

REVENGE AND VICTIM JUSTICE

A Philosophical Analysis and Evaluation

by

Charles Karoly Bako Barton

A thesis submitted for the degree of Doctor of Philosophy

of

The Australian National University

May 1996

Statement

This is to certify that, unless otherwise indicated, this thesis is entirely my own work. It is the result of research carried out by me while a candidate for the degree of Doctor of Philosophy, in the Division of Philosophy and Law, Research School of Social Sciences, at The Australian National University.

A handwritten signature in cursive script that reads "Barton".

Charles Barton

1 May 1996

Acknowledgments

In writing this dissertation I have relied on many people whose support I would like to acknowledge. Foremost among them are Richard Sylvan, Michael Tooley, Roy Perrett, John Braithwaite, and Karen van den Broek. I sincerely thank them for their time and effort, not only in terms of their comments and suggestions, but also in terms of their support and encouragement through difficult times. I also wish to thank the many people who have commented on my ideas and arguments in the context of various seminar presentations in Australia and New Zealand. Finally, I wish to thank my dear friends for their moral and practical support. It has been important for me to know that they were there, and that they cared. Many of these people, one way or another, have helped to bring about the successful completion of this project. It is now my supreme pleasure to share with them the credit and satisfaction.

Abstract

Contemporary Western society's legal ostracism and social condemnation of revenge could hardly be more complete, or stronger. Not only are there stiff penalties for 'taking the law into one's own hands' - never mind that the law may be misguided, impotent, corrupt or wrong and, therefore, unable or unwilling to deliver justice, - according to public belief in the Western world, revenge is always wrong and it has nothing to do with justice; or worse, revenge undermines justice; or still worse, revenge is the very antithesis of justice. The purpose of this dissertation is to show that such views are wrong, that revenge is a form of justice, and that serious consideration should be given to allowing institutionalised forms of revenge within the frameworks of criminal justice systems.

The thesis contains six chapters. The first three critically examine the nature of revenge and its relationship to other important concepts in the philosophy of punishment. In the first chapter I examine the most commonly held beliefs and rationales for the unquestioned rejection of revenge in contemporary Western societies. Much of the philosophical literature on revenge is dominated by religious, philosophical and culturally embedded beliefs which exaggerate the dangers and the retributive element in revenge to the point of vilification, if not hysteria. The task of the first chapter, therefore, is to show that there are no philosophically defensible foundations to underpin those fears, attitudes and dogmas. In the second chapter I provide context for, and motivate departure from, previous accounts of revenge by continuing the critical survey of the philosophical literature on revenge in relation to retribution and the principles of justice. I clarify the concept of retribution and argue that it is independent from the principles of justice, such as the principle that the innocent should not be punished. I also challenge attempts to identify retribution with revenge, as well as attempts to place them into

different moral categories. In the third chapter I discuss the modalities and the process of punishment and construct a map of the logical space of punishment with a view to defining revenge and giving a visual representation of its relation to other key punishment concepts in logical space. I define revenge as personal retribution which is typically accompanied by feelings of indignation, anger and resentment for wrongs suffered in one's personal domain of care and concern.

The remaining three chapters evaluate the rational and moral status of revenge and argue its relevance to justice. In Chapter 4 I first discuss the importance and forms of rationality in order to prepare the ground for defending the rationality of revenge. I then consider the collective rationality of group revenge among the Maori of New Zealand and the collective rationality of practices of revenge as social institutions of law enforcement in societies with no centralised system to rely on. In the fifth chapter I give detailed consideration to the claim that revenge cannot be rational from the individual agent's point of view. I argue on the basis of distinctions developed between various forms of rationality that revenge can be rational, and that it often is, both directly and indirectly, at both formal and substantive levels.

Finally, in Chapter 6 I defend the moral appropriateness of legal institutionalised revenge from a victim justice perspective. I argue that in many cases victim justice requires acknowledgment and satisfaction of the legitimate personal needs of victims in terms of adequate restoration, apologies and/or retributive justice for the wrongs committed against them. Such justice needs to be secured without victims being trivialised, moralised or patronised on account of their needs and feelings. This can be achieved in a controlled and civilised manner through legal institutionalised forums. Such forums of empowerment, I shall argue, can assure not only superior forms of justice to offenders and victims alike, and a scope for institutionalised revenge, but also the necessary checks and balances by means of which the problematic aspects can be eliminated from revenge.

