn't regard it as right that
he be killed, or he doesn't authorize anyone to kill
him. But the fact that **A** doesn't want to be killed
clearly doesn't show that he is claiming a right not to
be killed. A necessary condition of the latter is that
A regards it as at least prima facie wrong to kill him.
But in wanting not to be killed he may think nothing of
the sort. Meanwhile **A**'s not regarding it as right that
he be killed is not indicative of whether or not
another has the right to kill him. Lastly, **A**'s morally
authorizing another to kill him is probably not
something he can do anyway so his not doing it does not
seem to add any weight to his claiming a moral right
not to be killed. So, in all, violation of GR doesn't
appear to entail the assertion of a right.

The second question 'how is it that in doing the same
back to the wrongdoer we assert a right towards him?'
may be asked here too. Clearly if the foregoing was

250

designed to establish why both the actions of the
wrongdoer and the state amount to assertions of right
then it has not been shown how it is that the state
does assert a right.

Partly on the strength of the foregoing claims Reiman
affirms what he calls the retributivist principle: in
virtue of the equality and rationality of persons the
offender deserves, and the victim has the right to
impose, suffering on the offender equal to that which
he imposed on the victim. However this principle will
not serve to show that retribution as we defined it
earlier is deserved by the offender and the right of
the victim.

Firstly, according to the retributivist principle
(RP,) punishment according to the lex is deserved by
the offender. But such punishment is not retribution,
for it involves the imposition of equal suffering on
the offender. Yet clearly this may fail to be
retribution in either of the senses outlined earlier;
that the offender experiences equal suffering to the
victim of his offence, with the proviso that the
offence be a morally wrong action, obviously may not be
consistent with the requirement that the greater the
blameworthiness of the agent of the immoral action, or
the greater the immorality of that action, the greater
the punishment. So RP can only show that something
other than what on our understanding is retributive

punishment is deserved. For the same reasons RP cannot

show that the victim has a right to retribution (or

retributive punishment) as we have defined it.

Secondly, even if the punishment or imposition of

suffering cited in RP were retributive RP itself has

not been shown to be true. Punishment in accordance

with the lex is deserved, according to RP, because it

is just. The ground for thinking the punishment

advocated in RP is just is that it treats people as

equals. But this is not shown by Reiman. He fails to

show that RP punishment involves treating people as

equals because he does not show that such punishment

involves everyone following the Golden Rule, nor that

in following GR we treat one another as equals.

Moreover the conditional ' if the society employed the

lex in its punishment system then everyone would follow

GR and hence everyone would treat one another as

equals' is ambiguous in an important way. It may either

mean that if society uses the lex then all other

actions will be just or it may mean that if society

uses the lex then all actions in society will be just,

i.e. that employment by society of the lex is a

sufficient condition of all action in that society

being just. But then if the conditional is only true on

the former interpretation then it is crucially

deficient in demonstrating that the offender deserves

the punishment of RP. This is because we still need to

know that the punishment in accordance with the lex is just, not simply that every other action is just. For if RP punishment isn't just then surely it can't be deserved. So it has not been shown that RP punishment is deserved.

Likewise it has not been shown that RP punishment of the offender is a right of the victim. Again, for the same reasons canvassed, if we haven't shown that RP punishment is just we haven't shown that the victim has a right to it.

Secondly, even if it were true, as Reiman claims, that doing back to the offender what he has done to you asserts the same right towards him that he asserted towards you it is still mystifying how this joint assertion of a right could annul a violation of the equality of persons; that A and B assert the same right vis-a-vis one another doesn't mean they both have the right. Hence this state of affairs is not even one of equality with respect to a particular right.

So the Reiman account as it stands does not show that retributive punishment as we have defined it is deserved or a right of the victim, nor does it show that punishment in accordance with the lex talionis is deserved or a right of the victim of an offence.

(3) The lex, equality and rights.

A connexion between the lex talionis, equality and

rights may be found in Kant in the translation of his

work entitled 'The Philosophy of Law' (Tr. Hastie,1887, pp. 194-8.)

[1]According to Kant it is the principle   of equality which determines

how an individual is to be punished.

'It [the principle of equality] may be rendered by

saying that the undeserved evil which anyone commits on

another is to be regarded as perpetrated on himself.

Hence it may be said: If you slander another you

slander yourself; if you steal from another you steal

from yourself; if you strike another you strike

yourself; if you kill another you kill yourself. This

is the right of retaliation (jus talionis);and properly

understood it is the only principle which [...] can

definitely assign both the quality and quantity of a

just penalty'.

Now it seems clear that one generally has a moral and

(with certain exceptions) a legal right to pain or

injure oneself. But the problem is that this does not

show that another has the same right. It appears that

what Kant has in mind is that when B does back to A

what A did to B it is as if A is doing it to himself.

But of course it is not A doing it so this is a

mystifying justification of retaliation. Moreover, even

if B has a right to hit A because A has hit B and B's hitting A is like A hitting himself this wouldn't show that it was just to hit A, any more than it is just for me to hit myself.

A simpler way of interpreting the above is that because all persons are equal if A slanders B then there is a right to retaliate. But of course the retaliator must be the victim in which case only B has a right to slander A. But if we accept this it does not square with an earlier comment of Kant's that only the sovereign as head of state may punish. Of course it could be argued that B's right is a right to slander but not a right to punish. But then if we were to accept this it would appear that on Kant's account it was quite legitimate for A to undergo the pain of what he did to B more than once, once by the state as punishment and once by B as retaliation. But Kant obviously does not hold this. Hence we would do well to interpret the right of retaliation as simply the right to do to an individual the same as he has done to some other, ignoring that part of the content of the concept of retaliation which requires that only the victim or one who empathizes with him should perform it. But then what is the justification for the claim that if A slanders B the state has a right to slander or do something equivalently unpleasant to A? Kant says a little later in the passage just quoted from :   'But

how then would we render the statement ''If you steal from another you steal from yourself?'' In this way, that whoever steals anything makes the property of all insecure; he therefore robs himself of all security in property, according to the right of retaliation.'

One way of interpreting the above is that if A steals from B then this produces a general insecurity for all. But this need not be. If A is the only wrongdoer then he and no other may be safe, in which case it is not true that all exist in a state of general insecurity. But even if all were in a state of such insecurity and it were possible that anyone should suddenly find himself in a position to be stolen from it would not follow that anyone had the right to steal from A or from anyone else. A may have lost his security but it does not follow from this that he has thereby lost all right to it. So it still remains to be seen where the right of retaliation comes from.

In fact it is the equality of A and B which supposedly grounds the right of retaliation. It might be claimed that because individuals are equal in some sense then (1) they only have a right to something X provided that they respect the right of others to X. A variant of this would be that (2) A only has a right to X towards B if he respects the corresponding right of B to X. This latter means that B only has a duty to refrain from interfering with A's access to X if A

refrains from interfering with his, i.e. B's. Let us
call these two claims the reciprocity and partial
reciprocity claims respectively.

Now if the reciprocity claim is in fact true will
this show that there is a right to retaliation on the
part of the state? Clearly A does not respect B's right
to enjoyment of his property if he commits burglary of
his house. But then it is not the case that A respects
the right in question of all others apart from A. Hence
he forfeits that right. But the reciprocity claim does
not state for how long the right is forfeited-it is
forfeited indefinitely. Hence A loses the right to
enjoy his property indefinitely. This would then mean
that the state could interfere with A's enjoyment of
his property. But there are at least four problems with
this.

Firstly the state can interfere with the enjoyment by
A of his property indefinitely. But this means that we
need something more than the reciprocity claim to show
that action in accordance with the jus talionis as Kant
refers to it is just. For action in accordance with the
jus talionis will be of a certain kind and a certain
duration while action which is permissible because of
the forfeiting of the right in question by A will only
be of a certain kind, not of a certain duration.

Secondly, A may have no property to enjoy. However
because of the way the reciprocity claim is worded we

cannot substitute some form of treatment of **A** which is equivalent in respect of unpleasantness. Hence if a vagrant breaks into **B**'s house the state cannot punish in virtue of the right of retaliation.

Thirdly, if **A** forfeits his right to enjoy his property then this means that anyone can interfere with that property. This cannot be objected to on the grounds that interference by someone like **B** with the property of **A** would interfere with interference by the state with the property of **A** for the reason that the state does not rightfully interfere. Rather it is permitted to interfere in the absence of a right of **A**'s toward it. If there is limited access to some activity it may be that only one individual at a time can engage in that activity. For example, no one can walk down exactly the same part of the pavement at the same time as me. But if I'm there first I am entitled to occupy that space even though by doing so I prevent others from occupying it.

Fourthly, the reciprocity claim will not show that punishment by the state is just if it takes the form of retaliation but at best that it is permissible. However if it is permissible then it is also permissible for the state not to punish. Yet this would be contended by Kant. A state must punish for a crime (and by extension a wrongdoing) in accordance with the jus talionis because it is required by justice.

Moreover some of the above criticisms apply if we use the partial reciprocity claim with the additional problem that the state is not even permitted to punish since the right which has been forfeited by **A** is not towards it. At best **B** is permitted to interfere with **A** in the enjoyment of his property.

