In Defence of a Utilitarian Theory of Punishment: Punishing the Innocent and the Compatibility of Utilitarianism and Rights

MIRKO BAGARIC*

Retributivism has replaced utilitarianism as the main philosophical justification of punishment. There have been two main reasons for this. Utilitarianism commits us to the supposedly unthinking practice of punishing the innocent and, more generally, is thought to be inconsistent with the concept of individual rights. It is argued that both of these criticisms are unpersuasive. Punishing the innocent is no worse than other acts or practices which we condone in extreme situations, and hence does not so trouble our sensibilities that it is justifiable to conclude that any theory which approves of such an outcome must be flawed. Further, while rights are now the conventional moral currency, on critical evaluation, non-consequentialist moral theories (which underpin most retributive theories) are unable to justify the foundation and existence of rights. In fact, utilitarianism is the only moral theory which can provide a firm basis for rights, including the right of innocent people not to be punished.

Introduction

The Current State of the Punishment Debate and Sentencing Practice

Retributivism, under the banner of just deserts, has replaced utilitarianism, at least ostensibly,[1] as the prime philosophical underpinning of

---

* BA, LLB (Hons) (Monash), LLM (Monash); Lecturer, Faculty of Law, Deakin University. I am grateful to Professor Chin Liew Ten and wish to thank him for his many valuable and thoughtful comments and suggestions regarding earlier drafts of this paper.

1 I have argued elsewhere that in reality a utilitarian theory of punishment still best fits the relevant sentencing factors (at least in Victoria): see M. Bagaric, (1997) 21 (4) "The Disunity of Confiscation and Sentencing" *Criminal Law Journal* 191.

punishment,[2] in the Western world. It is also generally perceived that the philosophical leaning towards retributivism has permeated most sentencing systems,[3] despite the gulf that normally exists between theories of punishment and sentencing practice and the tendency of the sentencing systems of most jurisdictions to not adopt a primary rationale for sentencing.[4] In this regard, the comments of Andrew Ashworth about a decade ago have proved prophetic:

> Across the common law world and elsewhere, new sentencing systems are being introduced or recommended. For example, Sweden, the U.S. federal jurisdictions and several American states have already begun to operate new sentencing schemes, and there are important proposals on the table in Canada, the State of Victoria and the Australian federal jurisdiction... In planning a new system it is necessary to think seriously

---

2 For an overview of the academic and social trends in punishment, see R.A. Duff and D. Garland, "Introduction: Thinking about Punishment" in R.A. Duff and D. Garland (eds), *A Reader on Punishment* (Oxford University Press, Oxford, 1984) 1, 8-16; A. von Hirsch, *Past or Future Crimes* (Rutgers University Press, New Jersey, 1985), ch 1; N. Walker, *Why Punish?* (Oxford University Press, Oxford, 1991), ch 1; A. Ashworth, *Sentencing and Criminal Justice* (2nd edn) (Butterworths, London, 1995), 69-72.

3 The revival of retributivism is due in a large part to the work of A. von Hirsch, particularly *Doing Justice: The Choice of Punishments* (Hill and Wang, New York, 1976); A. von Hirsch, *Past or Future Crimes* (Rutgers University Press, New Jersey, 1985). In the United States the just deserts model was responsible for the move away from wide discretionary sentencing powers to laws aimed to promote greater certainty and consistency in sentencing, such as the Minnesota guidelines. Legislation in Washington and Oregon also expressly adopts a just deserts based philosophy. See also, A Ashworth, above n 2, ch 13.

4 The unprincipled nature of sentencing practice has led to what Ashworth labels a 'cafeteria system' of sentencing, which permits sentencers to pick and chose a rationale which seems appropriate at the time with little constraint: A. Ashworth, above n 2, 331. This is made significantly easier by the large number of discrete factors that the courts have identified as being relevant to sentencing. Two separate studies, about twenty years ago, determined that there were between 200 and 300 hundred such factors: J Shapland, *Between Conviction and Sentence* (Routledge & Kegan Paul, London, 1981), 55, identifies 229 factors, while R Douglas, in *Guilty, Your Worship* (Melbourne, LaTrobe University, 1980) in a study of Victorian Magistrates' Courts identified 292 relevant sentencing factors. The results of such studies were noted in *Pavlic v R* (1995) 5 Tas R 186, 202, where it was stated that 'it is impossible to allocate to each relevant factor a mathematical value, and from that, extrapolate a sum which determines the appropriate penalty.'

> about the purposes of sentencing, and it is at this stage that the "just deserts" approach has been influential in many of the jurisdictions mentioned.[5]

However, this has not long been the case. Only a few short decades ago, Mabbott stated that 'in the theory of punishment, retribution has been defended by no philosopher of note [for over fifty years] except Bradley. Reform and deterrence are the theories accepted in principle and increasingly influential in practice'.[6] In the 1975 Victorian decision of *Rv Williscroft*, Starke J stated that 'retribution as an element of punishment has by now, in my opinion, disappeared, or practically disappeared from our criminal law... Reformation should be the primary objective of the criminal law'.[7]

This paper considers the main theoretical attacks on utilitarianism which have resulted in a decline in the support for a utilitarian theory of punishment and concludes that they have been unduly persuasive. Before turning to substantive matters, I shall first outline the main criticisms of the utilitarian theory of punishment and then briefly attend to some house keeping in the form of clarifying some definitional matters.

Reasons for the Movement away from Utilitarian Punishment

Broadly, there have been two main reasons for the movement away from utilitarianism as the prime theory of punishment since about the 1970s: one is pragmatic and the other theoretical.

Pragmatic Reasons

The first problem was the perceived failure of penal practice and the treatment based goals of sentencing to measure up to the prime utilitarian objectives of deterrence[8] and rehabilitation.[9] Research findings relating to

5 A. Ashworth "Criminal Justice and Deserved Sentences" (1989) *Criminal Law Review* 340.

6 As cited in K.G. Armstrong, "The Retributivist Hits Back" in S.E. Grupp (ed.), *Theories of Punishment* (Indiana University Press, Ontario, 1971), 19-20.

7 *R v Willscroft* [1975] VR 292, 303-4.

8 For an overview of the literature on deterrence see J Q. Wilson, "Penalties and Opportunity" in R.A. Duff and D. Garland (eds), *A Reader on Punishment* (Oxford University Press, Oxford, 1994), 177, where he argues that the main factor relevant to deterrence is not the penalty level, but rather the perceived probability of apprehension. This does not necessarily diminish the importance of punishment. The likelihood of being caught is only undesirable because of the accompanying realisation that punishment

rehabilitation, in particular, were at one point so depressing, that a 'nothing works' attitude was pervaded.[10] Given the apparent failure to achieve such lofty and ambitious sentencing goals, the natural inclination was to set the sights on aims which were far more achievable. Future orientated goals of punishment, such as rehabilitation and deterrence, made way for backward looking considerations where the main goal was to ensure that criminals got what they deserved. Thus the aim of doing more good through the prison system was replaced by the goal of doing justice, where justice broadly equated to imposing punishment that was proportionate to the severity of the crime.[11] On this rationale, so long as the punishment fitted the crime, or was thereabouts,[12] the sentencing system was a 'success', irrespective of the

---

may follow. See also T. Tyler, *Why People Obey the Law* (Yale University Press, New Haven, 1990), 107, 175-6, where following a 1984 study of about 1 500 people who lived in Chicago about their contact with legal authorities, Tyler noted that normative issues are closely linked with compliance with the law. People do not merely obey the law because it is in their self-interest to do so, but also because they believe it is proper to do so.

9 A.E. Bottoms, "An Introduction to the Coming Crisis" in A.E. Bottoms and R.H. Preston (eds), *The Coming Penal Crisis: A Criminological and Teological Exploration* (Longwood Publishing, Edinburgh, 1980), 1.

10 R. Martinson, "What Works?—Questions and Answers About Prison Reform" (1974) *The Public Interest* 22. Following research conducted between 1960 and 1974, Martinson initially noted that empirical studies had not established that any rehabilitative programmes had worked in reducing recidivism. Martinson, however, softened his position several years later, concluding that some types of rehabilitation programmes, particularly probation parole, may be effective and that generally 'no treatment...is inherently either substantially helpful or harmful. The critical factor seems to be the conditions under which the program is delivered': R Martinson, "New Findings. New Views: A Note of Caution Regarding Sentencing Reforms" (1979) *Hofstra Law Review* 243, 254.

11 For example, see S. Cohen, *Visions of Social Control* (Blackwell, 1985).

12 Given the difficulties in defining the factors that are relevant to proportionality approximate just deserts is possibly the most that can be hoped for. For a discussion regarding the considerations relevant to proportionality, see A. von Hirsch and N. Jareborg, "Gauging Criminal Harm: A Living Standard Analysis" (1991) 11 *Oxford Journal of Legal Studies* 1. They state that there are several steps involved in gauging the seriousness of an offence. The first involves an appraisal of the types of interests which the paradigm instance of an offence violates or threatens to infringe upon. They identify four basic types of interests. In order of most to least important, they are physical integrity; material support and amenity (such as nutrition and shelter); freedom form humiliating or degrading treatment; and privacy and autonomy. Next it is necessary to determine the effect that violating the relevant interests typically has on the victim. Finally, it is necessary to consider the offender's culpability and the remoteness of the harm.

indirect consequences stemming from it. Retributivism was the clear beneficiary of such an approach.

Theoretical Attacks on Utilitarianism

The decline of utilitarian punishment and sentencing was also greatly accelerated by the fact that at the theoretical level there was a move towards rights based moral theories and widespread support for arguments that utilitarianism commits us to abhorrent practices, such as punishing the innocent. This paper shall focus on these theoretical attacks.

The main general argument in support of rights based moral theories is aptly stated by John Rawls who claims that only rights based theories take seriously the distinction between human beings, and protect certain rights and interests that are so paramount that they are beyond the demands of net happiness.[13]

Charges of this nature have been extremely influential. In the last half of the century, following the Second World War, there has been an immense increase in 'rights talk',[14] both in sheer volume and the number of supposed rights. The rights doctrine has progressed a long way since its original rather uncomplicated and noble aim of providing 'a legitimisation of...claims against tyrannical or exploiting regimes'.[15] There is now, more than ever a strong tendency to advance moral claims and arguments in terms of rights.[16] Assertion of rights has become the customary means to express our moral sentiments: 'there is virtually no area of public controversy in which rights are not to be found on at least one side of the question—and generally on both'.[17] The domination of rights talk is such that it is accurate to state that 'the doctrine of human rights has at least temporarily replaced the doctrine of maximising utilitarianism as the prime

13 J. Rawls, *A Theory of Justice* (Clarendon Press, Oxford, 1972).

14 By rights talk I also include the abundance of declarations, charters, bills, and the like, such as the *Universal Declaration of Human Rights* (1948); the *International Covenant of Economic, Social and Cultural Rights* (1966); and the *European Convention for the Protection of Human Rights and Fundamental Freedoms* (1966), that seek to spell out certain rights. There were numerous declarations, and the like, of rights prior to the Second World War, such as, the *Declaration of Independence of the United States* (1776) and the *Declaration of the Rights of Man and Citizens* (1789), however it is only in relatively modern times that such documents have gained widespread appeal, recognition and force.

15 S.I. Benn, "Human rights—For Whom and For What?", in E. Kamenka and A.E. Tay (eds), *Human Rights* (Edward Arnold, Melbourne, 1978) 59, 61.

16 Almost to the point where it is not too far off the mark to propose that the 'escalation of rights rhetoric is out of control': L.W. Sumner, *The Moral Foundation of Rights* (Clarendon Press, Oxford, 1987), 1.

17 Id.

philosophical inspiration of political and social reform'.[18]

The narrower theoretical objection to utilitarian punishment: that it permits punishment of the innocent, has been so persuasive that it alone has led many to reject utilitarianism as a general theory of morality.[19] The real force of this objection is found in the more general criticism that utilitarianism fails to protect basic individual rights and interests, and since it does not prohibit anything per se may lead to horrendous outcomes. However, in light of the amount of ground that critics of a utilitarian theory of punishment have taken with the punishing the innocent criticism, I shall consider it[20] first before discussing the more general objection that utilitarianism is inconsistent with individual rights and interests. But first I shall attend to definitional matters.

Definitional Issues

Retributive Theories of Punishment

A vast array of theories of punishment have been advanced that are classified as retributive.[21] Due to the diversity of these theories, it has proven remarkably difficult to isolate a distinctive feature of theories carrying the tag.[22] All retributive theories assert that offenders deserve to suffer, and that the institution of punishment should inflict the suffering they deserve, however, they provide vastly divergent accounts of why criminals deserve to suffer.[23]

---

18 H.L.A. Hart, *Essays in Jurisprudence and Philosophy* (Clarendon Press, Oxford, 1983), 196-7.

19 For a historical account of how punishing the innocent has become textbook commonplace on which rejection of utilitarianism is based, see F. Rosen, "Utilitarianism and the Punishment of the Innocent" (1997) 9(1) *Utilitas* 23.

20 A related problem for the utilitarian is that of exemplary punishment. However most of the moral difficulties raised by exemplary punishment are even more acutely invoked by the practice of punishing the innocent, hence will not be specifically addressed in this paper.

21 For an overview of many of the theories, see C.L. Ten, *Crime, Guilt and Punishment* (Clarendon Press, Oxford, 1987), 38-65; J. Cottingham, "Varieties of Retributivism" (1979) 29 *Philosophical Quarterly* 238; T Honderich, *Punishment: The Supposed Justifications* (rev ed, Penguin Books, Harmondsworth, 1984), 211.

22 See T. Honderich, *Ibid*, 211; D. Dolinko, "Retributivism, Consequentialism, and The Intrinsic Goodness of Punishment" (1997) 16 *Law and Philosophy* 507.

23 See A. Duff and A. von Hirsch, "Responsibility, Retribution and the 'Voluntary': A Response to Williams" (1997) *Cambridge Law Review* 103, 107.