Contents

<i>Statement</i>	<i>iii</i>
<i>Acknowledgments</i>	<i>v</i>
<i>Abstract</i>	<i>vii</i>
<i>Detailed Table of Contents</i>	<i>xi</i>
<i>Preface</i>	<i>xiii</i>
Introduction	1
1 The anti-revenge paradigm	9
2 Retribution, revenge and justice	45
3 Revenge: definition and domain	85
4 Rationality and the collective rationality of revenge	103
5 The individual rationality of revenge	157
6 Victim justice and the morality of institutionalised revenge	199
Summary and Conclusions	243
Bibliography	251

Detailed Table of Contents

Preface		xiii
Introduction		1
1	The anti-revenge paradigm	9
	1.1 Introduction	9
	1.2 Perversity, evil, vindictiveness and revenge	10
	1.2.1 Revenge as losing sight of its good end	11
	1.2.2 Revenge as a perversion of its good end	12
	1.2.3 Revenge as an evil vindictive passion	15
	1.3 The emotional tone of revenge	19
	1.3.1 Revenge and the taking of pleasure in the suffering of another <i>simpliciter</i>	21
	1.3.2 Revenge and the taking of pleasure in the suffering of an aggressor	25
	1.4 Punishment and revenge	29
	1.5 Conclusion	42
2	Retribution, revenge and justice	45
	2.1 Introduction	45
	2.2 The definition of retribution	46
	2.3 The principles of offender justice	49
	2.4 The relationship between retribution and revenge	59
	2.4.1 Revenge and retribution as identical	60
	2.4.2 Revenge and retribution as different species	66
	2.5 Conclusion	84
3	Revenge: definition and domain	85
	3.1 Introduction	85
	3.2 The modalities of punishment	85
	3.3 The process of punishment	87
	3.4 The logical space of punishment	88
	3.4.1 A motive-based typology of punishment	88
	3.4.2 The basic forms of punishment	91
	3.5 Revenge: definition and domain	95
	3.6 Conclusion	100

4	Rationality and the collective rationality of revenge	103
	4.1 Introduction	103
	4.2 Rationality	104
	4.2.1 The significance of rationality	104
	4.2.2 Formal and substantive rationality	110
	4.2.3 Direct and indirect rationality	125
	4.3 Collective rationality and revenge	146
	4.3.1 The collective rationality of <i>utu</i> as group revenge	147
	4.3.2 The collective rationality of revenge as an institutionalised practice	152
	4.4 Conclusion	155
5	The individual rationality of revenge	157
	5.1 Introduction	157
	5.2 The case against revenge	158
	5.3 The backward-looking element in revenge and rationality	160
	5.4 Honour and revenge	164
	5.5 Revenge and the material/non-material distinction	169
	5.5.1 Revenge and material interests	170
	5.5.1.1 Immediate costs and benefits in revenge	170
	5.5.1.2 Long term interests and indirect rationality in revenge	173
	5.5.1.3 Social sanctions and direct rationality in revenge	181
	5.5.1.4 Material versus overall rationality in revenge	184
	5.5.2 Revenge and non-material interests	186
	5.6 Conclusion	196
6	Victim justice and the morality of institutionalised revenge	199
	6.1 Introduction	199
	6.2 A defence of moral retributivism	200
	6.2.1 The thesis of moral retributivism	201
	6.2.2 A moral retributivist justification of punishment	205
	6.2.2 Mercy and retributive liability for punishment	214
	6.3 Victim justice and empowerment	217
	6.4 Victim involvement and the dangers of revenge	238
	6.5 Conclusion	241
	Summary and Conclusions	243
	Bibliography	251

Preface

If legal justice perishes, then it is no longer worth while for men to remain on this earth.

Immanuel Kant¹

Now my hypothesis is not so much that the court is the natural expression of popular justice, but rather that its historical function is to ensnare it, to control it and to strangle it, by re-inscribing it within institutions which are typical of a state apparatus.

Michel Foucault²

Justice is too important to be left to the judiciary.