Ross (The Right and the Good, 1930, Chapter 2,) questions the truth of the partial reciprocity claim. He accepts that the claim holds with respect to duties which arise out of contract but not with other duties. However the example he gives of duties arising out of contract does not in fact vindicate what we mean by the partial reciprocity claim. Ross argues that where **A** and **B** enter a contract, **A** to perform one action a and **B** another b, **A** and **B** respectively have 'contingent' duties to perform those actions. (They are contingent in the sense that failure by one or other of the contracting parties may absolve the other from his contractual duty.) Thus if **A** does not perform a **B** is under no obligation to perform b. But this does not show that partial reciprocity is at work here because this example is not one of a case in which **B** has a right against **A** that he, **A**, perform a certain action a, as a result of the non-performance of which **A** forfeits his right against **B** that **B** do a. For ex hypothesi **A** and **B** have duties to perform different actions. Rather the example Ross proposes is one in which **B** has a right against **A** to perform a and as a result of the non-

performance of a A forfeits his right against B to perform b.

However the example can be shown to serve the purposes Ross intends it for by simply noting that where A and B enter into a contract A has a right against B that B do what he has contracted to do while B has a right against A that A do what he has contracted to do. And if A fails to respect the right of B against him then indeed B is released from the right of A against him to do what he has contracted to do.

Ross notes that the partial reciprocity claim does not appear to fare so well in the case of telling lies. If A tells B lies then B is under no obligation not to tell A lies if the partial reciprocity claim is true. Yet, as Ross correctly points out, we believe that B ought not to tell lies to A even if A has told lies to B.

So far it has not been clarified whether moral rights or legal rights or both are at issue. Clearly if we have legal rights in mind the partial reciprocity claim does hold for the above example of contractual obligation. This is because A will only have a legal right against B that he do what he, B, has contracted to do where B has a like right against A.

Suppose we translate the second example about telling lies into the terminology of rights. But this may not

be possible. The point about the second example is that it concerns a general moral obligation as opposed to a contractual one. Of course it can be that **A** and **B** have legal rights toward one another that they be told the truth by the other. For example, where **A** and **B** are bartering and **A** says that his goose can lay golden eggs, **B** knowing all along that it cannot, it does not seem that **B** is required to tell the truth about the goat he is trying to barter. But quite generally, it is not true that **A** has a legal right to the truth from **B** concerning all matters, and that this legal right is contingent on **A**'s respecting **B**'s legal right to the truth from him, **A**. Yet if we construe the partial reciprocity claim as: **A** has a legal right against **B** to **x** only if **A** respects the corresponding legal right of **B** then the claim gives a necessary condition of possession of a legal right. But if there is no legal right to being told the truth generally then a claim about a necessary condition of a legal right is not falsified by the fact that it does not seem that the right (simpliciter) to the truth generally is lost by an individual vis-a-vis those to whom that individual does not tell the truth. So the partial reciprocity claim as a claim about legal rights is unaffected by the second example.

Before we can assess whether or not **A** has a moral right to **B** telling him the truth we have to be able to

determine whether the fact that **A** has a duty to tell **B** the truth licenses the inference that **B** has a moral right against **A** that **A** tell him the truth. Ross considers the correlativity thesis of (moral) rights and duties in the light of the cases of our duty to animals and the so-called duty of beneficence.

Consider the following pair of propositions: 1. A duty of **B** to **A** implies a right of **A** against **B**, and 2. A right of **A** against **B** implies a duty of **A** to **B**. Now if we accept that 3. we have a duty to animals but 4. they do not have a duty to us it cannot be that propositions 1 and 2 are both true. If we accept that we have a duty towards animals then they have a right against us by 1. But then by 2 animals have a duty towards us. Hence it can't be that 1,2,3 and 4 are all true together. As Ross notes in accepting 1 we appear to be embracing the view that a duty of **B** towards **A** is a sufficient condition of **A** having the corresponding right against **B**. Thus in Ross's words 'possession of a nature capable of feeling pleasure and pain is all that is needed in order to have rights'. On the other hand acceptance of 2 commits us to the view that only moral agents can have rights. Ross tentatively endorses the view that we should accept 2 rather than 1 since ' On the whole, since we mean by a right something that can be justly claimed, we should probably say that animals have not rights, not because the claim to humane treatment would

not be just if it were made, but because they cannot
make it.' But to say that a right is something which
can be justly claimed is open to more than one
interpretation. We may say that the claim amounts to
the conditional i) 'if x has a right to y then x can
justly claim y. But the consequent in (i) is itself
open to interpretation either as the conditional ii)
'if y were claimed by x then that claim would be just'
or as the conjunction iii) 'if y were claimed by x then
that claim would be just ' and ' x is able to make that
claim'. Clearly if we interpret the consequent of (i)
as (iii) then the fact that animals are unable to make
the claim in question means that the second conjunct of
(iii) is false and so that (iii) itself is false if we
substitute 'animals' for x and make the proper
grammatical corrections. Hence by MTT if we interpret
the conditional (i) as having the consequent (iii) then
it would follow that animals do not have rights.
However we are not forced to make this interpretation
by anything Ross has said and because he implies that
the claim made by animals would be just if made then on
the interpretation of (i) as having the consequent (ii)
the consequent of (i) is true not false with the
appropriate substitution for x. Hence it is not shown
that 1 must be false: we cannot infer that animals do
not have rights on the strength of a negation of the
consequent of the conditional (i,) yet nonetheless we

do have duties towards them. But in any case this is
more in keeping with the general strategy of Ross's
argument for the reason that if 1 were false then we
could not infer from his comments about the reciprocity
of duties anything about the reciprocity of rights.

In addition we seem to have good reason for not
falsifying 1 for the reason that if we are to make
moral agency a necessary condition of possession of
rights then the rights of the mentally ill and
handicapped as well as infants become more of a problem
to account for (assuming that they do have rights.) So
the upshot of this discussion is that, even if a right
is something which can be justly claimed, this doesn't
show that animals don't have rights. But if animals may
have rights then it may be consistent that we have a
duty towards them and 1 is true. Hence it is not shown
that 1 must be false.

However the proposition 1. 'A duty of **B** to **A** implies
a right of **A** against **B**' has still to be tested in the
light of the duty of beneficence.

Ross believes that the reason that one objects to the
idea of a right to beneficence is that this is confused
with a right to benevolence. While beneficence is to do
what is good benevolence is to act from a spirit of
goodwill. Now Ross proposes that while we can claim as
a right treatment which is just we cannot claim as a
right that we receive justice in the spirit of justice.

Rights concern actions rather than the spirit in which
or from which those actions are performed. But as Ross
himself concedes this does not account for why it might
be thought that although the individual who helps the
old lady across the road from fear of being thought
unkind does his duty and acts beneficently the old lady
cannot claim the action as a right.

Moreover it is difficult to see what the force of
'claiming as a right' in the context of moral rights
can be if this is something the content of which cannot
be drained by talking of what another ought or has a
duty to do. In the context of legal rights the claim of
a right may take the form of legal action. We might
think of a just claim in the context of a legal right
as signifying two procedures, the first a requirement
that there be a proper judicial determination of
whether in fact a legal right has been violated and the
second the taking of action by the legal machinery in
respect of the violation if any. By contrast if one
does not have a legal right there is no requirement
that some judicial determination of the merits of the
case be made, nor any requirement of a follow-up
procedure. But in the case of a moral right the idea of
a just or valid claim can have no such significance.

However it does seem to be the case that a moral
right must have some kinship with a legal right. One
possibility is that if one has a moral right one has a

just claim that ought to be heard in court. But this
seems dubious. For it may be that if everything to
which a moral right attaches were to be the proper
subject of judicial inquiry very few cases could
actually be heard, in which case many very important
abuses of rights which are at the moment dealt with by
the courts could no longer be so dealt with. But then
it would be very implausible that, given the
circumstances, such claims deriving from moral rights
ought to be heard in court. One might still maintain,
of course, that in a perfect world such claims ought to
be heard because their hearing would not interfere with
the hearing of more important abuses. This invites the
question of whether one thing which courts may well
dispense if they see fit, namely redress, is part of
the key to understanding the nature of a right and part
of why one objects to the idea of a right to
beneficence. For if a right were to be understood in
this way the failure of the individual to help the old
lady across the road would be something for which, in a
perfect world, some sort of redress could officially be
contemplated. That this might seem a tremendous
limitation on individual liberty may be one way of
circumscribing what we are prepared to regard as a
right. The problem with this sort of consideration
however is that it yields no readily applicable test to
determine whether or not something belongs to an

individual by right. Whether or not, to return to the earlier case, an individual has a right to be told the truth would depend on whether or not it would be a great limitation on human freedom to contemplate redress and ultimately provide it in a certain number of cases of failure to tell the truth.

I shall not pursue these sketchy remarks about redress and rights further. I do take it to be the case, however, that one does have a duty to help the old lady across the road. Moreover it is reasonable to think that if redress is part of the significance of rights talk, and all the deterrent effects on non-compliance with rights which would go with the possibility of redress, at least some things which we feel people ought to do would not be things which we would want to correlate the corresponding rights with.