Despite this, it has, somewhat ambitiously, been claimed that there are broadly three similarities which underlie retributive theories.[24] The first is that only those who are blameworthy deserve punishment and that this is the sole justification for punishment. Thus punishment is only justified in cases of deliberate wrongdoing.[25] This feature however, does not justify the institution of punishment; rather it acts as a constraint on the circumstances in which punishment may be administered, and fails to justify the link between crime and punishment.[26]

The second is that the punishment must be equivalent to the level of wrongdoing.[27] This is a claim enthusiastically endorsed by Andrew von Hirsch, one of the main contemporary proponents of retributivism, who asserts that

> sentences according to [the just deserts] theory are to be proportionate in their severity to the gravity of the criminal's conduct... In such a system, imprisonment, because of its severity, is visited only upon those convicted of serious felonies. For non-serious crimes, penalties less than severe imprisonment are to be used.[28]

He argues that the basis for proportionality, is essentially that 'punishment is the vehicle for condemnation and as a matter of fairness punishment must be proportionate since the severity of the sanction expresses the stringency of the blame'.[29] However, the view that punishment should be commensurate with the seriousness of the offence, does not provide a

---

24 J. Anderson, "Reciprocity as a Justification for Retributivism" (1997) *Criminal Justice Ethics* 13. As Anderson points out, all of these factors are present in what Hart refers to as 'crude retributivism': see H.L.A. Hart, *Punishment and Responsibility* (Oxford University Press, Oxford, 1963) 231-7.

25 J. Anderson, *Ibid*, 13, 14.

26 The other proposition implied in this first statement is that criminal guilt alone justifies punishment. This is made clearer in what Anderson claims is the third distinctive feature of retributivism, which is discussed below.

27 J. Anderson, above n 24. This claim is given legal expression in the form of the principle of proportionality, which is discussed below.

28 A. von Hirsch, above n 2, 10.

29 See A. von Hirsch, "The Politics of 'Just Deserts'" (1990) *Canadian Journal of Criminology* 397, 398. See also A. Von Hirsch, "Censure and Proportionality" in R.A. Duff and D. Garland (eds), above n 2, 115, 125. Elaborating on this he provides that 'were penalties ordered in severity inconsistently with the comparative seriousness of the crime, the less reprehensible conduct would, undeservedly, receive the greater reprobation': Id.

justification for punishment, rather it too, simply, acts as a restraint on it.[30] Such a claim is also not distinctly retributivist. Utilitarians have also been known to invoke the principle of proportionality. For example, Bentham argued that proportionality has a secure utilitarian foundation, and that in fact it has a central role in a utilitarian theory of punishment. Bentham asserted, as a secondary principle, that crimes should be punished in proportion to the harm done to the life and security of others in society.[31] If crimes are to be committed it is preferable that offenders commit less serious rather than more serious ones. Therefore sanctions should be graduated commensurate to the seriousness of the offence so that those disposed to crime will opt for less serious offences. Absent proportionality, potential offenders would not be deterred from committing serious offences any more than minor ones, and hence would just as readily commit them.[32]

Finally, it has been asserted that a distinctive feature of retributivism is that punishing criminals is itself just: it cannot be inflicted as a means of pursuing some other aim.[33] However, even a cursory consideration of some

30 It provides an upper and lower ceiling regarding the amount of punishment that should be inflicted for any particular offence: see R.G. Fox "The Meaning of Proportionality in Sentencing" (1994) 19 *Monash University Law Review* 489, 491.

31 J. Bentham, *Principles of Morals and Legislation* (1789), L.J. Lafleur (ed.) (7th ed, Hafner Press, New York), 178–88.

32 This argument has been persuasively criticised by von Hirsch, who points out that there is no evidence that offenders make comparisons regarding the level of punishment for various offences: A. von Hirsch, above n 2, 32.

33 J. Anderson, above n 24, 13. Some retributivist, such as Kant and Hegel, make the additional claim that punishment of wrongdoers is not only just, it is obligatory. However, it has been argued that even if punishing the guilty is an intrinsic good there is still no moral duty to bring about such a state of affairs: see D. Dolinko, "Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment" (1997) 16 *Law and Philosophy* 507, 518-22. Dolinko claims that there are two propositions which are defining of retributivism. He refers to them as the intrinsic good claim and the desert claim which provide that the guilty receiving their just deserts is an intrinsic good, and that punishment is justified solely by the fact that those upon whom it is inflicted deserve it, respectively. The first claim here really embodies two propositions. One, that punishing the guilty is an intrinsic good; the other, that the level of punishment should be in accordance with one's just deserts. It should be noted that the other supposed central plank of retributivism, the desert claim (as stated by Dolinko), adds nothing to the intrinsic good claim. In fact, rather than being a defining aspect of retributivism it appears to be at most a logical consequence of the intrinsic good claim: ie, the guilty deserve to suffer, *because* the guilty receiving their just deserts is an intrinsic good. In any event, Dolinko argues that while these claims are characteristic features of retributivism, they are ultimately incompatible because the intrinsic good claim suggests that bringing about

of the leading contemporary retributive theories reveals that few do not ultimately advert to extraneous reasons to justify punishment. Only one retributive theory (intrinsic retributivism)[34] claims that punishment is justified because it is intrinsically good to punish wrongdoers. All other retributive theories appear to rely, at least partially, on the instrumental effects of punishment to justify the practice. These include the capacity for punishment to convey blame or reprobation;[35] to induce repentance, self-reform, reparation, and reconciliation;[36] or to restore the fair balance of benefits and burdens which is disturbed by crime.[37]

Accordingly, it is difficult to identify a principle which represents a retributive pedigree. The true picture seems to be that there are many different theories of punishment wearing the retributive label. There is no

---

good is what justifies punishment, thereby invoking consequential considerations. while the desert claim provides that extraneous consequences cannot be relevant to the justification of punishment (at 516). However, this is only so if one adopts a consequential interpretation of the intrinsic good claim and (as Dolinko concedes) this is a claim which most retributivists, at least ostensibly, reject.

34 See T. Honderich, above n 21, 212; J. Kleing, *Punishment and Desert* (The Hague, 1973), 67; D.J.B. Hawkin, "Punishment and Moral Responsibility" in S.E. Grupp (ed.), *Theories of Punishment* (Indiana University Press, Ontario, 1971) 13, where he asserts that at the pre-reflective level it seems to be assumed that a guilty act deserves punishment.

35 A. von Hirsch, above n 2.

36 R.A. Duff, *Trials and Punishment* (Cambridge University Press, Cambridge, 1986).

37 W. Sadurski, *Giving Desert its Due* (Kluwer Academic Publishers, 1985) ch 8; J. Finnis, *Natural Law and Natural Rights* (Clarendon Press, Oxford, 1980) 263-4. See also, H. Morris, 'Persons and Punishment' in J.G. Murphy (ed.), *Punishment and Rehabilitation* (Wadsworth Publishing Co, Belmont, 1973); J.G. Murphy, *Retribution, Justice and Therapy* (Dordrecht, 1979) 82-115. Morris and Murphy have both subsequently moved away from this theory: see H. Morris, *A Paternalistic Theory of Punishment* (1981); J.G. Murphy, "Retributivism, Moral Education and the Liberal State" (1985) 4 *Criminal Justice Ethics* 3. A. von Hirsch also previously endorsed this theory as a partial justification for punishment, but has resiled from it because it is vulnerable to unjust society objections and fails to provide guidance on how much punishment is deserved: A. von Hirsch, *Past or Future Crimes* (1985) ch 5; "Censure and Proportionality" in R.A. Duff and D. Garland (eds), *A Reader on Punishment* (Oxford University Press, Oxford, 1994), 115, 116; "The Politics of 'Just Deserts'" (1990) *Canadian Journal of Criminology* 397, 408). For a modern defence of this version of retributivism, see G Sher, *Approximate Justice: Studies in Non-Ideal Theory* (Rowman & Littlefield, London, 1997). For a response to an earlier outline of Sher's theory, see C.L. Ten, "Positive Retributivism" (1990) 7(2) *Social Philosophy & Policy* 194.

distinctive badge worn by, or internal unifying principle running through, all of them. But they do have at least one thing in common: they are not utilitarian.[38] Thus retributive justifications for punishment do not turn on the likely achievement of consequentialist goals: punishment is justified even when 'we are practically certain that attempts [to attain conseqentialist goals, such as deterrence and rehabilitation] will fail.'[39] This alludes to another characteristic feature of retributive theories: they are essentially backward looking; punishment is an appropriate response to a past offence, irrespective of other incidental effects of it.[40] This is in contrast to utilitarianism which is concerned only with the likely future consequences of imposing punishment. The contrast with utilitarianism also explains why theories which rely on factors or virtues outside the parameters of the act of punishment itself are still regarded as retributive. The virtues invoked are not, at least expressly, consequentialist in nature, but instead are those commonly associated with a deontological account of morality.[41]

It is unclear whether this captures the full contrast between

---

38 However, it has been claimed that retributivism could be formulated as a consequentialist theory: M. Moore, "Justifying Retributivism" (1993) 27 *Israel Law Review* 15. For a convincing criticism of this see D. Dolinko, "Retributivism, Consequentialism, and The Intrinsic Goodness of Punishment" (1997) 16 *Law and Philosophy* 507; 509, who argues that a non deontological account of retributivism evinces a poor understanding of retributivism and that consequentialist retributivism is not a coherent form of retributivism. Dolinko must be accurate here. One of the key criticisms of many retributive theories is that logically they invoke consequential considerations, hence they do not provide a viable alternative to utilitarianism, for example, see C.L. Ten, above n 21, 38-65. If retributivism expressly invokes consequential considerations then arguably it has nothing left to present as a viable alternative to utilitarianism.

39 R.A. Duff, Trials and Punishments, above 36, 7. However, it has been argued that many retributive theories do implicitly rely on consequential considerations; see C.L. Ten, above n 21, 38-65.

40 Also it is claimed that future orientated considerations—the defendant's need for treatment, his or her likelihood of offending again, the deterrent effect of his punishment on others—have no role in determining the comparative severity of penalties: A. von Hirsch, Past or Future Crimes, above n 2, 10.

41 The retributive traits I advert to here are similar to those in the definition of retributivism adopted by D. Dolinko, who defines retributivism as any theory that 'explains either the rational justification of punishment, or its moral justification, or both, by appealing to the notion that criminals deserve punishment rather than to the consequentialist claim that punishing offenders yields better results than not punishing them': D. Dolinko "Some Thoughts About Retributivism" (1991) 101 *Ethics* 537. The retributive theory which most expressly endorses a deontological account of morality is rights retributivism, this is discussed below.

retributivism and utilitarianism. However, for the purpose of this paper the *precise* definition of retributivism is not critical. For it is argued that all theories of punishment which do not advert to consequential considerations are unsound.

The Utilitarian Theory of Punishment

The picture is far clearer in relation to the utilitarian theory of punishment. Utilitarianism is the theory that the morally right action is that which produces the greatest amount of utility. The utilitarian theory of punishment is merely an application of the general utilitarian theory of morality to the specific issue of punishment, and in this domain it does not matter significantly which version of utilitarianism is adopted.[42] Although utility has been defined in numerous ways, I shall adopt what I consider to be the most persuasive and coherent version of utilitarianism: hedonistic act utilitarianism, which provides that the utility which should be maximised is happiness or pleasure, which is the sole intrinsic good, and that pain is the sole inherent evil.[43]

---

42 C.L. Ten, above n 21, 4.

43 Apart from hedonistic utilitarianism, several other utilitarian theories have been advanced. Ideal utilitarianism is the theory that in addition to happiness there are other intrinsic goods such as knowledge, love and beauty (see G.E. Moore, in *Principa Ethica* (Cambridge University Press, Cambridge, 1903) and accordingly we should also attempt to maximise these virtues. Ideal utilitarianism is unstable and ultimately collapses into hedonistic utilitarianism. It is true that we generally pursue virtues such as love, beauty, knowledge, but we do not do so for their own sake. Rather we seek them because they generally tend to generate pleasure. To the extent that we desire other things such as money, power, virtue or fame it is only because they are generally a means to happiness, but this does not change the derivative attraction of such virtues: see J.S. Mill, "Utilitarianism" in M. Warnock (ed.), *Utilitarianism* (Fontana Press, Glasgow, 1986, first published 1861), 251, and D. Raphael, *Moral Philosophy* (Oxford University Press, Oxford, 1981), ch 4. The most recent substitution of note, is to define utility in terms of preference or desire satisfaction. The corresponding theory is called preference utilitarianism. Preference utilitarianism does not have the same degree of self-evident appeal as hedonistic utilitarianism. For example, it is unclear why we should seek to maximise desires which make people unhappy. Further, it is impossible to know which act will maximise desire satisfaction, given the overwhelming number of desires which will invariably need to be considered in any particular case. Also it may be argued that our ultimate fundamental desire is generally, if not always, to be happy and hence that preference utilitarianism, too, collapses into hedonistic utilitarianism. Preference utilitarianism is outlined in R.M. Hare in *Moral Thinking: Its Levels, Methods and Point* (Clarendon Press, Oxford 1981) and P. Singer, *Practical*

The utilitarian starting point regarding punishment is to consider the most direct and immediate effect of punishment, and from this perspective it is a bad thing because it causes unhappiness to the offender. It is only justified because of the wider contingent benefits it produces, which it is felt out weigh the bad consequences. The good consequences of punishment which are thought to outbalance the suffering inflicted on the offender, include discouraging the offender from re-offending and potential offenders from committing crimes in the first place, and once the offender is apprehended by rehabilitating him or her and where necessary incapacitating the offender.[44] If there are several forms of punishment which produce the same good consequences we must chose the one which imposes the least unpleasantness to the offender. Thus unlike retributivism, the utilitarian theory of punishment is forward looking: the commission of a criminal act does not justify punishment; rather punishment is only warranted if some good can flow from it.

I shall now turn to the theoretical attacks on utilitarianism, which have been so effective in diminishing its appeal as a justification for punishment.