Sir Ludovic Kennedy³

Justice is indeed very important and to nobody else is it more important, perhaps, than to innocent people who are accused, and sometimes convicted, of crimes they did not commit. Understandably, the lives of such people often become dominated by a passion to clear their names because it is next to impossible for them psychologically to leave the hurtful injustice behind and return to the quiet enjoyment of their lives. What they are looking for is justice by way of clearing their names and, possibly, being compensated for the wrongly imposed hardship on them. Since they can find no peace of mind without this, their stake in finding justice is obviously high.

An equally valid and important, but seemingly far less recognised, need to find justice exists in victims of crime. Especially in cases of serious crime, they also become preoccupied with the hurtful past and are unable to get on with their lives because they

¹Kant 1965, p. 100.

²Foucault 1980, p. 1.

³Kennedy 1989, p. 321.

can find no peace of mind until they are satisfied that justice is done. What they are looking for might consist in fair restitution, compensation and the just retributive punishment of the persons responsible for the wrongs committed against them. As with innocent people wrongly accused of crime, victims have a very high stake in finding justice in the sense just indicated.

Justice is also important to actual offenders and not only in that they should not be punished excessively and out of proportion to the moral gravity of their crime, but also in terms of being afforded constructive ways to make amends for their mistakes so that they can clear their conscience and make a fresh start.

Yet, in their purported mission to protect the innocent and punish the guilty in accordance with their moral desert, contemporary legal justice systems marginalise and disempower these three groups of people who, in a sense, are their clients. Most significantly, victims and accused are denied the opportunity to take an active role personally in the legal processing and final resolution of their cases. In effect, they are reduced to the status of idle bystanders in their own cases in what, after all, is *their* conflict. From a client focus point of view, this is an appalling situation. The judicial system seems to have been monopolised by legal professionals who make a handsome living out of creating drama, and much squabbling among themselves over how to resolve other people's conflict. Under the adversary system, for example, prosecution and defence are bent on vanquishing each other by fair means or foul. Unlike the case of the innocent clients of the legal system, what seem to be of overriding importance to the professionals are not the truth of the matter and justice, but whether they can triumph over each other by winning the case.

Through all this high drama of the courtroom, judges seem to be content to act as if they were mere referees in a boxing match or a cock fight, and justice in which victims and accused have such an important stake, is unnecessarily and recklessly jeopardised. It is part and parcel of this unsatisfactory situation that many professedly well-meaning

police officers are of the view that “the judicial system is a joke,” and they “would perjure themselves or do whatever else was necessary to secure a conviction.”⁴

Moreover, in the light of the mass violence created by the Rodney King trial in Los Angeles, for example, we cannot dismiss Braithwaite’s suggestion that the American criminal justice system might be

... more a cause of crime than a protection against it ... It is time to recognise that it, like other Western criminal justice systems, is an abject failure. In fact, the criminal justice system stands out as the greatest failure of any of America’s institutions.⁵

Kennedy, who recommends the adoption of the inquisitorial system of France and other European countries put the matter in the following way.

... the adversary system of criminal justice which we employ in this country [Britain] and which we have exported to the United States and Commonwealth and is essentially one of conflict, is not only extremely childish, but a most unsatisfactory way of attempting to dispense justice. In a situation where one side is doing its best to vanquish the other, truth is apt to fall by the wayside.⁶

Similarly, in his book *Trial by Voodoo*, Evan Whitton argues that under the adversary system the law defeats justice and democracy because it does not set out to establish the truth, and punish the guilty.⁷ The long list of historical examples cited by Whitton and Kennedy - all of them cases where the guilty have gone free while many innocent people have been convicted, due in the main to the adversarial nature of the law - leaves little doubt that reform of this particular aspect of judicial systems is long overdue.

However, without intending to question the virtues of the European model in other respects, traditional inquisitorial systems are still not satisfactory enough in terms of

⁴Judicial System a Joke: Detective. *The Canberra Times*. Wednesday, July 5, 1995, p. 5.

⁵Braithwaite 1993, pp. 33 & 34.

⁶Kennedy 1989, p. 322.

⁷Whitton 1994.

providing for substantial, grass-roots participation by victims and offenders in the legal processing and resolution of their cases. As Braithwaite points out,

The Western criminal arrest and trial is a sterile production line process dominated by experts (mainly lawyers) who disempower the communities that might be able to plan some solutions to the underlying problems.⁸

Fortunately, this is increasingly recognised all over the world and laws are being modified, even in Europe.⁹ This is as it should be. For, given the high stakes held by clients of the justice system, justice is far too important to be left completely in the hands of legal professionals.