So the duty of beneficence may not be correlated with a moral right to beneficence, in so far as it is at all clear what a moral right is. Hence where the right in 1 is construed as a moral right 1 may be false and so part of the correlativity thesis may be false. But this means that the duties of **A** and **B** to tell the truth aren't necessarily correlated with rights against each other to tell the truth.

Now we are considering whether even if **A** tells **B** lies **A** has a right to the truth from **B**, i.e. that we have a counterexample to the partial reciprocity claim for

moral rights. But, as we have seen, there is some doubt whether we can infer from B's duty to tell the truth that A has a right to the truth from B. So the fact that B seems to have, despite A's conduct, a duty to tell the truth does not of itself show that A has the kind of right to the truth from him which would show that the above counterexample holds, i.e. which would show that A can have a right against B in the absence of respect for the corresponding right of B. Hence the upshot of the discussion is that it hasn't been shown that the partial reciprocity claim is false where this claim is construed either to be about moral or legal rights.

A different account which might be thought to have a bearing on the partial reciprocity claim is offered by Nino. Nino purports to give an account of the justification of punishment in terms of the social protection which is provided by it. Now one problem which this poses is that a general justification of an action in terms of social protection, say quarantine for example, seems unfair to the individual who is a target of that action, the quarantined individual. However we can mitigate this unfairness by means of providing compensation for the quarantined individual. However in the case of punishment such an option of compensation is not open to us for this would defeat the object of the (supposedly) deterrent effect of

punishment. In the case of punishment the unfairness of
the situation, an unfairness which may either reside in
the fact that the punishee is being used as a means to
some end rather than being treated as an end in
himself, or because there results from punishment an
unfair distribution of benefits and burdens, is offset
or removed by the fact that the punishee has consented
to his liability for punishment, that is, that it is
legal to punish him. Since if, under certain
circumstances, we consent to something happening to us
as a result of someone's actions it is fair for that
someone to make that thing happen to us then under
certain circumstances punishment is not unfair.

The point about the account which has bearing on the
present difficulty is that forfeiture of legal right is
justified by consent under certain contractual
circumstances. The action constitutive of consent must
be voluntary and made in the knowledge both of the law
and the factual consequences of that act, the
bargaining power of the individual who consents must
not be very dissimilar to that of those with whom he
enters into the contract, and, in the case of consent
to liability for punishment, the rules, violation of
which creates liability for punishment, must themselves
be fair.

One problem which such an account appears to invite
concerns the nature of the punishment to which one

consents. For if punishment takes the form of, in part, the deprivation of certain legal rights, there appears to be no reason why we should not punish non-retributively. For example, provided that an individual consents under the appropriate circumstances to 15 years hard labour for stealing a VCR it would be fair, supposedly, to punish him with this, while one who murders may consent to only 5 days hard labour. Now Nino cannot block this type of scenario by pointing out that one could not properly consent to liability for such punishments on the grounds that the bargaining power of one who so consented would be attenuated in some way; nor can he point out that this is a case in which the legal rules violated are unfair - the legal rule prohibiting homicide may be fair even though the rule specifying the punishment for violation of that rule is itself unfair, and it is only the former rule that Nino specifies must be fair. In fact the only objection Nino can make here to non-retributive punishment is that it fails to secure the social protection which is its primary source of justification. But if the only criterion for determining which punishment is morally acceptable is whether it secures social protection at minimum cost then the difficulties attendant on a utilitarian-type justification of punishment beset Nino's account. For one thing, as with utilitarian justifications of

punishment, Nino cannot show that the right punishment will be retributive. More to the point here he will not be able to show that if **A** infringes **B**'s legal right to **x** then **A** only forfeits the legal right to **x**. For according to Nino it will be fair, and hence, presumably, no infringement of a legal right, for **A**, under certain conditions, to forfeit more (or less) than the legal right to **x** depending on whichever course of action delivers the best package of social protection.

Secondly, even if maximum social protection were only purchasable by means of punishment such that the punishee was only deprived of a right which he had infringed in another it is unclear why consent to this deprivation would make the punishment fair. For the consent is supposed to confer fairness on whatever punishment is necessary for some measure of social protection. Yet it is far from intuitively obvious that if **A** consents to be hanged for theft of a loaf of bread that there is some principle of fairness according to which what we consent to it is fair to bring about. The point is that the power of consent to confer fairness on a punishment is understood against a backdrop in which the punishment seems fair for independent reasons. But if this is so then consent cannot play the only part in showing that forfeiture of a legal right in accordance with partial reciprocity is fair.

So the partial reciprocity claim has not been shown to be morally acceptable; nor has it been shown that its justice would mean that the lex talionis is just. What, finally, are we to make of the claim that it is the fundamental equality of persons which justifies the partial reciprocity claim?

Firstly we need to decide what this fundamental equality is. Let us suppose that the claim of equality is to be interpreted as 'all persons are equal in respect of their value or inherent worth.' How are we to determine what the value of an individual is? We may speak of the value of certain things to certain individuals. Thus two vases owned by X may be of equal value to him. But then if all persons are of equal value they must be of equal value to someone. If one resists the temptation to answer 'to God' certain answers are clearly false. For example, if we choose an individual at random it is unlikely that all people will be of equal value to him or her. Usually those closest to us are of greater value than those of whom we know nothing on the other side of the world. Moreover if we could find some one individual for whom this was a true claim it would be difficult to see why the value beliefs of some one unknown individual confer on the claim the sort of respect it is usually accorded.

We might modify the claim somewhat to 'all persons have equal intrinsic value'. It might be thought that this claim would be invulnerable to the above criticism that those closest to us are of greater value to us. For it might be said that the reason for this is that those people closest to us give us things which are of value to us and so the value which such people afford us is instrumental rather than intrinsic. Thus it could be claimed that that part of the value to us of kith and kin which was extra to the value to us of strangers was instrumental rather than intrinsic.

Certainly something may be of intrinsic value to an individual. My own happiness may be of intrinsic value to me because at least part of its value to me may be grounded in nothing else. But it seems that anything of value to me which is of intrinsic value is going to be valuable in virtue of being some state or experience of mine, e.g. my health or happiness or pleasure. It is hard to see how another person could be of intrinsic value to me since such value would be mediated by some valuable state/experience I had as a result of him/her. Nor perhaps can I be of intrinsic value to me. If we say ' x is of intrinsic value to me' iff (i) x is of value to me and (ii) it is not the case that x is of value to me in virtue of its relation to something else y which is of value to me, then I will not be of intrinsic value to me because, in all probability, I am

of value to me in virtue of my relation to certain states or experiences I have and so condition (ii) will not obtain. Of course such considerations would, if correct, serve to reinforce the idea that people are of equal intrinsic value to everyone on the grounds that people are of equal zero intrinsic value to themselves and one another. But if this is why the claim of intrinsic equality is true it is not much of a claim on which to ground claims about justice.

A different way of interpreting the claim is as a stipulation that people are of equal positive value intrinsically. It might be thought that if we were to start with this sort of stipulation it would follow that if I could only save one out of five individuals then because all individuals were of equal intrinsic value it would be morally irrelevant which of the group I saved. But this would be a poor inference. I might have reason to believe that if a certain X were not saved his mother would die of heartbreak as a result, or I might have good reason to believe that X was on the verge of a breakthrough in the treatment of cancer; in either case it might be that I was morally obliged to save X and not some other. It is far from clear that instrumental value plays no vital part in moral evaluation of outcomes, rather the importance of intrinsic value is that it makes the point that no one person's happiness, valued in itself, is more important

than any other's. It would amount to the utilitarian injunction that each was to count for one, and not more than one.

So perhaps we may see some significance in the idea that persons are of equal value. But how does this help us to understand the partial reciprocity claim?

As we have noted before rights forfeiture in the partial reciprocity claim is about making certain actions permissible rather than right or morally required. Hence we are looking to show that the fact that people are of equal positive (intrinsic) value somehow explains that it is permissible (roughly) to do to **A** who violates **B**'s right to **X** something which interferes with the freedom of **A** to **X**.

Let PE be the principle of equality as we are interpreting it. Suppose, further, that PE shows that the following principle, PED, is true: if **A** causes **B** disvalue **d** then the most disvalue **B** is permitted to cause **A** is disvalue equal to **d**. If PE shows that PED is true does it show that the partial reciprocity claim PR is true?

Consider now a case in which **A** tries, unsuccessfully, to murder **B**. In such a case **A** has violated not **B**'s right to life but **B**'s right not to be the subject of an attempt on his, **B**'s, life. Thus **B** may try to kill **A**, but he is not permitted to be successful. Ex hypothesi, PE shows the truth of PED and so **B** is morally permitted

to cause **A** as much disvalue as he, **B**, was caused. But perhaps **B** was caused no disvalue by **A**'s action; perhaps he never got wind of it. But then the action permissible according to PR, viz. one violative of **A**'s right not to be the subject of a plot to kill him, is not necessarily permissible according to PED. The reason for this is that while some actions violative of **A**'s right not to be the subject of a plot to kill him may cause **A** disvalue some might not. So PR can't be justified by PE on the grounds that the latter principle verifies PED.