Ethics (2nd ed, Cambridge University Press, Cambridge, 1993). Singer, at p14, states that if happiness is defined broadly enough to include achieving what one desires (as I believe is the case), then there is no conflict between hedonistic and preference utilitarianism. A further distinction is made between act utilitarianism and rule utilitarianism. Act utilitarianism is simply the view that the correctness of an action is judged according to the degree of utility it promotes. Rule utilitarianism is the view that the rightness of an act is assessed by reference to its compliance with rules established to maximise utility. For the rule utilitarian the principle of utility is used as a guide for the rules we should follow, as distinct to the particular actions we should perform. Due to the difficulty in performing the utilitarian calculus necessary to determine which of a number of options we should choose it is claimed that a set of rules guiding us in our decisions would be more likely to achieve the desired goal. The main problem with rule utilitarianism is that it is inevitable that in complying with the rules there will be occasions when happiness will not be maximised. To refuse to break the rule in such circumstances constitutes 'rule-worship': see J.C.C. Smart, "An Outline of a System of Utilitarian Ethics" in J.C.C. Smart and B. Williams (eds), *Utilitarianism: For and Against* (Cambridge University Press, London 1973) 3, 10. It is no answer that in most cases it is beneficial to comply with the rule, otherwise we are putting the rule above its justification. If we do break the rule, we are still being guided by the ultimate principle: act utilitarianism. As I discuss later it is not that the act utilitarian does not see general rules as playing an important role in our moral decisions, but he or she will only act in accordance with the rules where it is felt that on each particular occasion this will generate most happiness.

44 Other benefits of punishment are outlined below.

Punishing the Innocent and Utilitarianism

Outline of the Objection

A famous illustration of the objection concerning punishing the innocent is McCloskey's famous small town sheriff example:[45]

> Suppose a sheriff were faced with the choice of either framing a negro for a rape which had aroused white hostility to negroes (this negro believed to be guilty) and thus preventing serious anti-negro riots which would probably lead to loss of life, or of allowing the riots to occur. If he were...[a] utilitarian he would be committed to framing the negro.[46]

Utilitarian Responses To Punishing the Innocent

(i) Punishing the Innocent Only a Theoretical Problem

There have been several attempts to counter this objection. First it has been suggested that examples which supposedly commit a utilitarian to punishing the innocent are impossible in the real world and hence need not be addressed.[47] Punishing the innocent may at times provide short term benefits, such as securing social stability, but these are always more than offset by the likelihood of greater long term harm due to the loss of confidence in the legal system and the associated loss of security to all members of the community who will fear that they may be the next person framed, once the inevitable occurs and it is disclosed that an innocent person has been punished. But with only a little imagination the above example can be tightened up, for instance by introducing considerations that significantly reduce or totally obviate the possibility of disclosure, so that the only logical utilitarian conclusion is to punish the innocent.[48] Even if the

---

45 Duff gives two other types of examples of what he terms punishing the innocent: punishing people for strict liability offences and punishing someone more severely than is commensurate with the seriousness of the offence: R.A. Duff, above n 36, 154-5. These situations will not be addressed in this paper. It is questionable whether such people are innocent, and in any event these situations do not represent the paradigm case of punishing the innocent.

46 H.J. McCloskey, *Meta-Ethics and Normative Ethics* (Martinus, The Hague, 1969), 180-1. A similar example to McCloskey's is provided in E.F. Carritt, *Ethical and Political Thinking* (Greenwood Publishing, Oxford, 1973), 65.

47 For example, see T.L.S. Sprigge, "A Utilitarian Reply to Dr McCloskey" (1965) 8 *Inquiry* 272.

48 As an example, McCloskey's hypothetical could be altered by providing that the town was an isolated one, hence there is no opportunity for help arriving before the riots occurred. Also the crime should be murder, not a rape, in

process of modifying the examples appears to far remove them from the real world, it is still a situation which the utilitarian must deal with. As Ten notes, 'fantastic examples', as he labels them, which raise for consideration fundamental issues, such as whether it is proper to punish the innocent, play an important role in the evaluation of moral theories since they sharpen the contrasts between them and illuminate the logical conclusions of the respective theories and in this way test the true strength of our commitment to the theories.[49] Thus, fantastic examples cannot be dismissed summarily on the basis that they are 'simply' hypothetical.

(ii) Definitional Arguments

An argument which has been used to buttress the punishing the innocent attack on utilitarianism is that such an outcome is inconsistent with the definition of punishment: 'punishment must not only be *of* an offender; it must also be for *her* offence';[50] 'even if the world gathered all its strength, there is one thing it is not able to do, it can no more punish an innocent [person] than it can put a dead person to death'.[51]

However this approach cuts both ways. If by definition punishment can only be imposed on the guilty, it follows that the issue of punishing the

which case there is one less person who could reveal the miscarriage of justice that has occurred, and thus the risk of a possible loss of respect and confidence in the law is not as significant. See also C.L. Ten, above n 21, 18; J. Rawls, "Two Concepts of Rules" (1955) 64 *The Philosophical Review* 3; R.M. Hare, *Moral Thinking: Its Levels Methods and Point* above n 43; 162-4. R.B. Brandt has argued that the rule utilitarian is not necessarily committed to punishing the innocent: R.B. Brandt, "Ethical Theory", *The Problems of Normative and Critical Ethics* (Prentice Hall, New Jersey, 1959), 490-5. For a counter see McCloskey, "A Non-Utilitarian Approach to Punishment" (1965) 8 *Inquiry* 239.

49 C.L. Ten, *Ibid*, 18-25, draws a distinction between a fundamental moral principle (a principle which is not justified by reference to some further moral principle) and a secondary moral principle (which has to be justified by appeal to some further moral principle), and makes the point that fantastic examples play an important role in relation to the evaluation of fundamental moral principles. While for the utilitarian the wrongness of punishing the innocent is a secondary principle, deriving its justification from the sole utilitarian fundamental principle that it is wrong because it would cause net unhappiness, for others the proscription against punishing the innocent is itself a fundamental moral principle. Accordingly fantastic examples have a role in testing this principle, and the fundamental utilitarian principle. Fantastic examples also allow us to ascertain whether or not a principle is fundamental or not.

50 R.A. Duff, above n 36, 152.

51 S. Kierkegaard, *Purity of the Heart is to Will One Thing* (Collins, Fontana, 1961) 85, as cited in RA Duff, *Ibid*, 152.

innocent is one which the utilitarian need not even begin to tackle. The definition of punishment which is adopted applies independently to the justificatory theory of it. Accordingly, it is open for the utilitarian to adopt the above definition and rest his or her case on the basis that one cannot be committed to that which is logically impossible.[52]

Ultimately, however, as Armstrong points out, definitional disputes are not likely to resolve normative issues. Irrespective of how punishment is defined, the utilitarian cannot side-step the problem of punishing the innocent, since the objection looses none of its force if the question is framed in terms of 'why shouldn't we do to the innocent that which, when it's done to the guilty is known as punishment'.[53] This requires a substantive, not a formal response.

But before turning to this, in a bid to highlight the futility of definitional arguments in this area, it is illuminating to note the caginess with which they may be, and have been, used. The fact that it makes sense to assert that 'he was punished for something he did not do'[54] has been used to take a cheap shot at retributivism, since it supposedly shows that punishment of the innocent is possible and thereby retributivism, which links punishment to a past crime, must be wrong.

(iii) Substantive Response—Hard Cases Lead to Hard Decisions

The more promising utilitarian response is not to attempt to deflect or avoid the conclusion that there may be some extreme situations where utilitarianism commits us to punishing the innocent, but rather to accept this outcome and contend that as horrible as this may seem on a pre-reflective level, on closer consideration it is not a matter that *really* insurmountably troubles our sensibilities to the extent that it entails that any theory which approves of such an outcome must necessarily be flawed.[55] By drawing comparisons with other situations in which we take the utilitarian option it is contended that punishing the innocent is not a practice which is necessarily unacceptable.

The view that punishing the innocent is the morally correct action in some circumstances is consistent with and accords with the decisions we as individuals and societies as a whole readily have made and continue to

52 See, A. Quinton, "Punishment" (1953-4) 14 *Analysis* 133.

53 K.G. Armstrong, "The Retributivist Hits Back" in S.E. Grupp (ed.), *Theories of Punishment* (Indiana University Press, Ontario, 1971) 19, 34. See also R.B. Brandt, "Ethical Theory", *The Problems of Normative and Critical Ethics* (Prentice Hall, New Jersey, 1959), 494-5.

54 K.G. Armstrong, *Ibid*, 19, 20—Armstrong, sensibly, rejects this criticism.

55 The distinction I am making between intuitive moral judgements and those formed after due reflection is similar to that made by R.M. Hare between intuitive and critical levels of moral thinking; see R.M. Hare, above n 43.

make when faced with extreme and desperate circumstances. Once we come to grips with the fact that our decisions in extreme situations will be compartmentalised to desperate predicaments and will not have a snowball effect and serve to henceforth diminish the high regard we normally have for important individual concerns and interests we find that when placed between a rock and a hard place we do and *should*, though perhaps somewhat begrudgingly, take the utilitarian option. In the face of extreme situations we are quite ready to accept that one should, or even must, sacrifice oneself or *others* for the good of the whole.

For example, in times of war we not only request our strongest and healthiest to fight to the death for the good of the community, but we often demand that they do so, under threat of imprisonment or even death. Quite often they must battle against hopeless odds, in circumstances where we are aware that in all probability they are not coming back.[56] And what is more: they *must* do so. Give their life. Not because they want to; not because they are bad; but, merely because it would be good for the rest of us—classical utilitarian reasoning.[57] Faced with the reality of the decisions we *do* make in such horrible situations the examples proffered against utilitarianism about the terrible things it entails, such as punishing the innocent, lose their bite. Horrible situations make for appalling decisions whichever way we turn, but in the death knell we do make the utilitarian choice because of our lack of true commitment to any higher moral virtue. By opting for the utilitarian line we are soothed by the one saving grace: at least the level of harm has been minimised. When the good of many or the whole is at significant threat we have no difficulty selecting certain classes of innocent individuals, whose only 'flaw' is their sex, state of health, and date of birth to go in to bat for the rest of us. Their protests that they should not be compelled to go because it impinges on their civil, legal or human rights to such matters as life and liberty, or their desperate appeals to other virtues such as justice or integrity, fall on obstinate ears. For this is serious stuff now—our lives (or other important interests) are at stake. Such appeals should be saved for rosier times. And when advanced in theory, we can all 'agree'.

The decisions we do actually make in a real life crisis are the best

---

56 Whilst this is not normally the case, ie. we normally like to think that we send our soldiers into situations with at least a fighting chance, there are countless reported instances of men being ordered to go or remain in situations which can only be described as suicide missions. For those brave men who voluntarily place themselves in such situations, it is rather illuminating that the proscription against suicide disappears. They are heroes rather than bad men—they followed the dictates of utilitarianism.

57 The classic deontological response to this and other examples, the doctrine of double effect, is discussed below.

evidence of the way we actually do prioritise important, competing, principles and interests. Matters such as rights and justice are important but in the end are subservient to, and make way for, the ultimate matter of significance: general happiness. Bad as it seems, framing the negro and imprisoning the innocent, are certainly no more horrendous than the decisions history has shown we have made in circumstances of monumental crisis.

A pointed example is the decision by the English Prime Minister of the day, Winston Churchill, to sacrifice the lives of the residents of Coventry in order to not alert the Germans that the English had deciphered German radio messages. On 14 November, 1940 the English decoded plans that the Germans were about to air bomb Coventry. If Coventry was evacuated or its inhabitants advised to take special precautions against the raid the Germans would know that their code had been cracked, and the English would be unable to obtain future information about the intentions of its enemy. Churchill elected not to warn the citizens of Coventry, and many hundreds were killed in the raid which followed. The lives were sacrificed in order not to reveal the secret that would hopefully save many more lives in the future.[58] Significantly, such decisions have subsequently been immune from widespread or persuasive criticism. This shows not only that when pressed we *do* take the utilitarian option, but also that it is felt that this is the option we *should* take.

Now, what we do actually do, does not justify what ought to be done. Morality is normative, not descriptive in nature: an 'ought' cannot be derived from an 'is'.[59] Still, the above account is telling because the force of

---

58 See, M. Velasquez and C. Rostankowski, *Ethics: Theory and Practice* (Prentice Hall, New Jersey, 1985) 103-6. A famous modern day example which comes closest to the dilemma of choosing whether to frame the innocent or tolerate massive abuses of rights followed the Rodney King beating in Los Angeles on 3 March 1991. The policeman who beat King was acquitted under state law of any offence regarding the incident. Riots ensued resulting in widespread looting, damage to property, and dozens of deaths. Shortly afterwards the government announced the almost unprecedented step that the policeman, who one must remember was found innocent of the alleged crime, was to be tried on federal charges regarding the incident. He was duly found guilty, despite the apparent double jeopardy involved. Whatever one's view of the government's motivation for committing the policeman on federal charges, it seems that justice took a back seat; for a while.

59 This has been used as an argument against a naturalistic view of morality. However, see C.R. Pigden, "Naturalism" in P. Singer (ed.), *A Companion to Ethics* (Basil Blackwell, Oxford, 1991) 421, 422-6, where he points that this phenomenon simply reflects the conservative nature of logic—you cannot get out of it, what you do not put in.

the punishing the innocent objection lies in the fact that it supposedly so troubles our moral consciousness that utilitarianism can thereby be dismissed on the basis that the outcome is so horrible that 'there must be a mistake somewhere'. But this loses its force when it is shown that punishing the innocent is in fact is no worse than other activities we condone.

Punishing the Innocent and Retributivism

Outline of the Objection

It has been pointed out that it is not only a utilitarian system of punishment that may permit punishment of the innocent.

> No practicable system of punishment can hope to punish only the guilty; in ensuring that we punish a reasonable proportion of the guilty, we will inevitably punish some who are in fact innocent. We could avoid punishing the innocent, by refusing to punish anyone: in maintaining a system which we know will sometimes punish the innocent we therefore cannot claim that we punish the innocent unintentionally; we must admit that we too are abusing it as a system of punishment.[60]

Retributive Responses to Punishing the Innocent

(i) Increase Level of Safeguards Against Wrongful Convictions

It is inevitable, given the fallibility of any institution, that any criminal justice system will at times inflict punishment on the innocent.[61] This difficulty could be largely[62] circumvented by increasing the amount and level of safeguards in the criminal justice process. For example, the standard of proof could be raised from beyond reasonable doubt to, say, beyond any possible doubt and admissible evidence could be limited to direct observations of the relevant act and a confession[63] could be made a

---

60 G. Schedler, "Can Retributivists Support Legal Punishment?" (1980) 63 *The Monist* 185. For this reason he concluded that retributivists simply cannot support the institution of punishment.