While the need for legal justice systems to be more client focused is the overall theme of this thesis, for the most part the emphasis will be on the importance of victim involvement and empowerment from a victim justice perspective, and on the relevance of revenge to the quest for justice in legal institutionalised contexts.

⁸Braithwaite 1993, p. 37.

⁹Severin 1995.

Introduction

No more tears now; I will think upon revenge.

Mary Stuart: Queen of Scots¹⁰

Revenge is a kind of wild justice, which the more man's nature runs to, the more ought law to weed it out.

Francis Bacon¹¹

Revenge is a universal phenomenon.¹² The need and the impulse to avenge personal injustices and wrongs remains a persistent feature of the individual human psyche. In spite of the law and public opinion in the Western world still being against it, in a climate of deepening dissatisfaction with the law's ability to protect the innocent and punish the guilty, revenge is poised to make a comeback onto the moral stage. This is confirmed by increasingly sympathetic portrayals of it in fiction and by real life situations where juries and judges are increasingly lenient on people who take it upon themselves to remedy the law's reluctance to punish their wrongdoers.¹³ The depth with which wronged individuals can feel about injustices committed against them, the intensity of their need to find satisfaction through redress, and the relevance and importance of revenge to the quest for justice, have been recognised by Bertrand Russell.

Suppose your daughter has been raped and then killed ... Ought you, in these circumstances, to preserve a philosophic calm? ... The primitive reaction to such disasters is revenge. When Macduff learns that his wife and children have been killed by Macbeth, he resolves to kill the

¹⁰On hearing of Riccio's death, 9 March 1566.

¹¹Bacon 1985, p. 72.

¹²Elster 1990, p. 862.

¹³Wallace 1995, p. 363.

tyrant himself. This reaction is still admired by most people, when the injury is great, and such as to arouse moral horror in disinterested people. Nor can it be wholly condemned, for it is one of the forces generating punishment, and punishment is sometimes necessary. Moreover, from the point of view of mental health, the impulse to revenge is likely to be so strong that, if it is allowed no outlet, a man's whole outlook on life may become distorted and more or less insane.¹⁴

Yet, many contemporary criminal justice systems not only ignore but explicitly dismiss such human needs through the marginalisation, silencing and disempowerment of victims in the course of processing and resolving their cases. While such legal systems can re-establish immediate legal order, they have proven unsatisfactory in terms of society's long-term interests in controlling and reducing crime, let alone in terms of being able to bring peace and satisfaction following a crime to those most affected by the whole situation, the victims and the offenders whom it marginalises and disempowers. In his report to the Queensland Corrective Services Commission (Australia) Severin stated the problem in the following manner.

Unfortunately, the criminal justice system does not provide an active role for victims and offenders. In the criminal justice system, victims and offenders are required to enact their roles as witnesses and defendants, but beyond that, they have to remain passive. Once they have stated their immediate cases, the system 'steals' their conflict and takes over. The system is very much concerned with re-establishing legal order, but fails to leave any room for the interpersonal resolution of conflict and the restoration of social peace.¹⁵

If peace is to be restored between offenders and victims (not excluding others who may also be affected because of their close relationship to them), and if a feeling of satisfaction is to be engendered in them with regard to the legal processing and final outcome of their conflict - both of which are important for their respective social-psychological restoration and for enabling them to get on with their lives, - then they must be empowered and given constructive opportunities to play an active role in working out a resolution to their conflict which is acceptable to both of them. This is the

¹⁴Russell 1961, p. 561.

¹⁵Severin 1995, p. 17.

restorative justice model for dealing constructively with criminal offences and Severin explains the idea as follows.