Secondly, PED is morally unsatisfactory in the sense that if we are only permitted to punish in accordance with PED we will not be permitted in some cases of attempted murder, say, where the murder subject is occasioned no disvalue, to inflict any punishment at all. But it is interesting that PR seems no less deficient than PED in this respect. This is because although, in attempting to murder **B**, **A** may violate rights of **B** other than his right not to be the subject of a murder project, it may nonetheless be that the action permitted by way of retaliation on the part of **B** owing to the rights which **A** has forfeited will not be of any disvalue to **A**. PR may fail to have teeth like PED.

Finally, it does not seem that PE is a principle with which one necessarily does not act in accordance where

one violates another's right, nor is it clear that one who acts in accordance with PED is respecting PE. If it is a necessary condition of acting in accordance with PE that one regards others as being of equal intrinsic value with oneself, and we assume what is by no means proven - that such an attitude entails that one regards the same states of others as being of equal intrinsic value with similar states of one's own - then the fact that **A** burgles **B**'s house is consistent with **A** regarding **B**'s comfort and security as no less (and no more) valuable than his own. Secondly, in acting in accordance with PED one may not regard the state of the violator of a right which is interfered with as being of equal intrinsic value with that of the one whose right was violated.

So we have failed to show a connexion between PE and PR via PED. Moreover, since we have not shown that PR is verified by PE, where the former is a claim about what it is permissible to do, then a fortiori it has not been shown that there is a right to retaliation or that Kant's jus talionis is just.

I think that in order to forge some connexion between the partial reciprocity claim and PE we need to be able to give some account of the relation between certain modes of treating an object and the value of that object.

It seems clear that it makes sense of a curator to say that a particularly fine picture deserves to be hung in better view of the public than it is, it being hung at present in some out of the way corner too high up to be properly appreciated. Here, roughly, it seems that if pictures of less value are being hung in good positions where they can be seen it is in some sense inappropriate to hang a very good picture in a less good position. The following principles suggest themselves: 1. where objects a and b are of unequal value the treatment which is received by that which is of lesser value ought not to be better (in some sense) than that which is received by that which is of more value; 2. where resources are limited, for example where gallery space is limited, and good treatment is only available for a few then it should be those of highest value which gain access to that good treatment; 3. it seems plausible that things of equal value should, where resources allow, receive equally good treatment.

The point of this is that there are cases in which it appears that we ought, in some as yet unexplained sense, to accord things certain kinds of treatment based on their value. Turning to the subject of people and what they deserve I think we can find cases in which, plausibly, differential treatment is accorded to individuals on the basis of estimations of their

intrinsic value. Some believe, and I find this abhorrent, that animals deserve less favourable treatment than humans, at least where resources are limited. Thus, those who think in such a way might follow the second principle above, namely, that where resources are limited, and only good treatment is available for a few, then those who get the good treatment ought to be humans. And their justification for thinking this way might be that they regarded (non-human) animals as being of less intrinsic value than humans. One might also cite examples in which differential deserved treatment for humans is at issue. One can imagine from an era of slavery the expression of a view to the effect that a certain mode of accommodation, though degrading for a freeman, was appropriate for a slave, i.e. not so bad that it could be objected to on the grounds that he deserved better. Doubtless Nazi Germany and Apartheid would provide other instances.

Principle 1 above entails principle 2. If we did not follow 2 we would contravene 1. On the other hand 3 doesn't entail 1; nor does it entail 2. 1 and 2 do not entail 3 and so 3 is logically independent of the other two principles. Thus we may reduce our principles to just 1 and 3.

But 3 poses problems. Suppose we embrace principle 3 and PE. Does this enable us to show that PR is a

principle we must endorse? Consider the assumption that all people are of equal intrinsic value and principle 3, viz. that where resources allow, things of equal value should (=ought to) receive equally good treatment. So imagine a case in which **A** violates a right of **B**'s by burgling **B**'s house. Suppose further that it is **B**'s comfort and security which is interfered with by **A**'s action. Now we are supposing that in virtue of the assumption that all people are of equal intrinsic value the comfort and security of **A** and **B** are of equal intrinsic value. But then their comfort and security, assuming unlimited resources, are, by 3, deserving of equally good treatment. Does it follow that **A**'s comfort and security may be interfered with to the extent that **B**'s has been?

Clearly not. For one thing the use of 3 appears to require our taking into account all treatment which the comfort and security of **A** and **B** have received. If, over the period t, the comfort and security of **A** and **B** have equal value, then during t these states should receive equally good treatment. But the fact that **A** violates **B**'s right and so interferes with his comfort and security does not show that by similar interference on the part of **B** their comfort and security will be treated equally over the period t. It might be that while **B**'s comfort and security had received, on the whole, very good treatment for the period t, **A**'s had

received very poor treatment on the whole for the same period. But then, far from moving us towards an achievement of equally good treatment for their respective states similar interference with **A** by **B** might move us towards greater inequality of treatment. So as they stand the assumption of the equal intrinsic value of people and 3 don't endorse PR.

Secondly, the assumption that where **A** and **B** are of equal intrinsic value similar states of **A** and **B** are of equal intrinsic value may well be problematic. We might try to show that there is some contradiction involved in the supposition that **A** and **B** are of equal intrinsic value and yet that some state **s** of **A** and **B** is not of equal intrinsic value. Suppose then that **s** of **A** is not equal in intrinsic value to **s** of **B**. But if **A** and **B** are of equal intrinsic value then, given the earlier stipulation about intrinsic value, **A** and **B** are of equal value, and this value is not dependent on **A** and **B** standing in a certain relation to something (other than **A** and **B**.) But then their value is not equal in virtue of the equality of respective states of **A** and **B**. But then the putative difference in value of these states cannot affect the equality of value of **A** and **B**. So we haven't shown a contradiction. Moreover it is dubious that a contradiction could be found since there is an essential gulf between the value of **A** and **B** and the value of states of **A** and **B** created by one of the

stipulated conceptual conditions of the nature of
intrinsic value.

Let us suppose that we may infer that states of **A** and
B are of equal value and, further, that the kind of
background assumption concerning the histories of
treatment of the comfort and security of **A** and **B** may be
made. Can we infer that it is morally permissible for **B**
or the state to retaliate in the form of equivalent
interference?

There are still problems with the inference however.
It has to be clarified whether the forfeiture of rights
in PR is of moral or legal rights. Secondly, the
meanings of 'ought' and 'good' in 3 have to be
clarified.

Consider the second problem first. Firstly, it must
be noted that in the picture in a gallery example it is
plausible that we are speaking of cases of relative
instrumental value. The value of a picture is probably
dependent on its relation to something other than
itself, viz. certain states of human spectators. But if
the principle in 3 is about instrumental value then the
inference will not go through since we only have a
premise about equality of intrinsic value.

Secondly, if we interpret 3 as a claim concerning
equality of intrinsic value it is still not clear what
the force of the 'ought' is. In the example of the
pictures, where instrumental value is at issue, it

might be suggested that the duty of finding the appropriate position for the picture in question is one of beneficence. By hanging the better picture where the worse one is people will get more from the experience of visiting the gallery. But clearly failure to abide by 3 need not evince a want of beneficence. Indeed 3 may be challenged in cases where no failure of beneficence would result from failure to apply it, as for example where I put one of the pictures in a better position than the other where no loss of enjoyment of the pictures will result. So the principle in 3 is itself questionable.

Likewise, where 3 is interpreted to be about intrinsic value and is applied to animate objects we might again question whether action in accordance with 3 is a requirement of beneficence. It is at least questionable that it is. For suppose **A** hurts **B** but that the hurting of **A** by **B** would serve no useful purpose. We may suppose that **B** would derive no vengeful pleasure from hurting **A**, no one would be deterred or reformed by its occurrence, etc.. But then surely the duty of beneficence in such a case would require that **A** be left alone since, one may suppose, **A** would be happier left alone, and so the total happiness of **A** and **B** would be higher if **A** were not hurt.

So where we try to construe the 'ought' in 3 as a moral 'ought', and we base the principle on the duty of

beneficence, the principle itself comes to appear inadequate as a justification for the permissibility of retaliation. Of course if 3 just is a claim about beneficence then there may be good reason for supposing that in most cases it would be a duty of beneficence, that is, it would maximize some desirable state of the individuals concerned, to retaliate or have an aggressor punished. But then 3 should be understood as something like a prima facie claim, one which would only demonstrate that it was right for an aggressor to be punished or be the object of retaliation, in most cases. But this 'on balance' interpretation of the force of 3 is certainly not what Kant had in mind, since we appear to be entering into utilitarian calculations the like of which he eschews.

So if the 'ought' in 3 is a moral 'ought' the principle of 3 may be inadequate for the purpose of endorsing PR.

And what of the idea that equally good treatment be administered? Firstly, 3 has to be understood in such a way that it is not disproved by certain obvious examples. To take one at random, that one schoolboy is more academically gifted than another is often thought to make it fair for one rather than the other to get a scholarship to an expensive school. But there is no question here of the children concerned being possessors of different intrinsic value. So 'equally

good treatment' looks as if it is going to be difficult to give anything but a highly abstract interpretation of if the principle in 3 is to be plausible. Certainly equally good treatment isn't the same treatment in this context, but then why should the same treatment that **A** administered to **B** be something which is argued for by our principle when it is administered in reverse, i.e. by **B** to **A**?