61 This point is also made by D. Dolinko, "Retributivism, Consequentialism, And The Intrinsic Goodness of Punishment" (1997) 16 *Law and Philosophy* 507, 510; see also, D. Husak, "Why Punish The Deserving?" (1992) 26 *Nous* 450-1.

62 Though perhaps no amount of procedural safeguards could ever fully prevent the conviction of at least some innocent people.

63 In circumstances where there was no possibility that it was procured by threat or inducement or was otherwise involuntary.

mandatory pre-condition to a finding of guilt. However, such a response is not open to retributivists. It would be self-defeating since it would result in more innocent people being harmed than is presently the case as a result of our imperfect criminal justice system.[64] Any retributive theory must have at its foundation some theory of morality, given that the prohibition against punishing the innocent is not a freestanding principle. The broader principle which logically flows from this prohibition is that people who are not blameworthy in any way should not be harmed.[65] The effect of radically increasing the amount of legal safeguards in criminal cases would result in very few guilty people being punished and thereby an increase in the amount of crime and innocent people being harmed.

(ii) The Doctrine of Double Effect

A common retributive response to the problem of punishing the innocent is that offered by Duff, who denies that punishing the innocent is a concern for the retributivist, since, unlike the utilitarian situation, punishment of the innocent is *not intended* and occurs despite the aims of a retributive system of punishment.[66] The credibility of this response turns on the persuasiveness of the distinction between consequences which are intended and those that are merely foreseen.

Outline of the Doctrine of Double Effect

Underpinning Duff's argument is the doctrine of double effect, which provides that it is morally permissible to perform an act having two effects, one good and one evil, where the good consequence is intended and the bad merely foreseen and there is proportion between the good and bad consequences which occur pretty much simultaneously.[67]

---

64 There may, however, be community support for increasing the level of safeguards to protect against wrongful conviction. A recent American study revealed that the vast majority of respondents supported a higher standard of certainty of guilt in cases involving the death penalty: D. Weinstock and G.E. Schwartz, "Executing the Innocent" (1998) *Criminal Law Bulletin* 330. However, given the finality and extreme nature of this form of punishment, it may well be that similar sentiments will not apply in relation to other forms of punishment.

65 An attempt to attenuate this principle, by confining harm to legally imposed sanctions would appear indefensible.

66 The same point is made by M. Moore, "Justifying Retributivism" (1997) 27 *Israel Law Review* 15, 20.

67 It is also sometimes contended that a further condition is that the act must not be intrinsically bad: see H.T. Engelhardt and J. Kenny, "Principle of Double Effect", in B. Brody and H.T. Engelhardt (eds), *Bioethics* (Prentice Hall, New Jersey, 1987), 160. However, given that the doctrine is commonly applied to very grave cases involving things such as the killing of innocent people and it is on the basis of the doctrine itself that such acts are sought to

The doctrine has a rich history and is frequently appealed to as a purported justification for acts or practices which produce foreseen undesirable consequences. For example, it is the reason why it is, supposedly, permissible to bomb an enemy's ammunition factory in wartime, even though it will result in the certain death of civilians, and why it is justifiable to kill an unborn baby where this is necessary to save the mother, and why self-defence is legitimate.[68] In the case of euthanasia, it is employed as a justification for alleviating pain by increasing doses of pain killers even when it is known that this will result in death—the intention is to reduce pain, not to kill.[69]

The legal status of the doctrine is unclear. In *R v Nedrick*,[70] the House of Lords held that foresight, even of near certainty, was not the same as intention, whereas in *Hyam v R*,[71] Lord Hailsham was of the view that one who blows up an aircraft in order to obtain money intends to kill.[72] However in relation to euthanasia the courts have endorsed the doctrine. In *R v Adams,* it was held that 'it is permissible to relieve suffering even if the measure...incidentally shortens life'.[73] This has, at least implicitly, been endorsed in subsequent cases.[74]

---

be justified, it begs the question to make such a condition an internal part of the doctrine. See also T. Nagel, *The View From Nowhere* (Oxford University Press, New York, 1986) 179, whose formulation of the doctrine of double effect essentially accords with the above.

68 Although there are also other justifications for excuse of self-defence.

69 For example, see P. Mullen, "Euthanasia: An impoverished Construction of Life and Death" (1995) 3 *Journal of Law and Medicine* 121, 127.

70 [1986] 1 WLR 1025.

71 [1975] AC 55.

72 To explain this incongruity, it has been suggested that where the motive is honourable there is room to distinguish between foresight and intention: D. Lanham, "Euthanasia, Pain Killing, Murder and Manslaughter" in J. McKie, *Active Voluntary Euthanasia: The Current Issues* (Centre for Bioethics, Monash University, Melbourne, 1994) 67, 73. However, this cannot be used to give a general account of the difference between that which is intended and foreseen, since this distinction is itself meant to be a test by which the moral status of an act can be evaluated. The doctrine would be redundant if the moral status of the act was clear from the outset.

73 [1957] *Crim LR* 365.

74 See *R v Cox* (1992) 12 BMLR 38, 39; *Airedale NHS Trust v Bland* [1993] 789, 867; *Auckland Area Health Board v Attorney-General* [1993] 1 NZLR 235, 248; *Re J (Wardship: Medical Treatment)* [1991] Fam 33, 46. The doctrine is the cornerstone of the *Consent to Medical Treatment and Palliative Care Act* 1995 (S A), which provides that where there is a conflict between proper control of symptoms and accelerating the timing of an inevitable death, then symptom control prevails. If a side effect of palliative care is death this is deemed not to constitute death in law, on the basis of the

Is There a Distinction Between Intended Consequences and Foreseen Consequences?

However the moral significance of the doctrine is much in dispute. Glover gives the example of a terrorist who for the purpose of making a [legitimate] political protest throws a bomb into a crowd which kills several people.[75] He correctly points to the difficulty in ascertaining whether the deaths are intentional or merely foreseen. The above examples illustrate that inevitability of the deaths cannot be used to impute intention, for the doctrine provides that foreseen consequences which are certain need not be counted as intentional. Thus the fact that the terrorist is possibly more certain to kill innocent people, than the institution of punishment is to punish innocent people is irrelevant.

It is also beside the point that the institution of punishment does not aim to punish specific innocent individuals. For not only would the terrorist be pleased if no person was killed, but as far as he or she is concerned the crowd consists of random unidentified individuals.[76] Thus there appears no principled reason to maintain that the terrorist intends to kill, whereas the institution of punishment does not intend to punish the innocent: in both instances if the respective objectives could be achieved without the harmful by-products the agents would be pleased. This alludes to the central flaw in the doctrine of double effect, which is that it is not possible to provide a general account of the distinction between what is intended and what is merely foreseen which applies in all circumstances.[77] It is illusory to claim that intentions are divisible, along the lines of good and bad consequences of an act.

The preferable view is that there is no inherent distinction between consequences that are intended and those which are foreseen. We are responsible for all the consequences which we foresee, but nevertheless elect to bring about. Whether or not we also 'intend' them is irrelevant.

---

doctrine.

75 J. Glover, in *Causing Deaths and Saving Lives* (Pelican Books, London, 1977), 88. It could be argued that doctrine does not apply in this situation because of a lack of proportionality between the good and bad effects of the act. However, this could be answered by altering the example so that only one person was killed in the explosion, and the protest was against a brutal regime which had a history of bowing to such acts of aggression.

76 See also D. Dolinko, "Retributivism, Consequentialism, And The Intrinsic Goodness of Punishment" (1997) 16 *Law and Philosophy* 507, 510-513, who considers Philippa Foot's example of the wicked merchants selling poisonous oil and thereby killing innocent people: P. Foot, "The Problem of Abortion and the Doctrine of Double Effect" in *Virtues and Vices* (Basil Blackwell, Oxford, 1978), 22.

77 See also, D. Dolinko, Id.

Underlying, and the only coherent basis for distinction adverted to by, the doctrine of double effect is nothing more than the consequentialist view that it is permissible to do that which is 'merely foreseen' if the adverse consequences of the act are outweighed by the good consequences that are 'intended'.[78] Utilitarianism deals with the difficulties that are sought to be overcome by the doctrine in a far more comprehensible and straight forward manner. The reason that the doctor who administers a lethal dose of pain killers and the legal system which punishes the innocent (believed guilty) are not blameworthy has nothing to do with the fuzziness relating to what is intended as opposed to foreseen, but simply follows because in all the circumstances the good consequences outweigh the bad. And, from the perspective of the innocent person who is punished, it certainly does not whether his or her punishment was intentional or merely foreseen: it hurts just the same. Notwithstanding this, an institution which causes such hurt is still morally justifiable, because it leads to a happier situation overall, than the alternative—abolishing punishment.

In the end, the motivation for the doctrine of double effect seems to be to provide a means for deontological theories which employ notions of absolute (or near absolute) rights to deal with the difficult, but inevitable, situations where there are conflicts between different rules or rights, or even different applications of the same rule or right.[79] The doctrine maintains absolutism by utilising the fiction of merely foreseen consequences and absolving liability for them.

(iii) The Innocent Not Used as a Means

However, there may yet be another way in which the retributivist may attempt to defend a system of punishment which, unfortunately, but invariably will result in the punishment of some innocent people. This adverts less crudely to the distinction discussed earlier regarding the identity of victims who are incidentally harmed as a by-product of what is thought to be a generally desirable act, and invokes the Kantian concept of means and ends.

The nature of this distinction is illustrated by the following example. It is necessary to build a bridge between two suburbs. Two different types of bridges are possible. If proposed bridge A is built, actuarial studies show

---

78 J. Rachels, in *The End of Life: Euthanasia and Morality* (Oxford University Press, Oxford, 1986), 94-6, argues that a person's intention is not relevant to determining whether an act is right or wrong, but instead is relevant to assessing the character of the person who does it. However, the difficulty with this is to coherently distinguish between the evaluation of the act and the agent; we normally judge people by their actions.

79 J.L. Mackie, *Ethics: Inventing Right and Wrong* (Pelican Books, Oxford, 1977), 161.

that it is certain that two people will die during the construction. If bridge B is constructed, it is known in advance that a particular workman will die. It is contended that the utilitarian on this information alone would elect route B.[80] On the other hand, a powerful deontological argument can be made in favour of bridge A, because unlike in case B no individual is being used simply as a means for a particular end and this follows from the fact that each person who is involved in the project or is in some way affected by it *may too ultimately benefit* from the project. For example, he or she may use the bridge or be paid a salary for working on its construction.[81] It may be argued that the terrorist example is analogous to situation B and the retributive system of punishment to situation A. Although the terrorist kills victims who are unknown to him or her, they are nevertheless specific people whose identity is ascertainable at the time of the act and they have no prospect of benefiting from the legitimate protest. Not so in the case of a retributive system of punishment which unintentionally punishes an innocent person. The identity of the 'offender' is not known at the time of conviction and sentence, and may never be known. Even though the innocent who are punished ultimately do suffer, they are part of a general practice through which they too may have prospered. In this sense, so the argument runs, they are not sacrificed for the good of the whole.

However, even putting to one side the difficulties associated with the means and ends distinction, this retributive approach to the dilemma is also unsatisfactory. At the time an innocent person is punished there is always at least one person who is aware of the injustice: the 'offender'. It is not to the point that the system is oblivious to the innocence of the 'offender' at the point of conviction and sentence. If the system was *really* concerned with the unfairness it would have taken measures to avoid the predicament eventuating; by implementing safeguards, of the type mentioned earlier, to prevent wrongful conviction. By persisting with such a defence of their theory, retributivists are expressing either feigned concern or blissful ignorance. Even more generally, it is immaterial that the 'offender' could

---

80 This conclusion is however by no means certain. A utilitarian could argue that ear marked deaths are worse than statistical ones because of the de-sensitisation that would follow if defined individuals were allowed to die. As an empirical fact, we seem to be built in such a way that when an identifiable individual is experiencing pain and suffering (or is in need of help) this impacts on us far more heavily than when it is experienced by faceless strangers. Thus in 1995 the Australian Government spent $5.8 million rescuing French sailor Isabelle Autissier who was stranded while on a solo frolic around the world, when the same money could have saved thousands of starving people around the world: Thus unless the number of statistical deaths is significantly more than earmarked ones, a utilitarian may prefer to opt for bridge A.

81 I thank Professor Ten for this point and example.

have potentially benefited from the institution of punishment. He or she did not, and is a victim of it and it is unrealistic to expect meaningful solace to be attained through such unrealised *potential*: in any meaningful sense of the word, he or she is being sacrificed for the good of the whole.

The Moral Relevance of Intentions

A related problem for the retributivist, especially in respect to the doctrine of double effect, is the absolute faith and reliance placed on the concept of intentions. Non-consequentialist moral theories of morality, invariably assert that intentions have intrinsic moral relevance: the intention to help others is worthy of moral praise, while the intention to harm justifies moral condemnation. On its face this may seem incontrovertible. However, the picture becomes less clear if one considers the case of 'Jack'.

Jack is generally a good person; more often than not he intends to assist others that he believes are not as fortunate as him. But he is not very bright. His parents (who unknown to Jack) are very wealthy and have always been extremely paranoid and untrusting of others; believing that others wish to exploit their wealth. Accordingly, they have been extremely vigilant to ensure that Jack is sheltered from the outside world, to the extent that Jack, despite being an adult, has never attended school (or received any other form of meaningful education) and, accordingly, has a very poor understanding of the empirical cause and effects systems which operate in the world. So poor, that he never manages to succeed in implementing his intentions so far as they affect his relationships with others, and in fact he always produces the morally opposite result. Thus when he wants to harm people, instead of robbing them, he gives them money (because he believes money is a cause of unhappiness) and when he wants to help he punches them (believing this to be a form of affection). Given that Jack's beliefs are so entrenched that they are beyond revision, even the most ardent non-consequentialist would prefer the 'nasty' Jack and would agree that it would be far better to live in a world of 'nasty' rather than 'nice' Jacks.