Restorative Justice goes beyond restitution and connotes a dynamic dimension and an interactive process of establishing justice and fairness. With its focus on conflict resolution and the re-establishment of peace (justice and fairness), restorative justice is essentially based on the voluntary and participatory nature of the conflict-resolving procedure.¹⁶

The early success and popularity of programs based on the concept of restorative justice in countries such as Australia, New Zealand, the United States, Austria and Germany where over the past decade the concept has developed at “breakneck speed,”¹⁷ are signs that the concept of restorative justice is destined to play a very important, if not central, role in the future development of legal justice systems. Although restorative justice is a winning principle for all parties involved, the current dissertation explores its importance and potential mainly from a victim justice point of view. Of special relevance in this regard is the almost magical ability of restorative justice models to accommodate constructively and unobtrusively the need, indicated above, which wronged parties have for retributive justice and revenge. Such victim needs may also be satisfied and accommodated by a similar involvement and empowerment of victims in modified versions of standard judicial systems. But whatever form victim involvement and empowerment might take, insofar as the model in question allows them the real possibility of some form of revenge, the defensibility of that model requires that the particular forms of revenge involved themselves be shown to be justifiable. Since both the restorative justice model and the modified judicial model for victim participation and empowerment allow for institutionalised forms of revenge, the current dissertation is focused on placing in proper perspective through conceptual analysis, and then defending, such forms of revenge.

¹⁶Severin 1995, p. 17.

¹⁷Severin 1995, pp. 16 - 17.

The first form of justice, throughout the world, has been revenge.¹⁸ By contrast, in modern industrialised Western societies the quest for justice has become confined to the domain of highly impersonal legal systems and the question arises whether in legal institutionalised contexts revenge remains at all relevant to that quest. This question, in turn, raises a more basic one: What is revenge? Saying, as Bacon does, that revenge is a kind of wild justice might capture some of its flavour, but it is no substitute for a proper analysis which the concept requires for an informed evaluation of its normative status with regard to rationality and morality. Such an analysis and evaluation are what I propose to undertake. I want to show that, instead of trying to eradicate revenge altogether, in the interests of victim justice the law should only seek to tame revenge, to weed the wildness out of it.

The earlier two aphorisms capture well the divergent attitudes that the idea of revenge is prone to invoke. They also represent an interesting polarisation of attitudes and two very different paradigms (or sets of paradigms) of justice. In the former paradigm revenge is held to be the ideal response to wrongdoing. In many cultures taking revenge is regarded as a genuine and legitimate moral option which is often open to victims of injustice. Indeed, there are many situations in which, even if it is not the only option, in view of the society in question it is *the only morally acceptable option* which is open to such persons. For example, on the basis of his survey of the various revenge practices of pre-industrial societies from around the world, Herbert Spencer points out that “In all these cases we see that either avowedly or tacitly revenge is considered a moral obligation.”¹⁹ Similarly, in his anthropological study of blood revenge in Montenegro, Christopher Boehm observes that:

After a first killing had occurred, the retaliatory homicide that followed was considered not only to be reasonable and proper but also to be *morally necessary* by traditional Montenegrin standards.²⁰

¹⁸See, for example, Spencer 1966, p. 362.

¹⁹Spencer 1966, p. 363.

²⁰Boehm 1984, p. 89.

By contrast, and in spite of its ubiquity on the individual level, in many societies there would, if anything, seem to be a norm against seeking revenge and, correlatively, injunctions to turn the other cheek.²¹

This is largely true of many industrialised Western societies in which strong anti-revenge attitudes have been developed under the combined influence of Christianity and policies pursued by powerful centralised States and their bureaucracies on crucial matters of law and order. The fact that most contemporary philosophers and thinkers would readily endorse Bacon's advice against revenge, and that they invariably do this by assumption rather than argument, are indicative of a pervasive and well established paradigm of justice which is uncompromising in its rejection of revenge. It is not uncommon, for example, to see polemical accusations that retribution is nothing more than revenge in disguise with the implicit assumption that retributivism is automatically knocked out as soon as it is found to be tainted by an element of revenge. Observations made by Nozick and Glover are illustrative.

The view that people deserve punishment for their wrongful acts ..., independently of the deterrent effect of such punishment, strikes some people as a primitive view, expressive only of the thirst for revenge.²²

Among objections urged against retributivism are that 'retribution' is a polite name for revenge, which is more generally recognized to be evil, ... hatred and pleasure ... combine so unpleasantly in revenge.²³

Similarly, C. S. Lewis warns that "the least indulgence of the passion for revenge is very deadly sin."²⁴ According to Lewis, revenge is a "vindictive passion ... [which] is evil and expressly forbidden to Christians."²⁵ In contrast to Lewis, Nietzsche rejects Christianity and accuses it of slave morality which cripples, rather than enhances,

²¹Elster 1990, p. 883.