One way of interpreting equally good treatment would be as ' equally right or wrong.' But to begin with it is not clear what could be meant by 'equally right' since right does not admit of degrees. Moreover if **B** is to do to **A** what is equally wrong to that which **A** did to **B** then the principle in 3 clearly cannot function with PE to justify PR. This is because when **B** treats **A** equally on such an interpretation of equal treatment he does something just as wrong to **A**. But since we are supposing that what **A** did was wrong in the first place all we would have succeeded in showing was that it was wrong of **B** to retaliate, not that it was morally permissible.

Alternatively, 'equally good treatment' may mean treatment which is equal in disvalue to that experienced by the victim. This might show that rights forfeiture was justifiable in some cases but we have already touched on the difficulties involved in

justifying PR by a principle to the effect that we do something equally disvaluable to the wrongdoer.

Lastly, we must clarify whether legal or moral rights are at issue in PR. It has not been shown that it is morally permissible to take retaliatory action. But then certainly the retaliator does not acquire a moral right to do as he does. Hence the initiator of rights violation does not forfeit the moral right the counterpart of which in another he has violated. However, we have not shown, nor could our arguments be used to show, that a legal right is or is not forfeited, since what is or is not morally permissible may have no bearing on the law. If it had been shown in our example that as a result of **A**'s actions against **B A** forfeited a legal right it would have become legally permissible for B to have retaliated in a certain way. But principle 3 contains, as we are supposing, a moral 'ought' which will not give us a conclusion concerning what is legally permissible. So either if we suppose PR to concern legal or moral permissibility it has not been vindicated by any of our arguments.

So we have not shown how PE can entail PR, either via PED, or with the aid of principle 3. Nor have we shown, it may be remembered, that PR is morally acceptable or that PR entails that the lex talionis is just. Hence it still has not been shown how one could begin to make good the Reiman claim that equality provides the

justification of punishment. Moreover, even if PR could be substantiated, this would not show that the lex talionis is just since on independent grounds it can be seen that this latter principle is not.

If we could vindicate PR would this show that retributive punishment was at least permissible?

To begin with vindication of PR (construed as a claim about moral rights) only demonstrates the moral permissibility of the one whose right has been violated performing a certain kind of retaliatory action. But retributive punishment as we are understanding it is meted out by the state. So PR at best would only get us part of the way to demonstrating the permissibility (and not the justice) of retributive punishment.

Suppose then two actions, **a1**, in which **A** violates the right **r1** of **B**, and **a2**, in which **A** violates the right **r2** of **B**. Now PR will (by assumption only) show the permissibility of the punishments **P1** for **a1**, where **P1** takes the form of an action which, previous to the performance of **a1**, would have constituted the violation of a similar right of **A**'s to that which was in fact violated in **a1**. Let us describe this by saying that **P1** takes the form of the putative violation of **r1** of A. Analogously, **P2** is punishment for **a2** where **P2** takes the form of the putative violation of **r2** of **A**. But it is apparent that **P1** and **P2** will only be retributive punishments (where the principle of simple

proportionality is construed as a principle about moral wrongs and distresses,) if, given that a1 is more morally wrong than a2, P1 is more unpleasant for **A** than P2. But if indeed a1 is more morally wrong than a2 if P1 and P2 are to be retributive it must be the case that where one violation of right is morally worse than another the former violation is more unpleasant for the possessor of the right than the latter. But what this amounts to is that P1 and P2 will only be retributive punishments if the morally worse an action is the greater the distress it causes its victim. Yet it is evident that this needn't be the case and so the punishments shown to be permissible by PR won't be, as a class, the class of retributive punishments. To see this take the example of **A** and **B**, both ex-pugillists and rose fanatics, who know nothing of one another. **A** cuts down one of **B**'s roses and **B** retaliates by cutting down one of **A**'s. Alternatively, suppose **A** hits **B** and **B** retaliates by hitting **A**. Now the second action of **A** is more morally wrong than the first on the plausible assumption that we judge the immorality of an action by the harm intended by it, while each of the second round of actions may be of less distress to their targets than each of the first. So the morally worse second actions are apparently not the more distressing ones in such an imaginary case. So PR doesn't help to show the

permissibility of retributive punishment on one of our definitions of retributive punishment.

And what of the other definition of retributive punishment, where retributive equivalence requires that the more blameworthy the agent of an action the greater his punishment for it ought to be?

Imagine the preceding scenario of the actions **a1** and **a2** and the punishments **P1** and **P2**. Let us suppose that **A** is more blameworthy in the commission of **a1** than **a2**. Now this, it will be noted, doesn't entail that **a1** is a morally worse action than **a2** or that a more important right is violated in **a1** than **a2**. But then if they are to be retributive **P1** must be more distressing for **A** than **P2**. But now **P1** may involve putative violation of a less (morally) important right than **P2** does, and so might be less distressing for the violator. Hence it cannot be inferred that **P1** is greater than **P2** and so it can't be inferred that **P1** and **P2** meet one of the necessary conditions for these punishments to be retributive.

So it appears that even if PR, conceived as a principle about moral rights, had been established it couldn't have been used to demonstrate the permissibility of retributive punishment on either of our definitions for such punishment.

(4) Retribution and connexion with correct values.

An interesting proposal made by Nozick (1981, op. cit., pp. 374-88,)
is that retribution (generally) connects a wrongdoer with the
correct values. This connexion occurs where correct
values, qua correct values, have some significant
effect on the wrongdoer's life. An example of correct
values having this effect, but not qua correct values,
would be where a book on morality fell on a wrongdoer's
head so administering the retribution he deserved. The
following are necessary conditions of correct values,
qua correct values, having a significant effect on a
wrongdoer's life: (i) the message 'what you did was
wrong' is sent to the wrongdoer by retribution; (ii)
the wrongdoer understands the message: (iii) there must
be an attenuation of the wrongdoer's gladness that he
performed the wrong action. (i) obtains only if the
wrongdoer is punished because of and in virtue of the
wrongness of his act, where it is the intention of the
punisher that the wrongdoer realize this and it is also
the intention of the punisher that the wrongdoer
realize that he was intended to realize this.

One problem with this account is that (i) at least
does not look like a necessary condition of correct
values, qua correct values, having a significant effect

on the wrongdoer's life. This is because (i) specifies that the message that what he did was wrong must be sent to the wrongdoer by retribution. But surely this message may be sent by other means. One may come to see the error of one's ways and go to work in the oft-quoted leper colony. This certainly looks like a case in which the correct values, qua correct values, have a significant effect on the life of the individual in question. But since it is a sufficient condition of connecting someone with the correct values that the correct values, qua correct values, have a significant effect on the life of the individual in question then where some x, having undergone the appropriate transformation, devotes himself heroically to some mission of self-denying philanthropy, it is reasonable to suppose that he is connected with the correct values. However, if the sole justification of retribution is to be found in (re-)connexion with the correct values retribution meted out to our philanthropist will not be justified. One cannot effect reconnection where connexion is already in place.

This is probably not a very telling objection. I think Nozick might accept with equanimity that retribution exacted of our imaginary philanthropist would be unjustified. But this does mean that (i) is not a necessary condition of either the correct values, qua correct values, having a significant effect on a

life or of reconnection with the correct values. In the example of the philanthropist it appears that (i) is replaced by something like: the wrongdoer receives the message that what he did was wrong.

Doubt over whether (i) is a necessary condition of reconnection with correct values prompts a further worry. If we may justify other responses to wrongdoing, on the grounds that such responses effect reconnection with the correct values, scepticism is invited as to whether such reconnection can be a sufficient condition of the justification of such a response. For suppose that **A,B,C** and **D** are measures which effect reconnection of a wrongdoer with the correct values. Further, suppose **A** is a hundred times more costly than **B** and **D**, while **B** is a thousand times more painful than **D** and **A**. Surely under such circumstances, where use of either **A,B,C** or **D** individually is equally possible we would not suppose that one is justified in using **A** and **B** where **D** was possible. In so far as there is a duty of beneficence one has a duty to be economical with resources and the infliction of suffering. But in that case it might be argued that it was wrong to use **A** or **B** where **D** was possible. Yet surely use of **A** or **B** could not be justified and morally wrong. Hence it looks as if reconnection with correct values cannot be the sole justification of a response to wrongdoing. In other words the criteria for the assessment of a putative

response to wrongdoing must include more than that of whether the response effects reconnection with correct values. So it is dubious that reconnection with correct values is a sufficient condition of a response to wrongdoing being justified.

Could something like retributive punishment be justified by Nozick's idea of reconnection? To do this we would have to show that only retributive punishment transmits the correct values qua correct values such that they have a significant effect on the life of the wrongdoer, and hence only this punishment effects reconnection with the correct values.

Let us suppose that it is only via some procedure like punishment that the correct values are going to have a significant effect on the wrongdoer's life. So it remains to be shown that conditions (i), (ii) and (iii) which are necessary (and sufficient) for reconnection only obtain if we punish retributively.