The only reason that we generally view intentions as being inherently worthy of praise or blame is that most us have sufficient factual knowledge about the empirical processes in the world to set in train the appropriate causal processes to achieve our intentions, hence there is a *very close connection* between intentions and consequences. If it transpired that intentions generally had no connection with consequences they would promptly become morally irrelevant. The above account of Jack may seem far-fetched, but the point that the example seeks to drive home, is already entrenched in the context of other mental states we experience. We are not responsible or culpable for other mental states we experience which do not produce harmful consequences. We are not condemned for the aspirations

or intentions we experience while dreaming or for our private wishes which we do not act upon. To the extent that we may be criticised when our dark private wishes became public, this is merely because it is assumed that they reflect upon sinister personal traits which may in the future guide our conduct and lead to undesirable consequences. But absent the possible connection between our private wishes and ultimate consequences, they are not objects of praise or blame. Thus the only basis for ascribing moral relevance to intentions is because of their close link with consequences. When this link is severed, it becomes apparent that at the bottom the only thing which really matters is consequences, and the appeal of distinctions or doctrines which bank on the purported significance of intentions readily dissipates.

The point I wish to make here is not as revisionary as might first appear. It is not contended that intentions and other types of mental states, such as recklessness and negligence, are irrelevant and that accordingly we ought to abandon the heavy reliance generally placed on them, and thereby, for example, implement a strict liability system of law.[82] As an empirical fact, as I have stated, there is a close connection between our intentions and actions and therefore the person who intentionally brings about a harmful act is more blameworthy than one who does so due to, say, indifference or mistake. Even though the immediate and direct consequences are identical, the person who deliberately sets in train a causal process which results in harm to another deserves greater blame and punishment because such behaviour in general is likely to lead to more suffering long term and thus stern measures must be implemented to deter similar behaviour in the future.

The Law and Intentions

Mental states do have a role, however they are not the ultimate considerations which are relevant to moral responsibility. And despite the general significance attached to mental states by our legal system, whereby substantial emphasis is attached to precise mental states; such as recklessness, negligence and carelessness, ultimately the law recognises that mental states per se are irrelevant. No matter how pervasively wicked a person may be or how resolutely they may intend that a certain harmful state of affairs should eventuate no legal responsibility is ascribed until and unless such mental states are accompanied by actions. The only possible exception to this is the law relating to attempted criminal offences.

---

82 Especially of the type proposed by B. Wootton in *Crime and the Criminal Law* (2nd, Stevens, London 1981), where she contends that the function of the criminal law is to prevent socially harmful acts, and that therefore mens rea is not relevant to criminal liability; although it does have a role in sentencing. For criticism of her views, see C.L. Ten, above n 21, 115-22.

However, even here the degree of intrusion into the principle that intentions are per se irrelevant is only marginal, if at all. For liability to occur it is necessary for the offender, as well as possessing the requisite mental state, to perform *actions* which are very close to committing the substantive offence: the actions must be immediately, and not merely remotely, connected with the completed offence.[83]

(iv) Deliberately Punishing and Inadvertently Punishing the Innocent

The retributivist could yet contend that a system which deliberately punishes the innocent is nevertheless worse than one where this occurs inadvertently,[84] because it is surely likely to lead to *more* innocent people being punished, and that the former system is also less preferable because of the corrupting influence it will have on those involved in the practice.

However, the first objection overstates the likelihood of the utilitarian calculus actually coming down in favour of punishing the innocent. While the second alludes to some of the reasons why it is extremely rare that utilitarianism will condone such an outcome.

The extent to which we should be troubled by the conclusion that utilitarianism may in some circumstances allow punishment of the innocent is quelled following a proper appreciation of the utilitarian purposes of punishment. A true utilitarian picture reveals that pragmatically it is rare if ever that such an outcome is permissible. On the utilitarian calculus, punishment is only justified where it is outweighed by the benefits flowing from it. These benefits include not only deterrence, but also extend to such things as the rehabilitation of the criminal, satisfying the vengeful desires of victims, satisfying the community by apprehending criminals,[85] re-enforcing the wrongness of crime, and increasing community safety.

While punishing the innocent may in rare circumstances promote general security and general deterrence, it does nothing to advance the other

---

83 This is termed the proximity test; see *R v Mohan* [1976] QB 1; *R v Smith* [1975] AC 476. See also *Crimes Act* 1958 (Vic), s 321N.

84 The utilitarian could argue that such an objection is not one that the 'pure' retributivist can fairly invoke. This numbers based approach, after all, involves comparing one set of *consequences* with another. However, given that at least implicitly many retributive theories do appeal to considerations beyond the immediate act of punishment to justify the practice, I shall outline a substantive utilitarian response to this objection.

85 See also F. Rosen, "Utilitarianism and the Punishment of the Innocent" (1997) 9 (1) *Utilitas* 23, who argues that the reason that utilitarianism *supposedly* permits punishing the innocent stems from the flawed assumption that utilitarianism justifies punishment on the basis of deterrence alone, whereas deterrence is only one of several relevant factors in the utilitarian theory of punishment.

core utilitarian aims of punishment. The absence of these virtues on the punishment side of the scale, substantially lightens this end of the load.[86]

Additionally, in all probability, punishing the innocent would lead to serious community unrest and turmoil. A recent example, is the widespread civil unrest in Malaysia in October and November 1998 following the arrest, beating, and detention of the opposition leader, Anwar Ibrahim, for what were widely assumed to be fabricated criminal charges of sexual misconduct (homosexuality) and related corruption offences. While it is impossible to exhaustively articulate in advance the circumstances in which the utilitarian is committed to punishing the innocent, it is evident that one pre-condition to this is certainty, or near certainty, that the innocence of the 'offender' will never be disclosed. Realistically, given the large number of officials involved in bringing a person to 'justice' (due to the separation of administrative and judicial power in most jurisdictions) it is extremely rare that this requirement will be satisfied.

If, following a proper setting of the utilitarian scales, where *all* of the above variables where correctly weighed (including the corrupting affect that punishing the innocent would have on the officials involved), the scales came down in favour of punishing the innocent, then it is open for the utilitarian to assert that this does not reveal a shortcoming in the theory, since such an outcome is appropriate after all. To attack this response in a manner which does not beg the question, the retributivist must provide as a reason for the supposed wrongness of punishing the innocent, something beyond the mere assumption that it is abhorrent. This is best done by invoking the concept of rights.

The Intrinsic Wrongness of Punishing the Innocent and the Supposed Incompatibility of Utilitarianism and Rights

It has been asserted that utilitarian responses to the charge that utilitarianism permits punishing the innocent all fail because no matter how they are framed they miss the full force of what is wrong with punishing the innocent. The wrongness of punishing the innocent is not a question of weighing up the contingent consequences, but is evident from the act itself.[87]

There are two ways this argument can be developed. First, that it is necessarily always the case that punishing the innocent is wrong and that

86 *Ibid.*

87 C.L. Ten, Crime, above n 21. See also R.A. Duff, *Trials and Punishment*, above n 36, 160-1.

this is apparent from our intuitions. At the pre-philosophical level this argument is appealing, however as I have stated above is not decisive.[88] Intuition is a very poor guide: no doubt two hundred years ago it was generally considered appropriate to enslave negroes and fifty years ago it seemed fitting to forcibly remove aboriginal children from their parents.

To this end, the better argument is that punishing the innocent is wrong because it violates some fundamental virtue: namely, the right not to be punished without having committed an offence. Such a right is recognised in some form or another in numerous international covenants and charters.[89] This direct and somewhat narrow attack on a utilitarian theory of punishment is only one of three ways in which the concept of rights may be used to attack a utilitarian theory of punishment.

The second manner is by resorting to the concept of rights to underpin a retributive theory of punishment. Many retributive theories, to varying degrees, have attempted to seize on the notion of rights. The retributive theory which relies most heavily on the notion of rights is rights retributivism, which provides that punishment is justified because where an offender violates the rights of his or her victim the offender thereby forfeits some of his or her rights. The theory also contends that punishment must be proportionate to the offence and provides a formula for achieving this: the offender should be deprived of the same or equivalent rights to those that have been violated by the crime. Rights are equivalent when people would be indifferent to preferring the rights violated to those lost through punishment. In terms of drawing the justificatory link between crime and punishment, this is said to be found in the fact that the offender has violated the rights of another.[90]

There are, however, several problems with this theory. For example, Honderich makes the point that rights retributivism does not advance the

---

88 Intuitions can only be resorted to in the most limited of circumstances. The persuasion of such 'truths' is roughly commensurate with the incongruity of an assertion to their contrary. Given that, as we have seen, that there is no absurdity in the view that sometimes it is permissible to punish the innocent, to prove to the contrary it is necessary to substantiate the principle on the basis of some other virtue. In making this point, I am obviously rejecting an intuistic picture of morality, which proposes a plurality of principles without a unifying principle. I also do not believe that a reflective equilibrium approach (of the type proposed by J. Rawls, in "Outline of a Decision Procedure for Ethics" (1951) 60 *Philosophical Review* 177; and a *Theory of Justice* (Oxford University Press, Oxford, 1972)) is persuasive, because it too ultimately relies on intuitions to do all the hard work.

89 For example, see the *Universal Declaration of Human Rights* (1948), Article 9; *International Covenant of Civil and Political Rights* (1967), Article 9.

90 A.H. Goldman, "The Paradox of Punishment" (1979) *Philosophy and Public Affairs* 1.

justificatory link between crime and punishment any further than intrinsic retributivism.[91] The claim that one has violated the rights of another, provides no further reason beyond the simple assertion that one has acted wrongly to justify punishment. It is not as if certain wrongs, those that involve infraction of rights, are any more or less deserving of punishment. Certainly rights retributivism adverts expressly to a particular moral theory, however there is nothing inherent in the concept of rights which mandates or permits a punitive response for violation of them.

These cursory observations aside, I shall not focus directly on either of the above two rights based arguments.[92] Rather the discussion will address the third and broadest, and most persuasive, argument that has been utilised by rights proponents against utilitarianism. In its broadest form, this is simply the claim that rights based theories are the soundest moral theories and accordingly, all other moral theories, particularly utilitarianism must be rejected. If rights based moral theories are correct, it follows that utilitarianism in all of its applications, including the practice of punishment, must be rejected. In order for the utilitarian theory of punishment to take back ground, it is necessary to discredit the plausibility of rights based moral theories. If this can be done the other more specific rights based objections discussed above will also be debunked.

First I shall outline the nature of rights based theories.

(a) The Nature of Rights Based Theories

Numerous rights based theories have been advanced and as a result of the colossal, and apparently ever increasing, amount of ethical language which is expressed in the form of rights, such theories present the greatest challenge to utilitarianism. Rights talk transcends all areas of moral discourse. It is fast becoming, if it is not already, the conventional moral currency. There is no shortage of rights based theories. The main differences between them typically being the precise rights which are acclaimed, the basis of the rights, and the absolutism with which they apply. The main role of rights in deontological theories is to protect people from being compelled to do something against their wishes for the good of another or the general good. I shall mainly look at arguably the two most influential contemporary rights theories, those of Ronald Dworkin and Robert Nozick. However, many of the observations I make in relation to these theories are applicable to most other rights based theories.

---

91 T. Honderich, above n 21, 217-8.

92 That is, rights retributivism or the right not to be punished without committing an offence.

Dworkin—Concern and Respect

For Dworkin, rights are 'political trumps held by individuals',[93] which protect them from the pursuit of common goods: 'the prospect of utilitarian gains cannot justify preventing a man from doing what he has a right to do',[94] and the general good is never an adequate basis for limiting rights. He asserts that people have rights when there are good reasons for conferring upon them benefits or opportunities despite a community interest to the contrary.

According to Dworkin, in order to take rights seriously, one

> must accept one or both of two important ideas. The first is the vague but powerful idea of human dignity. This idea, associated with Kant ... supposes that there are ways of treating a man that are inconsistent with recognising him as a full member of the community, and holds that such treatment is profoundly unjust. The second is the more familiar idea of political equality.[95]

Observance of these ideals leads to the fundamental right of equal concern and respect, which is the foundation of Dworkin's rights thesis:[96] it makes sense to say a man has a right if that right is necessary to protect his dignity or his standing as equally entitled to concern and respect. To treat one with concern is to treat one as a human being, capable of suffering and frustration,[97] and to accord respect is to recognise one as a human being capable of forming and acting on intelligent conceptions of how life should be lived.[98]

Nozick—Rights Which Exist in a State of Nature

Robert Nozick's rights theory stems from his analysis of the legitimate role of the state.[99] I am not so much concerned here with the end

---

93 R. Dworkin, *Taking Rights Seriously* (4th edn) (Harvard University Press, Massachusetts, 1978), xi.

94 *Ibid*, 193.

95 *ibid*, 198.

96 *Ibid*, 199, See also R. Dworkin, "Liberalism", in S. Hampshire (ed.) *Public and Private Morality* (Cambridge University Press, Cambridge, 1979) 127, 136.

97 R. Dworkin, above n 93, 200.

98 Id.

99 He begins by imagining that no state exists and goes on to detail the type of state that is legitimate and which he believes people would mould consistent with their moral rights. Through this process he claims that we would arrive at the minimal state: a position between anarchy and a redistributive state. In this state fetters on freedom are few. Individuals have power to own and transfer property and to hire the labour of others. The state has an extremely minimalist role; its functions being confined to those which are essentially

product of this state, but rather with his picture of morality which underpins it. Nozick believes that morality is founded on rights. For him, the rights we have are those which supposedly exist in a state of nature and derive from our natural liberty. This gives rise to several distinct rights: the right to absolute control over ourselves; the right to be free from all forms of physical violations; and the right to acquire property and other resources as a result of the proper exercise of our personal rights.[100] These rights are contingent upon not violating the same rights of others. We also have the right to exact retribution against, and compensation from, those who violate our rights. Moral rights are said to act as side constraints on the actions of others and cannot be violated even to achieve greater goods.[101] Thus on Nozick's account moral rights are negative rights; there are no positive rights such as the right to welfare or health care.[102]

---

protective in nature. Basically, the state can only protect against such matters as force, theft and enforcement of contracts and so on. It cannot implement paternalist measures or coerce citizens to aid others. Thus the state cannot assume private property or impose taxes in order to, say, re-distribute resources to the disadvantaged. Roles such as this, if they are to be undertaken. must be left to private individuals and enterprises. This is the type of state, a pure form of capitalism, which it is claimed will emerge through an 'invisible hand process' by rational people acting in a self-interested manner. Nozick claims that this type of minimal state is the best manner to ensure that rights are not violated. A more powerful state would impinge upon individual rights and is hence unjustifiable unless people unanimously waive some of their rights to establish such a state: R. Nozick, *Anarchy, State and Utopia* (Blackwell, Oxford, 1974).