²²Nozick 1981, p. 366.

²³Glover 1970, p. 145.

²⁴Lewis 1957, p. 109.

²⁵ Lewis 1957, p. 82.

people's character.²⁶ Nevertheless, he concurs with Lewis and Glover in rejecting revenge.

... for humankind to be redeemed from revenge - that is for me the bridge to the highest hope, and a rainbow after long storms.²⁷

I, for one, do not share these extreme attitudes to revenge and regard both paradigms of justice mentioned above to be unsatisfactory. The first one is lop-sided because its concern for justice is focused predominantly on the wronged party's need for retribution. The second one is similarly lop-sided because it ignores the needs of victims for restoration following an offence committed against them and focuses mainly on the offender and on society's need to deter crime. Moreover, while an idealisation of revenge and its public encouragement through social expectations and norms are likely to generate unnecessary dangers, hardships and disharmony for both individual interests and the public good, it is a mistake to consider revenge to be incompatible with morality and justice. As Wallace points out, revenge can be measured and proportionate and "there is nothing in the concept of revenge that necessitates injustice to the revengee."²⁸ Accordingly, I shall argue that revenge can be a way to ensure justice and that making lawful allowances for the legitimate personal needs of victims of wrongdoing and crime in terms of retributive justice leads to superior legal systems which make provisions for institutionalised forms of revenge.

²⁶Nietzsche 1968.

²⁷Nietzsche 1993, p. 122. (*Thus Spake Zarathustra*, Chapter 29: "The Tarantulas.") It must be noted that Nietzsche mostly uses the word revenge to indicate a psychological state of resentfulness which is born of envy and jealousy but which takes on the cloak of goodness through self-deception, rather than to mean acts of revenge as we commonly do. He talks with disdain only about people who self-deceptively regard themselves to be good on account of not retaliating when maltreated by others, but who, at the same time, deceive themselves about the impotent resentment they feel and are too cowardly to express openly. They decide instead to hide it even from themselves by attributing their lack of action to some imagined virtue, such as humility or long-suffering. By contrast, in some places he talks disparagingly of the doglike people who allow themselves to be maltreated and who are too afraid to revenge themselves. His examples include the person "who does not outrage, who harms nobody, who does not attack, who does not requite, who leaves revenge to God." (Nietzsche 1989, p. 46.) In such passages, where he talks of individual acts of revenge, rather than the destructive and contorted psychological state of *ressentiment* and the actions which spring from it, his attitude to revenge seems to be more accepting, if not approving.

²⁸Wallace 1995, p. 369.

One of the important tasks of this thesis, therefore, is to show that, contra Bacon, revenge need not be wild and dangerous, for there are perfectly acceptable ways to tame and civilise it through appropriate models of institutionalisation. Of equal importance is to demonstrate the superiority of resultant systems of justice over standard judicial systems. This superiority consists, as I shall argue, not only in their outstanding ability to ensure justice to victims but also in their much improved ability to control crime by assisting, rather than hindering, the correction and social reintegration of offenders. In the course of developing and defending these ideas I shall seek to retain what is worth preserving from both the above paradigms of justice. At the same time, I shall seek to lay the philosophical foundations of a new paradigm of justice which encompasses a new way of thinking about revenge, and which escapes the excesses of both extremes.

Finding a constructive way of breaking free of the clutches of an established paradigm is rarely simple or easy. The task at hand is no exception. Plausible and stable reference points must be found for constructing a new frame of reference in the light of which a new understanding of the nature, rationality and morality of revenge can be developed and, correlatively, more accepting attitudes to revenge can be seen to be sensible and desirable. A defence of moral retributivism by reference to moral responsibility for wrongdoing and the drawing up of a map of the logical space of punishment which helps to define revenge and clarify its place in relation to other forms and concepts of punishment will serve as important reference points in these regards. A third essential reference point is provided by victim justice which, in all fairness, should be at least as important to the legal justice system as doing justice to offenders. I shall argue that, in turn, doing justice to victims requires a legal justice system which, instead of silencing, marginalising and disempowering victims, acknowledges and respects their legitimate need for retributive justice, empowers them to speak up for themselves and encourages them to participate in the legal resolution of the offences committed against them. Giving such considerations to victims, I shall argue, is necessary for their social-psychological restoration and is in part what it means to treat them with justice.