I think, however, that we may permit ourselves an alternative here. If it can be shown that retributive punishment effects the most economical reconnection then retributive punishment will have been justified. Briefly, then, we seek to show either:

1. only retributive punishment meets conditions (i), (ii) and (iii) or

2. retributive punishment, of all the responses which meet (i), (ii) and (iii), is the most economical.

Let us consider 1 first. Certainly punishments which, by our definitions, are non-retributive will meet condition (iii). For example, the infliction of distress which is far too great for retributive equivalence to obtain might well be expected to meet (iii). If our aim is simply to make people less happy with their misdeeds then savage punishment is presumably equally likely to effect this as retributive punishment is. We shall permit ourselves the assumption here that an individual is less glad with something he did at $t1$ at a later time $t2$ if, as a result of what he did at $t1$ he experiences distress at $t2$. (Strictly speaking this won't pass muster as a sufficient condition of the attenuation of the gladness felt at the performance of some action. For example, having fallen in love and lost one may feel that, despite considerably disvaluable experiences which resulted from the loss, one was nonetheless no less glad at having fallen in love than one would have been had the whole affair had a more happy ending.)

Does non-retributive punishment fail to meet condition (i)? According to Nozick (i) is met only if the wrongdoer is punished because of and in virtue of the wrongness of his act. But this suggests that Nozick sees 'because of' and 'in virtue of' as having different meanings here. Let us try to make sense of the distinction at work.

Suppose the boss of a company dislikes **x** in virtue of **x**'s good looks and charm. However the rules of the company do not allow him to dismiss **x** for being good-looking so he waits for his chance and dismisses him instead for his scruffy attire. Here we might say that **x** was dismissed because of but not in virtue of his scruffy attire. While what follows 'because of' must be some violation of a rule sufficient for the action taken, in this case dismissal, what follows 'in virtue of' must be the reason or motivation for whatever action is taken.

Can we apply this sort of distinction to the case of punishment? The distinction at work seems to be one between which rule permits the taking of some action and which motivation is at work in the taking of it. Now if we punish non-retributively does it follow that we either fail to punish because of the wrongness of his act or in virtue of it? Clearly we may nonetheless punish non-retributively because of the wrongness of an act. A rule may specify punishment for some wrongdoing which does not satisfy the requirement of retributive equivalence just as it may specify a punishment which does meet it. So punishment, retributive and non-retributive alike, appears theoretically capable of satisfying the requirement that it be meted out because of the wrongness of an act. Secondly, if we are to regard that in virtue of which the punisher punishes as

his real motivation in punishing it seems no more likely that the punisher will punish in virtue of the wrongness of the act (which punishment is ostensibly for) where punishment is retributive than where it is non-retributive. In either case the punisher may really punish because he has bills to pay or because he wants, symbolically at least, to reverse the role he plays in a recurring nightmare he has in which he is soundly beaten.

A different kind of case might illuminate the distinction. Suppose the police know that x has committed a certain crime c1 but they realize that they have insufficient evidence to convict him. On the other hand they can concoct evidence to get him convicted of a crime c2 which he did not in fact commit. Perhaps here we would say that x is punished because of c2 but in virtue of c1. So we get: 'x was punished because of a' where for 'a' we may substitute any punishable offence, and 'x was punished in virtue of b' where for 'b' we may substitute something which is both a punishable offence and something which a has actually committed.

But again this does not show that retributive, but not non-retributive, punishment will meet condition (i). We may give someone the electric chair for double parking, i.e. punish them non-retributively both because of and in virtue of their double parking.

Someone may be punished by electrocution in virtue of having double parked provided that double parking is both a punishable offence and something of which a is actually guilty, and because of double parking where he is found guilty according to legal rules of that offence. It is clear that one may be punished according to legal rules for a crime where a non-retributive punishment is specified for that crime and that one may actually have committed that crime. There appears to be no reason to suppose that one may not be punished either retributively or non-retributively because of and in virtue of some wrongdoing (though this wrongdoing would have to be an offence also.) So it seems that non-retributive punishment can meet condition (i) as well as its retributive counterpart.

Other conditions for (i) to obtain are offered by Nozick. In addition to the condition that the wrongdoer be punished because of and in virtue of the wrongness of his act, it must also be the case that it be the intention of the punisher that the wrongdoer realize this and, further, that it be the intention of the punisher that the wrongdoer realize that it was his (i.e. the punisher's) intention that he (the wrongdoer) realize this.

But surely there is nothing in this to show why the capacity of the wrongdoer to appreciate either of these

things should be affected by whether the punishment is retributive or non-retributive.

Finally, in virtue of the above conditions for (i) to obtain it would seem that the following are conditions for (ii) to obtain, that is, are conditions which must be met if the wrongdoer is to understand the message being sent to him.

1. He, the wrongdoer, realizes he is being punished because of and in virtue of the wrongdoing.

2. He realizes it was the intention of the punisher that he realize this.

3. He realizes that it was the intention of the punisher that he (the wrongdoer) realize that it was his (i.e. the punisher's) intention that he realize this.

As we have seen there is no reason to suppose that 1 couldn't obtain if punishment were non-retributive. But neither is there reason to suppose that where some punishment out of all proportion to the wrongdoing is meted out that the wrongdoer couldn't realize that it was the intention of the punisher that he (the wrongdoer) realize that he was being punished because and in virtue of his wrongdoing. Lastly, even where the wrongdoer gets the electric chair for double parking why shouldn't he realize that it was the intention of the punisher that he (the wrongdoer) should realize that it was his (the punisher's) intention that he (the

wrongdoer) should realize that he was being punished because of and in virtue of his wrongdoing?

Of course, it might be objected to all three conditions that if the punishment was ludicrously disproportionate then the wrongdoer might well realize that he wasn't being punished because and in virtue of his wrongdoing but rather was being used as a scapegoat. If so, i.e. if he realized that he wasn't being punished because and in virtue of his wrongdoing, then he couldn't realize that it was the punisher's intention that he realize that he was being punished because and in virtue of his wrongdoing. Just as I can't intend that someone know something false I can't intend that they realize something which is not the case. So condition 2 wouldn't hold either. But on the same grounds 3 couldn't hold either. To see this we need only remark that if it is not the intention of the punisher that the wrongdoer realize that he is being punished because and in virtue of his wrongdoing then the clause in square brackets in the following is false:

He realizes that it was the intention of the punisher that he realize [that it was the intention of the punisher that he realize this.]

But if the bracketed clause is false then, by the application of the same considerations concerning the concept 'to realize' he, the wrongdoer, cannot realize

that it was the intention of the punisher that he
realize the content of the bracketed clause. Hence
where 1 doesn't obtain in the manner described neither
do 2 or 3. We have a kind of epistemic house of cards
or domino effect.

But why should we suppose that the wrongdoer
appreciates that he is not being punished because and
in virtue of his wrongdoing just because he is not
punished retributively?

It could be that he is punished non-retributively
against the backdrop of a punishment system in which
retributive punishment is the norm. But this will not
show that non-retributive punishment only won't meet
conditions 1, 2 and 3. For we may imagine a punishment
system in which non-retributive punishment is the norm
so that the occasional retributive punishment alerts
the punishee to the fact that he is not being punished
because of and in virtue of his wrongdoing. Nor can it
be correct to suppose that because non-retributive
punishment will be haphazard there would be no way of
distinguishing the retributive punishment from the
rest. The non-retributive punishment employed normally
as part of the practice of the system of punishment in
question might take a regular form. For example,
punishment would be non-retributive where we punish
extremely severely for a minor offence (i.e. minor in
respect of its moral gravity,) and decreasingly

severely the more heinous the moral nature of the offence. Surely the wrongdoer could distinguish a retributive punishment against this sort of backdrop (though not, perhaps, in the middle range of cases.)

So it is not evident that it is only in the case where we punish non-retributively that the wrongdoer may realize that he is not being punished because and in virtue of his wrongdoing. So it is equally not evident that only where we fail to punish retributively conditions 1, 2 and 3 are not met. So it hasn't been shown that condition (ii) that the wrongdoer understand the message is not met specifically where punishment is non-retributive. But then it has neither been shown that where we punish non-retributively rather than retributively our response fails to meet conditions (i), (ii) and (iii). This means that we have failed to show that non-retributive punishment, as opposed to retributive punishment, fails to effect a reconnection of the wrongdoer with the correct values.

So we have failed to show that we may justify retributive punishment on the grounds that it alone reconnects the wrongdoer with the correct values.

And what of the alternative 2, viz. that we show that, of the responses which effect reconnection, retributive punishment does the job more economically?

I shall only consider retributive punishment in the light of suffering. Is it true that non-retributive punishment is less economical of suffering?