100 Nozick claims that the right to ownership of property and resources is derived from the fact that the person we each own is not just the tangible body parts but also consists of certain abilities which permit us to utilise resources in the world, which we will only do if we are permitted to enjoy the benefits of our inputs. Without the opportunity or possibility of appropriating and enjoying the products of our input, our talents would not be exercised and hence we would effectively be denied the full exercise of the rights to our person. However, as is correctly pointed by J. Waldron, in *The Right to Private Property* (Clarendon Press, Oxford, 1988), the desire to exercise one's talents is not contingent upon being able to benefit in some economic fashion from them and one's self-ownership confers no right to exercise one's talent for one's own benefit.

101 Nozick believes that the paramountcy accorded to the right of self-ownership and liberty is necessary to protect people from the burdensome demands of competing moral theories such as utilitarianism. For example, he believes only his rights theory can protect people from such ghastly violations as forced organ donations where the donations would maximise happiness by saving the lives of many or assisting those most in need: *Anarchy State and Utopia*, 206-7.

102 Nozick, unlike Dworkin, goes on to develop a (retributive) theory of

The rights explosion gives a running start to rights based moral theories or claims or protections couched in such language. An ethical theory or moral principle which is clearly rights based is ostensibly at an enormous advantage over other theories, such as utilitarianism, which give no natural weight to individual rights.

(b) The Case Against Rights Based Theories

One of the main problems with rights is that there appears to be no basis to stop their expansion. It seems the number of alleged rights has blossomed exponentially since the basically protective rights of life, liberty and property were advocated in the seventeenth century. Nowadays all sorts of dubious claims have been advanced by reference to them. For example, 'the right to a tobacco-free job', the 'right to sunshine', the 'right of a father to be present in the delivery room', the 'right to a sex break',[103] and even

punishment from his general moral theory. Nozick advances a communicative theory of punishment in which he claims that punishment is justified on the basis that it reconnects the offender with the correct values from which his or her wrongdoing has disconnected him or her. Punishment conveys a message from those with appropriate values to offenders, whose conduct shows that they possess incorrect values, that their conduct was wrong. The message aims to affect the criminal in a way that corresponds with the magnitude of the offence: *Philosophical Explanations* (Oxford University Press, Oxford, 1981), 363-397. Thus Nozick's theory has many similarities with Duff's theory of punishment, with a significant difference being that Nozick's theory is even more purely non-consequential, since in order for punishment to be justified on Nozick's account there is no need that this message should achieve moral transformation of the offender: punishment is 'right or good in itself; apart from the further consequences to which it might lead': *Philosophical Explanations,* above, 374. There are several specific problems with Nozick's theory of punishment. It has been noted that if it is irrelevant whether or not punishment changes the offender or not then we are still left wondering why the message must be conveyed in the first place: see N. Walker, *Why Punish?,* 81. Nozick provides a hint when he states that through punishment the correct values have some significant effect on the offender's life and makes the offender less pleased for his actions and in this way the offender is encouraged to regret the values he or she held. However as Ten points out, this is in effect no more than a subtle way of stating that the aim of punishment is to encourage regret and to achieve deterrence, which are clearly consequentialist considerations which Nozick is disqualified from resorting to: see C.L. Ten, above n 21, 42-6. Objections to any *particular* retributivist theory are not central to this discussion, thus these preliminary observations aside, I shall not elaborate on the above discussions.

103 These examples are cited by J. Kleinig, "Human Rights, Legal Rights and Social Change" in E. Kamenka and A.E. Tay (eds), *Human Rights* (Edward Arnold, Melbourne, 1978), 36, 40.

'the right to drink myself to death without interference'.[104] The 'right to die'[105] is also arguably a member of such an incredulous group. Due to the great expansion in rights talk, rights are now in danger of being labelled as mere rhetoric and losing their cogent moral force: 'an argumentative device capable of justifying anything is capable of justifying nothing'.[106]

In order for rights proponents to capitalise on the wave of support currently enjoyed by rights based theories, and for such theories to be capable of having a persuasive and meaningful role in post-philosophical moral discourse, it is necessary to provide rights a foundation which can be used to solve several key problems concerning them. These include, what is a right?; where do they come from?; what is their justification?; how can we distinguish real from fanciful rights?; when, if ever, can rights be overridden?; and which right takes priority in the event of clashing rights? Overall, rights based theories fair poorly in meeting the challenges posed by such questions.

(i) The Definition of a Right

A difficulty which has persistently plagued rights based theories is that of defining exactly what is meant by the concept of a right. Following the work of Hohfeld,[107] there is no shortage of definitions which have been advanced. McCloskey believes rights to be simply entitlements,[108] while in Sprigge's view 'the best way of understanding...that someone has a right to something seems to be to take it as the claim that there are grounds for complaint on their behalf if they do not have it'.[109] Still further, rights have

---

104 S.I. Benn, "Rights", in P. Edwards (ed.), *Encyclopedia of Philosophy* (Collier-MacMillan, 1967) vol 7, 196.

105 This supposed right has gained widespread support in the context of the euthanasia debate. One leading Australian politician even defined it as the 'most fundamental human right of all': Gareth Evans, "Death Bill Advocates to Focus on Senate" *Age* 11 December 1996, 7. The right has also received recognition in the courts: 'dying is an integral part of living...it follows that the right to die with dignity should be as well protected as is any other aspect of the right to life. State prohibitions that would force a dreadful, painful death on a rational but incapacitated terminally ill patient are an affront to human dignity': *Rodriguez v A-G British Columbia* [1994] 85 CCC (3d) 15, Cory J.

106 L.W. Sumner, above n 16, 8-9.

107 W.N. Hohfeld, defined four categories of rights: claim-rights, privileges, powers and immunities. He qualifies this by stating that only a claim-right accords with the proper meaning of the term: W.N. Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning in and other Legal Essays,* in W.W. Cook (ed.), (Yale University Press, 1919)).

108 H.J. McCloskey, "Rights: Some Conceptual Issues" (1976) 54 *Australian Journal of Philosophy* 99, 115.

109 T.L.S. Sprigge, *The Rational Foundation Of Ethics* (Routledge, 1987),

been defined as: claims and entitlements to benefit from the performance of obligations;[110] 'those minimum conditions under which human beings can flourish [as moral agents] and which ought to be secured for them, if necessary by force';[111] and the liberties each man hath, to use his own power, as he will himself, for the preservation of his own nature'.[112] Finally, Galligan defines a right as a 'justified claim that an interest should be protected by the imposition of correlative duties'.[113]

For all that, I shall not get weighed down on the issue of what is a right. Having acknowledged at the outset the important role rights have in morality, such an inquiry could only be of minor significance to the discussion at hand. Whatever the outcome of the exercise, it is highly improbable that it would have an impact upon the role and proliferation of rights in moral discourse. Even if it was concluded that it is not possible to provide a coherent definition of rights, but, rather that rights were, say, some multiformal types of claims with no common feature, it is still necessary to adequately fit the concept into an account of morality. Rights have become such an entrenched feature of the moral (and legal)[114] landscape that it is now too late to simply dismiss them on the basis that they are merely the product of faulty analysis or logic or are 'nonsense on stilts'.[115] Even if such claims may have been tenable at some earlier point, rights now have such an inextricable connection with morality that they have possibly re-shaped its meaning. Any moral theory which failed to account for such a notion is likely to be readily dismissed as being irrelevant. Thus I shall accept a right is a coherent concept.

However, for the sake of completeness I believe the following to be the correct definition of a right. A right is a presumptive benefit or protection one can assert against others.[116] Presumptive, because it is never indefeasible or absolute. By benefit, I mean a positive entitlement such as

216-7.

110 G. Marshall, "Rights, Options and Entitlements" in A.W.B. Simpson (ed.), *Oxford Essays in Jurisprudence* (Clarendon Press, Oxford, 1973) 228, 241.

111 J. Kleinig, above n 105, 44-5.

112 T. Hobbes, *Leviathan* (1651) (Blackwell, 1946), 84-5.

113 D.J. Galligan, "The Right to Silence Reconsidered" (1988) *Current Legal Problems* 69, 88.

114 See M Bagaric, "The Diminishing 'Right' of Silence" (1997) 19 (3) *Sydney Law Review* 366, 374 where I argue that the High Court has also been heavily influenced by the rights movement.

115 J. Bentham, "Anarchical Fallacies" (1824) in J Bowring (ed.), *Works* vol 2.

116 This is similar to the interest theory of rights which provides that 'A's having a right to something means that there is an aspect of A's well-being that [ie, an interest of A's] important enough to justify imposing a duty on some other person(s) in respect to that interest': A. Marmor, "On the Limits of Rights" (1997) 16 *Law and Philosophy* 1, 3.

the right to welfare. A protection is a negative entitlement, such as the right to be free from a particular violation. I do not agree with Hart's view that a right necessarily requires that the holder must be in a position to elect whether or not to exercise it.[117] It does not seem to be overly straining the language to assert that children, the mentally handicapped or even animals have rights. For example, it is appropriate to speak of the mentally disabled as having the right to have children, or children as having the right not to be physically abused, and such issues are normally discussed in terms of rights without even the hint of incoherency.[118]

(ii) The Justification For Rights and Situations Involving Clashing Rights

The Basis of Rights—Concern and Respect?

A much more serious problem which plagues rights theories is that of justifying the existence of rights. While on its face Dworkin's theory sounds tenable, if we look just below the surface we find a conspicuous lack of substance. Sure it is comforting and agreeable to claim that we are all entitled to concern (since we can all suffer) and respect (since we have the capacity to make intelligent decisions about how to live our lives) and even more comforting that this should be in equal amounts. But ignoring pleasantries for a moment, the question is why, beyond perhaps wishful thinking, are we so entitled? Furthermore why does this form the core of morality?

Concern and respect are no doubt desirable virtues, and ideally the more the better, but they would not appear to be any more important and desirable than, say, sympathy, compassion, courtesy, love, and honesty. Dworkin contends that the right to equal concern and respect is a fundamental right because it does not conflict with another person having

117 H.L.A. Hart, *Are there any Natural Rights*? (1955) LXIV *Philosophical Review Quarterly* 175.

118 See also, G. Marshall, "Rights, Options and Entitlements" in A.W.B. Simpson (ed.), *Oxford Essays in Jurisprudence* (Clarendon Press, Oxford, 1973) 228, 235; C Arnold, "Analyses of Right" in E. Kamenka and A.E. Tay (eds), *Human Rights* above n 103, 74, 80-1 (who points out that lunatics do have rights and argues that while duties must be enforceable, there is no reason that the right-holder must be the one who chooses to initiate the proceedings); and T. Regan, *The Case for Animal Rights* (Routledge, 1983). On this issue of the relationship between rights and duties, it has been argued, contrary to the Holfeldian conception, that a right is not correlative to a duty, but rather that, pursuant to the interest theory of rights, rights in fact justify the imposition of duties. Thus rights are regarded as being grounded in the interests of the right holder and duties are derived from rights: see A. Marmor, "On the Limits of Rights" (1997) 16 *Law and Philosophy* 1, 3-4.

the same right. But, as Mackie has argued, the right to be treated in a particular way is further dependent upon the right to certain opportunities of living.[119] Further, numerous other vague ideals if attributed to all, such as the 'right to be treated with compassion and honesty', would also not cause conflict, but no reason is given why they are not selected as the basis for all other rights.

More generally, Dworkin provides that 'a man has a moral right...if for some reason the state would do *wrong* to treat him in a certain way, even though it would be in the general interest to do so (emphasis added)'.[120] However, this is merely to swap one piece of rhetoric for another. Wrong: by what standard? Dworkin would do well to attempt to justify the right to equal concern and respect by developing the notions of dignity and equality which supposedly underpin this fundamental right. However he refuses to take up this challenge. Even though he frankly concedes that dignity is a 'vague'[121] ideal, he provides that he 'does want to defend or elaborate these ideas [the notions of equality and dignity], but only to insist that anyone who claims that citizens have rights must accept ideas *very close* to these (emphasis added).[122]

Accordingly, when it comes down to establishing the foundation of rights Dworkin's theory is seriously deficient. He frankly concedes that the existence of rights cannot be demonstrated, and attempts to mitigate the harm from this by merely stating that because a statement cannot be demonstrated to be true does not mean that it is not true.[123] However, while this may be so, the same reasoning could be used to defend claims about unicorns and witches.[124] 'Philosophers frequently introduce ideas of dignity, respect, and worth at a point at which reasons appear to be lacking, but this is hardly good enough. Fine phrases are the last resort of those who have run out of arguments';[125] which Dworkin (nearly) has.[126]

---

119 J.L. Mackie, "Can There Be a Right-based Moral Theory?" in (ed.), R. French et al., *Studies in Ethical Theory* (1978, vol 3).

120 R. Dworkin, above n 93, 139

121 *Ibid*, 198.

122 *Ibid*, 99. In R. Dworkin, in "Liberalism", S. Hampshire (ed.), *Public and Private Morality* (Cambridge University Press, Cambridge, 1979) 127, 136 Dworkin outlines what he believes is entailed by the notion of equality, but still does not articulate how and why this could form the basis of morality.

123 R. Dworkin, above n 93, 81.

124 See A. MacIntyre, "A Critique of Gerwith and the Notion of Rights" in L.P. Pojman (ed.), *Ethical Theory* (Wadsworth Publishing Co, Belmont, 1995), 715, 717.

125 P. Singer, "All Animals are Equal", in P. Singer (ed.), *Applied Ethics* (Oxford University Press, Oxford, 1986) 215, 228.