The relevance of revenge to a new victim justice focus in legal justice settings has also been recognised by Robert Solomon.

To the dangers of vengeance unlimited it must be countered that if punishment no longer satisfies vengeance, if it ignores not only the rights but the emotional needs of the victims of crime, then punishment no longer serves its primary purpose, even if it were to succeed in rehabilitating the criminal and deterring other crime (which it evidently, in general, does not). The restriction of vengeance by law is entirely understandable, but, again, the wholesale denial of vengeance as a legitimate motive may be a psychological disaster.²⁹

It is important to stress that, just like Solomon, I am not advocating the unconditional, inherent moral desirability of revenge in legal institutionalised contexts. Rather, I am arguing for the moral *permissibility* of institutionalised revenge with a view to promoting victim justice. This does not imply that we should actively encourage people to be retributive or revengeful within limits allowed by law. What it allows for, however, are forums where victims can express their legitimate feelings of resentment and outrage, forums where they can demand satisfaction in terms of adequate material restoration, apology and/or retributive justice in a controlled and civilised manner for the wrongs committed against them without being moralised or patronised and without feeling that they are thought less of by others for doing so. An important dimension of such victim empowerment consists in according victims their rightful status to be the primary bearers of the prerogative and power to forgive and be merciful. It stands to reason that the forgiveness of wrongs should lie primarily with those against whom those wrongs have been committed and their decisions on such matters should be subject to veto only when they show unreasonable leniency or harshness. Victim empowerment in the above terms is important, I shall argue, for purposes of victim restoration and is an integral part of providing due justice to them.

²⁹Solomon 1990, p. 294.

1

The anti-revenge paradigm

1.1 Introduction

I mentioned in the main introduction that in many contemporary cultures, such as industrialised Western societies, revenge is held in very low esteem and that such cultures are often characterised by norms prohibiting revenge and by corresponding injunctions to forget, forgive, or turn the other cheek. I also suggested that such strong anti-revenge attitudes are part of a paradigm at the core of which there seem to lie certain religious, philosophical and culturally embedded beliefs that exaggerate the dangers and the retributive element in revenge. I argue in this first chapter that there are no philosophically defensible foundations to these beliefs about revenge.

I proceed as follows. In the next section I question views of revenge which regard it to be a perversity and an evil vindictive passion. Following this, in section 1.3 I challenge suggestions that revenge is characterised by a morally repugnant emotional tone: pleasure in the suffering of another. Finally, in the last main section of this chapter I explore the question whether revenge can be properly regarded as a form of punishment. I argue in support of an affirmative answer to this question, but point out that, ultimately, this question is important only if the word 'punishment' is unduly ennobled by an assumption that whatever qualifies as punishment is thereby automatically justified.

1.2 Perversity, evil, vindictiveness and revenge

C. S. Lewis warns that “the least indulgence of the passion for revenge is very deadly sin.”³⁰ According to him, revenge is a “vindictive passion ... [which], of course, is evil and expressly forbidden to Christians.”³¹ Moreover, according to Lewis, revenge is a “perversion”:

The good thing of which vindictive passion is the perversion comes out with startling clarity in Hobbes’s definition of Revengefulness; “desire by doing hurt to another to make him condemn some fact of his own.”³² Revenge loses sight of the end in the means, but its end is not wholly bad - it wants the evil of the bad man to be to him what it is to everyone else.³³

These are somewhat confusing passages but it is worth trying to make Lewis’ message less obscure. He makes at least three allegations regarding the nature of revenge:

- (a) Revenge has a good end but it loses sight of it
- (b) Revenge perverts, or is a perversion of, this good end.
- (c) Revenge, by its very nature, is a vindictive passion and is evil.

Unfortunately it is not clear how Lewis conceives the relation between these claims. Therefore, insofar as it is possible, it will be best to consider each on its merits. Before evaluating these claims, however, some clarifying remarks regarding the Hobbes/Lewis definition of revenge are necessary. Supposedly, the purpose or the good end of revenge is some sort of future state of affairs in which the wrongdoer comes to recognise the wrongness of his own actions. The means to this good end is the infliction of suffering (in whatever form) on the wrongdoer. But so understood, the Hobbes/Lewis definition of revenge can be given at least two readings: one of these is consequentialist, the other is not. I shall discuss the plausibility of the first two claims (a

³⁰Lewis 1957, p. 109.