Traditionally, there are difficulties involved in trying to show that, for example, what is right from the utilitarian viewpoint in respect of punishment just is in accordance with certain criteria for retribution. To take a topical case it is often felt that the sentence for rape should be equal to that for murder, on the grounds that rapists have a high tendency to reoffend. By keeping a rapist mandatorily in prison for life the benefits to those at risk outweigh the disadvantages to those rapists who would be reformed by punishment less than that for murder and so would receive less punishment overall. There are others who, however, argue that it is foolhardy to make punishment for stealing a lamb the same as that for stealing a sheep. If a rapist is to get a life sentence anyway then he may be encouraged to kill his victim since this might make his detection less likely. The point to be made here is that neither argument is about what retributive equivalence requires. Starting from something like a retributive requirement that (not unarguably) murder is always worse than rape, and hence that the punishment for the former should always be more or at least generally be more unpleasant than for the latter, we move on to the discussion of a cost

benefit analysis of suffering and well-being. It is far from clear that, given the serpent-winding nature of the considerations in play in this sort of discussion, we would independently come up with retributive punishments. Moreover in some cases, e.g. strict and vicarious liability, and, at one time at least, the punishment for having intercourse with a girl under the age of fourteen, we effectively pronounce that the best results don't always come from adherence in all cases to the requirements of something like retributive equivalence.

So if we were to have retributive and non-retributive punishment systems which effected reconnexion with the correct values equally the decision as to which is most economical with regard to suffering would be estimated by reference to consequences other than reconnexion. But such discussion in itself usually leaves it open whether retributive or non-retributive punishment is the most economical in respect of suffering. So we can't justify retribution by means of this second stratagem.

(5) Retribution and the expressive function of
punishment.

Feinberg (Doing and Deserving, 1970, Chapter 5,) proposes a different sense in which
punishment must fit the crime. Punishment is distinct
from penalty in so far as the former but not the latter
has a condemnatory aspect. By punishing an individual
we do not merely subject him to hard treatment, i.e.
the imposition of pain, but we also express our
disapproval of what he has done. For punishment to fit
the crime the disapproval expressed by that punishment
must be such that it is in accord with the requirement
of justice that the greater the crime the greater the
disapproval of it expressed by punishment of it. In
this context a crime is to be assessed by reference to
the amount of harm it generally causes and the degree
to which people are disposed to commit it. But it is
merely a convention that we punish people by harder
treatment, to use Feinberg's expression, when we wish
to express greater disapproval.

Consider first Feinberg's idea that how great a crime
is is to be determined by reference to two factors: the
harm that it generally causes and the degree to which
people are disposed to commit it. We might look at this
principle in connexion with the crimes of murder and
manslaughter.

Certainly the harm caused by individual crimes of murder will vary - there is a difference in harm between murdering one and murdering a dozen people. It is for this reason no doubt that Feinberg speaks of the harm generally caused by a crime. But how are we to determine the harm generally caused by, say, murder? One suggestion is that we consider the kind of harm which generally results from the crime. In the case of murder we necessarily have the harm resulting from homicide. But then if we make the simplifying assumption that this is the only kind of harm generally caused by murder (what of the harm to those close to the murder victim, for example?) there seems no reason to suppose that the harm generally caused by manslaughter is any different, since the distinctive harm resulting from the latter crime is that of homicide also.

Perhaps manslaughter and murder are different in respect of the degree to which people are disposed to commit them. Suppose x is more disposed to do A than B if, in most cases, x would prefer to do A rather than B. But then one way of understanding the claim that people are more disposed to do A than B is that most people in most cases would prefer to do A than B. But is it true that in most cases most people would prefer to commit murder rather than manslaughter?

One ready answer here is that, given that the punishment for murder is greater than that for manslaughter and that most people are rational, most people in most cases would probably prefer to commit murder rather than manslaughter. One generally prefers to suffer less rather than more. But the problem with this reply is that it can only be made where we already have a means for determining which crimes are greater. What we have here is the conditional: if murder is more serious than manslaughter then people are more disposed to commit the latter than the former. But if we know that in respect of the harm factor murder and manslaughter are equal, the consequent of this conditional can be used, given this background equality of harm factor and the Feinberg principle for determining the seriousness of a crime, as the antecedent of the following conditional: if people are more disposed to commit manslaughter than murder then manslaughter is more serious than murder. Hence, by the additional assumption that Feinberg regards greater disposition to commit A as a reason for supposing A a more serious crime, we derive an absurdity. But, just as important, in order to determine the degree of disposition to commit a crime, according to the first conditional we need to know antecedently how serious that crime is. But then, if Feinberg's principle is a way of determining seriousness of crime it is self-

defeating, and if it is a way of explaining why some crimes are more serious than others since the explanandum figures in the explanans, this way is unilluminating.

So it does not look as if the method proposed by Feinberg for determining the seriousness of a crime is adequate. Certainly it will not show that murder is a more serious crime than manslaughter, which is at least plausible, since when applied to such crimes it either draws us into absurdity or into falsehood. Moreover if we were to ignore this absurdity and use Feinberg's principle it would follow from what he has to say about punishment that, given our conventions for the expression of disapproval and the assumption that we punish more harshly for murder than for manslaughter, we ought to punish more harshly for manslaughter than murder.

Secondly, what are we to make of the claim that if the crime c1 is more serious than the crime c2 it is just that we express greater disapproval of c1 than c2? Let us suppose that the 'we' here refers to society in general. Further, I shall suppose that 'to disapprove of **x**' is to have an unfavourable attitude towards **x** in virtue of some moral quality of **x**.

Now the first thing which we may infer from the fact that c1 is more serious than c2 and that we have a duty to express more disapproval for the more serious crime

is that we have a duty to express disapproval of c1 at
least. The reason for this is that we cannot express
less disapproval than no disapproval at all. But then
even if we suppose that we have a duty to express no
disapproval of c2 we will have a duty to express some
disapproval of c1 for only some disapproval, however
little, will be more than no disapproval. So if the
claim about the justice of expressing proportional
disapproval, call this PD, is true, then in the case of
c1 we have, i.e. society has, a duty to express an
unfavourable attitude that we have towards c1, where we
have this attitude in virtue of some moral quality of
c1. But is it the case that, in respect of some crimes
at least, we have a duty to express disapproval
thereof?

One possibility is that since, of all the options,
expressing disapproval will result in fewer crimes
being committed than other options, we have a duty to
express disapproval. But this is hardly plausible -
society might express its disapproval in all kinds of
ways other than punishment. But then it is unlikely
that expression of disapproval simpliciter will be
justified. Secondly, allied to this claim, we have the
further problem of showing that proportional expression
of disapproval would have optimal results. This would
be difficult. For example, the maxim 'one might as well
be hanged for a sheep as a lamb' cannot be used to

reinforce the idea that it is advisable to have
proportional disapproval. Even if we assume that having
proportional disapproval will mean having proportional
punishment, in an environment in which we punish as
severely for flashing as for murder this would only
make the option of murder over flashing rational if we
wanted to murder more than we wanted to flash. But
there is no reason to suppose, as we have already
noted, that we may legitimately infer from the greater
moral gravity of some offence that such an offence is
one which we are more disposed to commit.

Of course, it might be suggested that only punishment
as an expression of disapproval is likely to have the
effect of minimizing crime. But this shows that we have
a duty to punish (albeit not necessarily
proportionally,) not that we have a duty to express
disapproval.

Yet it might be objected that punishment to which we
somehow add the expression of disapproval is more
effective in reducing crime than 'mere' punishment.
This may be so. But as a means of deriving the duty to
express disapproval this is only valuable if it has
been shown that a necessary condition of minimizing
crime is expressing disapproval in some way. But
clearly showing that punishment plus disapproval is
more effective than punishment alone doesn't establish
this. The former is consistent with its being a good

idea to abandon punishment and disapproval entirely. So we cannot argue from the beneficial consequences of expressions of disapproval to the duty to express disapproval.

A different line of reasoning would be that we have a duty to express disapproval of c_1 because, if we were not to, this would entail that we didn't regard c_1 as wrong.

Firstly, if the above is true so is its contrapositive in the form: if we regard c_1 as morally wrong then we express disapproval of c_1. Yet this is clearly false for one may regard something as morally wrong and yet deliberately refrain from any expression of disapproval of it - in order not to lose a friend perhaps or be thought illiberal by a peer. It is more likely that the conditional: 'if we regard x as morally wrong then we disapprove of x' is true but this of course only yields the contrapositive to the effect that failure to disapprove, as opposed to failure to express disapproval, of c_1 entails that we don't regard c_1 as morally wrong.

Secondly, the general thrust of this argument is that we ought to express disapproval of c_1 because we ought to regard what, for the sake of argument, is morally wrong as being morally wrong. But this duty to regard what is wrong as wrong is not self-evident. If we suppose that we have a duty to refrain from doing what

is wrong we clearly can't infer from this that we have a duty to regard the wrong action as wrong. Our duty is not to do it, not to entertain certain ideas about it or attitudes towards it. Moreover, even if it could be shown that by regarding wrong actions as wrong we minimize their occurrence it still has to be shown why it is that what is arguably not an overall duty for an individual should be one for society itself. The duty to maximize good or minimize evil is not one which unarguably applies to individuals at all times. If it were to then there would be no supererogatory actions. It has to be shown why society must always do whatever maximizes good.