126 Perhaps noting the difficulty in identifying what lies at the core of the rights, some have moved, perhaps somewhat hastily to the issue of what it is about

Dworkin has one more attempt to explain where rights come from and why they exit:

> *So if rights make sense at all*, then the invasion of a relatively important right must be a very serious matter. It means treating a man as less than a man, or as less worthy of concern than other men. The institution of rights rests on the conviction that this is a grave *injustice*, and that it is *worth paying the incremental cost* in social policy or efficiency to prevent it (emphasis added).[127]

The first point to note is that it is not open for Dworkin to simply assume that it makes sense to speak of (non-consequentialist) rights. Secondly, it is also not playing by the rules for him to resort to the, nebulous, notion of justice to justify (or buttress) his theory of rights without a meaningful elaboration of this concept. Finally, he gives no indication about the currency he is employing when he asserts that it is worth paying 'the cost' to acknowledge rights. This last point is also fatal to his suggestion regarding when it is permissible to limit rights.

Dworkin and Clashing Rights

Dworkin states that there are three situations where a right may be limited.

> First, the Government must show that the values protected by the original right are not really at stake in the marginal case, or are at stake in some attenuated form. Second, it might be shown that if the right is defined to include the marginal case, then some competing right, in the strong sense defined earlier, would be abridged. Third, it might be shown that if they were so defined, then the cost to society would not be simply incremental, but would be of a degree far beyond the cost paid to grant the original right, a degree great enough to justify whatever assault on dignity or equality might be involved.[128]

The first of these suggestions is of little guidance, since it, effectively, deals with situations where rights are on second glance not applicable. The second suggestion refers to Dworkin's supposed distinction between rights in a strong sense and in a weak sense. Strong rights are those which are

---

a particular demand that makes it an instance of a more fundamental right: see A. Harel, "What Demands are Rights? An Investigation Into The Relation Between Rights and Reasons" (1997) 17(1) *Oxford Journal of Legal Studies* 101.

[127] R. Dworkin, above n 93, 199.

[128] *Ibid*, 200.

wrong to interfere with and weak rights refer to activities which are not wrong to pursue.[129] However, even if one ignores the questionable nature of such a distinction,[130] it hardly addresses the issue here since even on Dworkin's account rights in the sense of entitlements which place limits on actions of others are all strong rights. The final stipulation is the most far reaching and thus promising, but in the end is devoid of content, since it prescribes a balancing process, however does not assign a unit of measurement by which the 'cost' can be measured.

Nozick—The Basis For Rights and Clashing Rights

In regard to identifying the content of rights, Nozick's theory is far more precise. As was detailed above, Nozick particularises the rights which he claims we possess. He also has a straight forward answer regarding the problem of conflicting rights. He asserts that given that rights are negative in character, serving merely as side-constraints, requiring others to refrain from certain types of actions; as opposed to requiring agents to perform certain acts, they are absolute and never clash.

However, Nozick suggests nothing which overcomes arguments which have been persuasively advanced against the notion of absolute rights.[131] The absurdity of absolute rights is evidenced by the extreme lengths some have gone to in order to attempt to justify such a notion. For example, in search of an absolute right it has been stated that 'the right of a mother not to be tortured to death by her son is absolute'.[132] However even such extreme examples fail. One could hardly begrudge a son torturing his mother to death if this is only way to save the lives of all his other relatives whom the mother was about to unjustifiably kill.

Further, Nozick's theory fails at the first hurdle. It fails to provide any justification for his set of fundamental rights. He asserts that his list of rights permit a person the capacity to shape his or her life in order to have a meaningful life.[133] This, however, ignores the fact that in order to have anything approaching a meaningful life requires the provision of certain

---

129 *Ibid*, 188-191. The example Dworkin gives of a weak right is the right of a captured enemy soldier to attempt to escape, even though we are not wrong to attempt to foil the escape attempt.

130 Weak rights are not accompanied by duties, and it is arguable whether such interests are rights in the conventional sense (see discussion above regarding the definition of a right) as opposed to merely being privileges.

131 For example, T. Campbell, "Democracy, Human Rights and Positive Law" (1994) 16 *Sydney Law Review* 195, 200-1, who points out that absolutist formulations of rights are fine in political rhetoric, but in the real world must be qualified, especially in cases of conflicting rights.

132 A. Gewirth, *Human Rights: Essays on Justification and Applications* (The University of Chicago Press, 1982), 233.

133 R. Nozick, above n 99, 50.

necessities, such as shelter, food and healthcare; which Nozick denies the right to. The claim that certain rights are supposedly to be found in a state of nature is also dubious.[134] The same could be claimed of any other so-called right, such as the right to a sex break. And just because something is natural is not of itself morally significant. It is also natural to display jealously and anger, but this hardly makes it justifiable to act upon or encourage such impulses. To call a right a natural right is no more compelling than to label it a human right.[135] Ultimately, rights appear to lack a foundation and an origin. Not surprisingly, even some rights proponents have been conceded that rights are 'inherently controversial'.[136]

The other significant defect of Nozick's account is that it is too revisionary. Acceptance of it would require the abandonment of too many established moral principles and duties.[137] The maxim of positive duty (which provides that we must assist others in serious trouble, when assistance would immensely help them at no or little inconvenience to ourselves),[138] would obviously be the first to go. On Nozick's view morality does not require us to save the baby in the puddle. While the maxim is not set in concrete, we would need far more persuasive reasons than Nozick's theory to retract our commitment to it.

---

134 See also T. Scanlon, "Nozick on Rights, Liberty, and Property" (1976) 1 *Philosophy and Public Affairs* who convincingly argues that the rights defined by Nozick are not those that are likely to be found in a state of nature.

135 The term 'human' right is a unhelpful. It suggests that species membership alone is adequate to create and safeguard a right. As Benn points out, 'if such rights are human rights, [why] should someone's, [say], being guilty of a crime deny him enjoyment of them? Do people enjoy such rights as men and women, or only as well behaved men and women—*quamdiu se bene gesserint*?: S I Benn, above n 15, 62-63. Bentham argued against the existence of natural moral rights on the basis that in cases of dispute there is no objective means to ascertain the content and scope of such rights. This is in contrast to legal rights, where disputes are resolved by the terms of the relevant law or by the courts. It is also in contrast to utilitarianism which provides a formula for determining the priorities of competing interests (see discussion below). Bentham claimed that all rights were contingent upon and derived from positive law, and that the view of rights independent to law was contradictory, like 'cold heat': J. Bentham, "Supply Without Burthen", in *Bentham's Economic Writings* (W. Stark edn. 1952) 279).

136 A. Marmor, "On the Limits of Rights" (1997) 16 *Law and Philosophy* 1, 16.

137 This point is also made by Marmor, *Ibid*. 7-8.

138 See M. Bagaric, "Active and Passive Euthanasia: Is There a Moral Distinction and Should There be a Legal Difference" (1997) 5 *Journal of Law and Medicine* 143.

Marmor and Clashing Rights

Other attempts to deal with the problem of clashing rights fair no better than the above accounts. For example, it has been suggested that rights can be limited or overridden only in order to secure some other more important or pressing right. This has been dubbed the Newtonian conception of the limit of rights, because it operates in a similar way to the Newtonian law of inertia: 'a right will continue to be in force so long as it does not collide with another right which conflicts with it'.[139] Thus clashes of rights are resolved by securing 'the set of rights that would *maximally* satisfy the most extensive set of rights for each person, which is compatible with the similar rights of others'.[140] However, the problem with such an approach is that there , again, is no barometer which can be used to measure the respective importance of rights. In discussing the paradox of the right to do something which a person ought not to, Marmor gives the example of the right to get married even where it seems that it is a 'wrong' decision because it will ruin both lives. The Newtonian conception of rights explains this on the basis that the right of personal choice is so paramount that 'we think that it is more *important* for a person to chose her spouse for herself than to choose correctly (emphasis added)'. But the question remains: important by what standards? It could be equally argued that one has the 'right' to prevent a foredoomed marriage, where the marriage would ruin the lives of the prospective husband and wife, because it is more *important* that people have overall fulfilling and enjoyable lives than be permitted to select their partner.

The Newtonian conception is ultimately rejected by Marmor, on the basis that rights are not limited only when they clash with other rights, but have internal limits imposed well before this on account of the fact that rights impose duties on others and these burdens must be taken into account at the outset when the interests of the potential right holder and others are compared.[141] This consideration not only determines the limits of a right, but also whether it exists at all.[142] 'A's right to x can only be justified, initially, if we think that A's interest in x is important enough to warrant imposing a duty on others, and *only to that extent.* Namely to the extent that the burden involved in the imposition of the duties does not out-weigh the importance of the interest in question'.[143] However, there are two problems with this analysis. First, there is again no standard by which importance is to be evaluated. The tenable argument that the importance of a right is commensurate with the its value to the right holder is rejected by Marmor

[139] A. Marmor, above n 136, 1-2, 11.

[140] *Ibid*, 9.

[141] *Ibid*, 10-1,1-2

[142] *Ibid*, 11.

[143] *Ibid*, 10.

on the grounds that rights benefit not only the right holder but promote the common or general good.[144] This highlights the second difficulty, which is that in weighing the interests and impositions that Marmor claims are integral to the determination of a right, a utilitarian calculus must be engaged in. He denies this on the basis that the 'cost-benefit analysis is not necessarily a quantitative matter; the *intrinsic* values and relative importance of the interests in question matter too'.[145] However, by failing to elaborate on the nature of these 'intrinsic' values the theory becomes vacuous.

What Rights do we Have?

Another difficulty with rights based theories is that we are typically left with no guidance concerning the specific rights we have. Dworkin acknowledges the obvious: that 'it is much in dispute...what *particular* rights citizens have (emphasis added)',[146] however offers surprisingly little to deal with this. For Dworkin, the rights that we have are those which ensure the principle of equal concern and respect is upheld. He states that the central issue is what inequalities are justified, and goes on to state that two rights flow from the abstract right to equal concern and respect. The first is the right to equal treatment; defined as the right to the same distribution of goods or opportunities as anyone else. The second, which he claims is more fundamental then the first, is the right to treatment as an equal, that is the right to equal concern and respect in how these goods and opportunities are distributed.[147]

However this fails to advance his position any further. The two more specific rights Dworkin expounds are just as vague as his general principle. Both employ the notion of concern and respect and the formula is even more generalised by resort to the ill-defined notion of equality. Such vague aspirations 'dissolve into generalised moral values which cannot function as rights by giving us a relatively objective and politically uncontroversial way of determining entitlements by reference to an authoritative system of norms'.[148]

The vacuousness of rights based theories and their propensity to confuse is no more apparent than in the punishment debate where a 'right to be punished' has been suggested.[149] Morris declares that the right to be

144 *Ibid*, 11-2.
145 *Ibid*, 13.
146 R. Dworkin, above n 93, 184.
147 *Ibid*, 273.
148 T. Campbell, *Justice* (MacMillan Education, London, 1988), 56.
149 See R.A. Duff, above n 36; H. Morris, Persons and Punishment in S.E. Grupp (ed.), *Theories of Punishment* (Indiana University Press, Ontario, 1971) 76, 92.

punished stems from 'a right to be treated as a person which is a fundamental human right belonging to all human beings by virtue of their being human. It is also a natural, inalienable, and absolute right'.[150] This misses the point that at the minimum a right is a positive attribute. It also serves to highlight the fuzzy reasoning and fantastic claims that can be made due to the lack of coherency and precision of rights based theories, which simply offer no workable mechanism to distinguish between real and illusory rights.

(c) The Instability of the Rights Thesis—Consequences The Ultimate Consideration

A further flaw in the theories of both Dworkin and Nozick is that by conceding that in some situations consequences prevail, their respective theories become unstable. Despite his absolutist tones, Dworkin accepts that it is correct for a government to infringe on a right when it is necessary to protect a more important right, or to ward off 'some great threat to society'.[151] In a like manner, Nozick states that teleological considerations would take over to 'avert moral catastrophe'.[152] But both fail to state, even loosely, at what point we reach a great threat to society or a moral catastrophe and consequentialist considerations legitimately 'kick in' to guide conduct.

By making this concession, which is necessary to avoid the even less plausible position that rights are absolute, the theories become irrelevant. When consequential considerations are admitted as being relevant the theories become hybrid and the main theoretical advantage of a deontological theory, the absolute protection given to people against certain intrusions, is forsaken. The problem is heightened because, in both cases, we are given no guidance as to when consequentialist considerations become overriding. Due to the lack of specificity in this regard, it could be argued that even Dworkin and Nozick accept that punishing the innocent is permissible where the lives of many are at stake. At this point rights theories collapse: they can neither rely fully on the theoretical justifications

---

150 H. Morris, Id. This a not an uncommon retributive theme; for example, see also RA Duff, in *Trials and Punishment* 263-271. At the heart of most retributive claims regarding the right to be punished is the view that this right stems from the dignity of each person, which requires that each person is to be treated as a responsible moral agent. This, however, *assumes* that the natural and appropriate response to crime is punishment—which is exactly the link which retributivists must *prove*. Duff also who makes the further claim that criminals not only have the right to be punished, but also *want* to be punished. This however, is completely at odds with the extreme lengths offenders normally resort to in order to avoid apprehension and detection.

151 R. Dworkin, above n 93, 199-202.

152 R. Nozick, above n 103, 95.

of deontological or consequentialist theories.

In summary, the present state of affairs regarding rights based moral theories is still best summarised by Hart: 'it cannot be said that we have had...a sufficiently detailed or adequately articulated theory showing the foundation for such rights and how they are related to other values. Indeed the revived doctrines of basic rights are...in spite of much brilliance still unconvincing'.[153]

Utilitarianism and Rights

Utilitarian Justification For Rights

As I have stated, the core of perhaps the most damaging criticisms of utilitarianism is that it is antagonistic to the concept of rights. It is claimed that utilitarianism does not take seriously the distinction between human beings, because it prioritises net happiness over individual sacrifices, and hence fails to protect certain rights and interests that are so paramount that they are beyond the demands of net happiness.[154] It is not difficult to see the basis for this criticism. Utilitarianism is a maximising principle, the aim being to maximise the net happiness. On the other hand the notion of rights is individualising, the purpose being to accord each individual certain interests.