Yet even if the more serious crimes, according to the Feinberg principle for determining seriousness, were morally graver than the less serious ones, and it was shown that it is a requirement of justice that we express greater disapproval the more serious the crime it would only follow that we ought to adopt a certain convention in accordance with PD (the principle that it is just to express proportional disapproval.) As it stands our convention is to express proportional disapproval by punishment. But we could choose whatever convention we liked. As far as the Feinberg account is concerned the particular convention we adopt is morally irrelevant. Thus we would be acting justly if henceforth we adopted a new convention capable of

expressing proportional disapproval. We might adopt the convention of clapping to express disapproval, expressing such disapproval proportionally by clapping more the more serious the crime. Yet surely our intuitions are that such a practice, even given that clapping did not have its existing associations and conventional expressiveness, would be unjust.

Does the Feinberg account offer some justification for a practice of punishment where this conforms to either of the proposed concepts of retributive punishment?

Consider retributive punishment in which the maxim must be observed that the greater the immorality of the crime the more distressing to the punishee must be the punishment for it. Clearly the Feinberg account needn't justify such a practice. The Feinberg principle for determining the seriousness of a crime yields either absurd or counterintuitive results. If it does the former then the account is incomplete anyway; if the latter then, to take our earlier example, one who is guilty of murder has committed a less serious crime than one guilty of manslaughter. But given existing conventions then according to Feinberg and using his terminology punishment for the latter should be harder than that for the former. Clearly this is not in accordance with what we have proposed as a necessary condition of retributive punishment.

Alternatively, retributive punishment must be greater the more blameworthy the punishee. But if we suppose that a charge of murder may be reduced to manslaughter on the grounds of diminished responsibility, or because the unlawful homicide was not an act likely to cause harm of a grievous nature, or was provoked or the result of gross negligence, it is plausible that, in some cases at least, less blameworthiness attaches to one guilty of manslaughter than to one guilty of murder. But then again in proposing harder punishment for the one guilty of manslaughter than for one guilty of murder Feinberg is putatively justifying punishment which is not retributive, and not justifying punishment which is retributive.

So it appears that the Feinberg account as it stands cannot provide justification for retributive punishment on either of our definitions thereof.

(iv) Conclusion.

A necessary condition of retributive punishment, as
we are understanding it, is that it meet the condition
SP; a further condition is that the maxims of crude
proportionality be met also. However, it has not been
shown that a system of punishment must meet even the
first of these conditions for it to be fair (Morris,)
for it to be deserved or a right of either the offender
or victim (Reiman,) for it to meet a requirement of
equality, or for it to be justified as the most
economical reconnexion of offenders with the correct
values or as the appropriate expression of moral
condemnation. It does not follow from this, of course,
that retributive punishment is not fair, a right of the
victim, etc..

A different problem concerns what might be referred
to as the moral exhaustiveness of the concept of
retribution we are using. What I mean by this is that
it is a requirement of SP that the more morally wrong
an offence the greater must be the punishment for it.
But then it follows, of course, that if we decide not
to punish venial offences we are not punishing

retributively since there will be some pair of
immoralities for which we do not punish such that
though the first is graver than the second, our
punishment for the first is no greater than that for
the second. Now it may be that it would constitute an
intolerable infringement of liberty to punish for, say,
failure to keep a dental appointment; but in that case
it might be argued that a system of retributive
punishment which did meet SP would be ruled unjust
anyway - it would simply be a non-starter. I shall not
explain how it could be done but I hope it is fairly
obvious how SP, and indeed the maxims of crude
proportionality, could be modified to describe relevant
conditions of a retributive punishment system where
only offences of a certain moral gravity were
punishable. However, this point does not affect the
status of the putatively justificatory accounts
considered since our discussion has concerned how such
accounts would enable us to deal with offences which
presumably would have to be punishable if one had
decided to have a punishment system at all.

Finally it is worth noting again that an account of
the justification of retributive punishment, as we are
understanding this latter, must meet the maxims of
crude proportionality. Now while any account which
meets these conditions necessarily meets SP the
converse is not true. Of course, for the reasons

already given, it seems that we can never know that an actual system of punishment meets these maxims of crude proportionality, but one cannot infer from this the impossibility of finding an account which would show why it is important that such conditions be met. All that could be inferred was that even if such an account were forthcoming one could never know that one's actual punishment system was retributive in the desired way. For one thing, if retributive punishment is the only justified form of punishment we could never be sure that some system of punishment actually employed was justified.

It might be wondered whether such a lengthy articulation of a principle, as the retributive principle has here received, should precede attempts to justify it. After all, surely it may be that even a punishment system which meets SP has no justification. In the final analysis only a satisfying justificatory account of retribution, as conceived here, can properly answer this doubt. But even if such an account were forthcoming one must not expect it to show that retributive punishment is always the punishment to be preferred. One can surely expect this prima facie retributive principle to be outweighed by other considerations. Rather we are entitled to seek, given the strength of the intuitions regarding retributive proportionality between punishment and offence, some

model which shows why we have such a prima facie principle, and why, in a morally imperfect world, it may be that this principle is outweighed. It is this model which is conspicuously absent from the accounts considered.

Bibliographical References

Bedau, H.A. Classification-based Sentencing: Some
Conceptual and Ethical Problems in Pennock, J.R. and
Chapman, J.W. (Eds) Nomos XXVII, (New York University
Press, N.Y. 1985) Part 2, Chapter 4.

Benn, S.I. An Approach to the Problems of Punishment
(Philosophy 1958)

Bentham An Introduction to the Principles of Morals
and Legislation (Hafner, N.Y. 1948,) Chapter 14, Secs
1-26, taken from Ezorsky, G. (Ed.) Philosophical
Perspectives on Punishment (N.Y.U.P. Albany, 1972.)

Cooper, D.E. Hegel's Theory of Punishment, in
Pelcynski, Z.A. Hegel's Political Philosophy: Problems
and Perspectives (C.U.P. Cambridge, 1971)

Cottingham, J.G. Varieties of Retribution
(Philosophical Quarterly, 1979)

Davis, M. - Why Attempts Deserve Less Punishment than
            Completed Crimes (Law and Philosophy, 5,
1986)
          - How to Make the Punishment Fit the Crime
            (Ethics, 93, July 1983,)
          - Setting Penalties: What does Rape Deserve?
            (Law and Philosophy, 3, 1984,)

Duff, R.A. Auctions, Lotteries and the Punishment of
Attempts (Law and Philosophy, 9, 1990,)

Ewing, A.C. The Morality of Punishment (Kegan Paul,
London, 1929)

Feinberg, J. Doing and Deserving (Princeton University
Press, Princeton, N.J. 1970) Chapter 5: The Expressive
Function of Punishment

Fowler, H.W. and Fowler, F.G. (Eds) The Concise Oxford
Dictionary (Clarendon Press, Oxford, 1972)

Goldman, Alan The Paradox of Punishment (Philosophy and Public Affairs, 9, 1979.)

Goldman, Alvin A Theory of Human Action (Princeton University Press, Princeton, N.J., 1970)

Hampton, J. A New Theory of Retribution in Liability and Responsibility (C.U.P. New York,

1991) Chapter 11

Hart, H.L.A. Prolegomenon to the Principles of Punishment: Punishment and Responsibility

(Clarendon Press, Oxford, 1968)

Honderich, T. Punishment: The Supposed Justifications (Polity Press, Cambridge, 1989)

Kant, I. The Philosophy of Law, Pt II, translated by W. Hastie (T.T. Clark, Edinburgh, 1887, pp 194-198) taken from Ezorsky, G. (op. cit.)

Kaufman, A. Mr. Quinton on Punishment (Analysis, October 1959)

Mabbot, J.D. Punishment (Mind, April 1939)

Mackie, J.L. Persons and Values (Clarendon Press, Oxford, 1985, pp 207-8) cited in Ten, C.L. Positive Retributivism (op. cit.)

Michael, M.A. Utilitarianism and Retributivism: What's the Difference (American Philosophical Quarterly, 29ii, April 1992,)

Moore, G.E. Principia Ethica (C.U.P. Cambridge, 1903) cited in Ewing, A.C. (op. cit.)

Morris, H. Persons and Punishment (The Monist 1968)

Nino, C.S. A Consensual Theory of Punishment (Philosophy and Public Affairs 12(4) 1983)

Nozick, R. Philosophical Explanations (O.U.P. Oxford, 1981 pp363-397)

Quinton, A.M. On Punishment (Analysis June 1954)

Rawls, J. Legal Obligation and the Duty of Fair Play in Hook, S.J. Law and Philosophy (New York University Press, N.Y. 1964)

Reiman, J.H. Justice, Civilization and the Death Penalty: Answering Van Den Haag (Philosophy and Public Affairs, 14ii Spring 1985)

Ross, W.D. The Right and the Good, Chapter 2, (Clarendon Press, Oxford, 1930)

Royal Commission on Capital Punishment, Cmd 8932, para. 53, 1953, cited in Hart, H.L.A. (op. cit.)

Sher, G. Desert (Princeton University Press, Princeton N.J. 1987)

Ten, C.L. - Crime, Guilt and Punishment (Clarendon Press, Oxford, 1987)
         - Positive Retributivism in Paul, E.F. et al. (Eds) Crime, Culpability and Remedy (Blackwell, Oxford, 1990)

Walker, N. Why Punish? (O.U.P. Oxford, 1991)