Rights do however have a place in a utilitarian ethic, and what is more it is only against this background that rights can be explained and their source justified. Utilitarianism provides a sounder foundation for rights than any other competing theory. For the utilitarian, the answer to why rights exist is simple: recognition of them best promotes general utility.[155] Their origin accordingly lies in the pursuit of happiness. Their content is discovered through empirical observations regarding the patterns of behaviour which best advance the utilitarian cause. The long association of utilitarianism and rights appears to have been forgotten by most. However, over a century ago it was Mill who proclaimed the right of free speech, on the basis that truth is important to the attainment of general

---

153 H.L.A. Hart, *Essays in Jurisprudence and Philosophy* (Clarendon Press, Oxford, 1983), 195.

154 J. Rawls, above n 13.

155 According to Mill, rights reconcile justice with utility. Justice, which he claims consists of certain fundamental rights, is merely a part of utility. And 'to have a right is to have something society ought to defend...if [asked why]...I can give no other reason than general utility': J.S. Mill, above n 43, 251.

happiness and this is best discovered by its competition with falsehood.[156]

Difficulties in performing the utilitarian calculus regarding each decision, make it desirable that we ascribe certain rights and interests to people, which evidence shows tend to maximise happiness[157]—even more happiness than if we made all of our decisions without such guidelines. Rights save time and energy by serving as shortcuts to assist us in attaining desirable consequences. By labelling certain interests as rights, we are spared the tedious task of establishing the importance of a particular interest as a first premise in practical arguments.[158] There are also other reasons why performing the utilitarian calculus on each occasion may be counter productive to the ultimate aim. Our capacity to gather and process information and our foresight are restricted by a large number of factors, including lack of time, indifference to the matter at hand, defects in reasoning, and so on. We are quite often not in a good position to assess all the possible alternatives and to determine the likely impact upon general happiness stemming from each alternative. Our ability to make the correct decision will be greatly assisted if we can narrow down the range of relevant factors in light of pre-determined guidelines. This is precisely the practice which is employed by our legal system.

The aim of the system is to attain justice. In order to achieve this, judges do not pursue this result by whatever means they believe appropriate to the case at hand. They are bound by procedural guidelines, such as the adversarial system and rules of evidence, and substantive principles, such as the presumption of innocence and the notion of unconscionability,[159] which represent processes which experience has shown if followed will generally produce the just result at the end of the day.[160] Similarly in the case of

---

156 J.S. Mill, above n 43.

157 These rights, however are never decisive and must be disregarded where they would not cause net happiness (otherwise this would be to go down the rule utilitarianism track).

158 See J. Raz, *Morality of Freedom* (Oxford University Press, Oxford, 1986), 191. Raz also provides that rights are useful because they enable us to settle on shared intermediary conclusions, despite considerable dispute regarding the grounds for the conclusions. See also A. Marmor, above n 136, 1, 15.

159 For example, see *Hospital Products Ltd. v United States Surgical Corporation* (1984) 156 CLR 41; *Muschinski v Dodds* (1985) 160 CLR 583; *Baumgartner v Baumgartner* (1987) 76 ALR 75.

160 This is not dissimilar to Rawls' concept of pure procedural justice: J. Rawls, *A Theory of Justice* above n 13, 453-462. Although, the set of legal norms is not derived from 'the original position', like Rawls' theory, substance is essentially compromised for form: the end result is supposedly justified simply on the basis that what is thought to be a desirable process has been undertaken. Thus even when the guilty go free, this is justified where the accused has been through the trial process.

utilitarianism. History has shown that certain patterns of conduct and norms of behaviour if observed are most conducive to promoting happiness. These observations are given expression in the form of rights which can asserted in the absence of evidence why adherence to them in the particular case would not maximise net happiness.

Thus utilitarianism is well able to explain the existence and importance of rights. It is just that rights do not have a life of their own (they are derivative not foundational), as is the case with deontological theories. Due to the derivative character of utilitarian rights, they do not carry the same degree of absolutism or 'must be doneness' as those based on deontological theories. However, this is no criticism of utilitarianism, rather it is a strength since it is farcical to claim rights are absolute. There are times when it is permissible to kill, lie and cheat. Any theory alleging the contrary is guilty of dogmatism.

Another advantage of utilitarianism is that only it provides a mechanism for ranking rights and other interests. In event of clash, the victor is the right which will generate the most happiness. The interests which are normally targeted by criminal sanctions, such as freedom, reputation and property ownership, are generally very high on most people's scale of things that make them happy, thus there are strong utilitarian reasons for not encroaching on them. Significant counter benefits (such as deterrence and incapacitation, and so on) must be produced to justify the violation of such interests. This is almost infinitely more so where a person is not guilty of an offence, given the potential violation of other interests which weigh heavily on the happiness register of most, such as the need for transparency and accountability in the legal system. Thus there is not only a place, but a high ranking for the right not to be punished without a prior determination of guilt in the utilitarian calculus. It is rare that happiness will be maximised by violating this right. This underpinning and approach to the problem of punishing the innocent provides the greatest degree of protection against such a practice.

Utilitarianism and the Separateness of Persons

This type of reasoning extends to fend off not only problems associated with rights, but also to counter more generally attacks which are based on the claim that utilitarianism does not accommodate the separateness of persons. Take for instance, Williams' classic Jim and Pedro example which aims to show that utilitarianism fails to accord sufficient weight to our integrity. Jim is a botanist on an expedition in a small South American town who the ruthless government regards as an honoured visitor from another land. He goes into town and sees twenty Indians tied up. Pedro, the captain in charge, explains that the Indians are a random group of inhabitants who after recent protests against the government are about to be executed to

deter others from protesting. Since Jim is an honoured guest, Pedro offers him the 'privilege' of killing one of the Indians himself. If he accepts, as a special mark of the occasion, the other Indians will be spared. If he refuses they will all be killed. Jim realises it is impossible to take the guns and kill Pedro and the large number of other soldiers. The Indians and other soldiers understand the situation, and the Indians are begging for him to take up the offer.[161]

Williams argues that if Jim was a utilitarian he would be required to kill the Indian. Williams' quarrel is not necessarily with the result that utilitarianism commits one to (in fact he has subsequently stated that he too would shoot the Indian), but with the reasoning process employed by the utilitarian to resolve the dilemma, and consequently the fact that one can be *certain* that killing the Indian is indeed the right choice. Williams contends that utilitarianism cuts out considerations which most would think integral to such cases, such as the idea that each of us is specially responsible for what he or she does, rather than what others do. This makes integrity unintelligible, because it fails to appreciate the relationship between a man and his projects. Utilitarianism fails to accept that 'among the things that make people happy is not only making other people happy, but being taken up or involved in a vast range of projects...[such as being] committed to a person, a cause, an institution, a career, one's own genius, or the pursuit of danger'.[162]

However this fails to recognise the important role that the pursuit of projects and commitments and their accomplishment have in promoting happiness. True it is that, like rights, in the utilitarian scheme of things projects have no intrinsic or absolute value. But, surely this must be correct. An aim or pursuit is not justified simply because one describes it as a project. Hitler had a project; but so what.

Williams accepts that the general aim of maximising happiness does not require the direct pursuit of this goal at every point along the way and that people with projects are perhaps happier than those without, and accordingly that utilitarianism can ascribe some weight to projects. However, he argues that this is not an adequate response, because ultimately our ability to pursue our projects is subject to the innumerable projects of others which our actions may affect:

> The utilitarian response is to neglect the extent to which [one's] actions and his [or her] decisions have to be seen as the actions and decisions which flow from the projects and attitudes with which he [or she]

---

161 B. Williams, "A Critique of Utilitarianism", in J.C.C. Smart and B. Williams (eds), *Utilitarianism: For and Against* (Cambridge University Press, 1973), 99.

162 *Ibid*, 112.

is most closely identified. It is thus, in the most literal sense, an attack on his integrity.[163]

Thus at the heart of Williams' objection is that generally we should only be responsible for the consequences we have orchestrated, and that we cannot be expected to drop or comprise projects which may be so strongly held to be defining of our lives simply because the utilitarian sum may happen to come down against us.

But utilitarianism does give considerable weight to such considerations, even more than Williams is prepared to accept. The pursuit of projects is integral to the attainment of personal happiness. And we cannot at every single point be expected to save the world: we simply do not know; and an attempt to do so would be self-defeating. But one thing we do know is what works for each of us and accordingly the collective pursuit of our individualist aims is at most points likely to be the best method of maximising happiness. This is one reason that utilitarianism attaches an enormous amount of weight to personal liberty.[164] This virtue, like all others, is not absolute and must be forsaken on rare occasions. One being where the maxim of positive duty applies. And it is for this reason that Jim ought to shoot. He can demonstrably assist a large number of others by positively interjecting.

Another supposed utilitarian failing that Williams claims is exposed by the Jim and Pedro example is the assuredness and conclusivity that it deals with complex moral dilemmas. Williams claims that resolution of the Jim and Pedro dilemma requires consideration of several difficult issues, including the distinction between 'my killing someone, and its coming about because of what I do that someone else kills them' and how much it matters that the people at risk are actual as opposed to future or elsewhere.

---

163 *Ibid*, 116-7.

164 The most famous statement of this is by J.S. Mill: 'the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. The only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant': J.S. Mill, above n 43, 135. The courts too have heavily endorsed the central role of personal liberty: 'the right to personal liberty is...the most elementary and fundamental of all common law rights. Personal liberty was held by Blackstone to be an absolute right vested in the individual...he warned "of great importance to the public is the preservation of this personal liberty: for if once it were left in the power of any...magistrate to imprison arbitrarily...there would soon be an end of all other rights and immunities"': *Williams v R* (1986) 161 CLR 278, Mason CJ and Brennan J at 292. More recently, see Lord Mustill in his dissenting judgement in *R v Brown* [1993] 2 WLR 556, 600.

On the contrary, the fact that utilitarianism provides definite answers to difficult moral situations is an enormous advantage of the theory; not a weakness. It is not to the point that utilitarianism 'cuts out a kind of consideration which for some others makes a difference to what they feel about such [a case]'. Rather, utilitarianism is the only theory which cuts through the verbiage and distractions to provide a coherent answer to complex moral dilemmas.

The Decisiveness of Consequences

As we have seen, ultimately all moral theories at some point bow to the weight of consequences. Logically it must follow that at the bottom this is what matters most. This being the case it is nonsensical to allow other considerations to have a primary role in our moral reasoning.

Soccer, Rights and Consequences

In this respect, an analogy may again be drawn with soccer. In evaluating how good a soccer team plays many different considerations are relevant: how much speed and endurance the players have; their skill in passing the ball; their ability to cross the ball; their caginess in playing the offside rule; their natural brilliance and flair; their ability to withstand pressure (including penalty shoot outs); their defensive, offensive and midfield proficiency; the tactical prowess of the coach; and so on. While all of these considerations are important in determining how good any team is, in the end, in respect to any particular match the only thing that matters is the scoreboard. Maximum points can be gained in relation to all the relevant indicia that go to making a team a good one and indeed to that team playing a good game, but ultimately if the side fails to put the ball in the net more times than its opponent this comes to nought. All that really matters is the scoreboard *outcome* at the end of the game. Every player and passionate fan would gladly prefer to win a game despite fairing poorly on the indicia that are generally used to evaluate a team, rather than playing a 'perfect' match and having an unlucky loss. For example, in the quarter final of the 1990 World Cup, Brazil (who were rated as clearly the best side in the world) totally dominated Argentina in general play, with almost the total game being played in its attacking half. Brazil played superbly, and as a result had countless scoring opportunities, and on several occasions hit the cross bar. Argentina were totally outplayed, however, managed to convert its single counter attack into an unlikely goal to win the match. Despite Brazil's on field dominance, virtually the whole nation went into mourning after the match, while the Argentineans celebrated long and hard.

Everyone knew that Brazil was the better side; they should have won and thoroughly deserved to do so. But in the end, all that mattered was the outcome. So too with morality: interests such as rights, integrity are all important barometers regarding the appropriateness of our moral

judgements, but ultimately they must make way for the ultimate gauge—the consequences of our actions.

While on soccer, it also emphasises why utilitarianism does not require that at every point we should seek to maximise happiness. A team, no matter how talented, would be unlikely to win a single match if each player shot at goal on every occasion he or she had control of the ball.

Strangely, nearly moral theories which ostensibly place a premium on virtues other than consequences accept that at some point these other virtues must be subordinated to consequential considerations, however persist in denying that consequences are ultimately the dominant moral consideration. This, I suspect, is because of a belief that consequences are somewhat fluky and have little connection with processes and principles observed along the way. However, this ignores that nearly always the side with the best defence, midfield, attack, and with the most skilful players, and so on will score more goals. Thus we are still justified in attaching considerable weight to matters such as skill, speed and technical proficiency. And most importantly, the basis for according weight to such matters is evident. So too, a utilitarian theory of morality provides a secure foundation for interests such as integrity and rights, including, of course, the right not to be punished unless one is guilty of an offence.

Conclusion

The main theoretical criticisms of a utilitarian theory of punishment are unconvincing. Sure, the utilitarian is committed to punishing the innocent in rare circumstances. However, on a post-philosophical consideration of this, one's conscience is not so unduly disturbed that it must follow that the underlying theory must be erroneous. For, as we have seen, we are prepared in other contexts to make decisions which even more seriously violate the interests of individuals for the good of the whole. Further, retributivists too must accept that any system of punishment will invariably result in some innocent people being punished. The response that this is permissible because it is unintended, though foreseen, is wanting because there is no principled basis for a morally relevant distinction between acts which are intended and those which are foreseen.

The criticism which underlies most attacks on utilitarianism is that it fails to pay sufficient weight to individual rights and interests, and thereby permits all types of horrible practices. This has lead to a proliferation in, and widespread support, for rights based moral theories. Despite their promise, such theories lack substance. Ironically, only utilitarianism adequately justifies the notion of rights and provides coherent answers to difficulties confronted by rights based theories.

Accordingly, the main theoretical attacks on a utilitarian system of

punishment are unconvincing. The time is now right to re-assess the empirical data regarding the success of utilitarian based sentencing objectives, such as deterrence and rehabilitation. If this is promising, the utilitarian theory of punishment should once again prevail and pave the way for future sentencing practices and guidelines.