³¹ Lewis 1957, p. 82.

³²The quote is from *Leviathan*, Pt. I, Chap. 6.

³³Lewis 1957, p. 82.

& b) by concentrating first on the consequentialist reading and then, very briefly, I shall consider these two claims under the non-consequentialist reading before taking on Lewis's third claim (c) that revenge is evil.

1.2.1 Revenge as losing sight of its good end

There are a number of problems with the claim that revenge loses sight of its own end. One of them is that, while Lewis says that revenge has a worthwhile goal (the goal of making the evil man condemn his own evil), he fails completely even to indicate *how* revenge loses sight of it. It is not hard, of course, to imagine cases where one loses sight of one's proposed goal and there certainly is no reason to believe that this cannot happen to the avenger. Therefore, if it is to make a point of substance and importance, Lewis's charge must be more than this. He must be claiming that losing sight of the goal in the means is a defining, necessary feature of revenge. But this claim is mistaken. The problem is due to the assumption - which in the definition is turned into a stipulation - that revenge has (to have) an end or a goal. As I shall show, this assumption is false and consequently the definition these authors offer of revenge becomes implausible and is unable to serve a useful purpose on any reading of it - be it consequentialist or not.

Under the *consequentialist* reading of the definition which is the one I think follows most naturally from the quoted passage, revenge is understood to be necessarily *forward-looking* in character. It has a forward-looking goal by definition, the goal of making the evildoer recognise his own evil. Intuitively, such a definition of revenge is completely out of place and this is reflected by the fact that many writers mentioning revenge seem to be reluctant to grant even the possibility of forward-looking reasons or goals to it. In fact, it is precisely for this reason that it is regarded by some as irrational. It is because revenge is commonly assumed to be *entirely* backward-looking by its very nature and that it is not there to benefit anyone but only to inflict more harm and suffering, that more often than not it is perceived as pointless aggression.

This is an extreme position to take and, no doubt, a mistaken one because even if one accepts a purely “ends-means” conception of rationality, the fact remains that revenge can be taken, at least in part, for consequentialist considerations. Notwithstanding, it is quite surprising to see that by radical contrast, Lewis and Hobbes swing to the other extreme and regard revenge as *necessarily* being characterised by forward-looking considerations. But, surprises or no surprises, the crux of the matter is that having a goal in mind is not at all a prerequisite for revenge. Contra Lewis and Hobbes, goal-orientedness, is not a defining feature of revenge. For, revenge can often be totally backward-looking, done only for reasons that lie entirely in the past. For example, to the question: “Why did you punch Fred in the nose?” one might just get the reply: “Because he insulted me in public and anyone who insults others in public to humiliate them, deserves such treatment.” Or take the real life example of a Sydney rape victim who “stabbed his attacker to death seven years later.”³⁴ His explanation for the murder was that he was raped by this person. Nothing more, nothing less.

In spite of there being no mention of any goals for the acts in question, these explanations are readily understood and seem complete, if not satisfactory, from a retributive standpoint. The only appeal they make are to general principles of desert. There isn’t even mention of abstract goals, such as honour-saving and definitely no mention of the wrongdoer’s reformation or the deterrence of others. In view of this consideration, Lewis’s first claim that revenge loses sight of its good end can be clearly seen as downright irrelevant to the matter at hand.

1.2.2 Revenge as a perversion of its good end

Now, if Lewis’s crucial assumption is wrong and goal-orientedness is not a defining feature of revenge, then his second claim that revenge is a perversion of its own end cannot be right either. It may, perhaps, be true in those cases where revenge is taken

³⁴Man Stabbed Over Rape, Court Told. *The Canberra Times*. Wednesday, May 29, 1991, p. 4.

with this particular aim in mind - that is, with the purpose of making that person "condemn some fact of his own." However, only "perhaps", for even if there are cases where revenge is taken with such, or similar, goals in mind, it is not at all clear how and whether revenge perverts these goals. For even if we grant that an avenger who initially intends to inflict suffering on the other in order to get him to recognise the wrongness of his own action may, in some cases, lose sight of this respectable end and come to treat instead the suffering of the other as an end in itself, it is still not clear why losing sight of one's goal in this way should amount to *perverting* that goal. This question, of course, may be impossible to settle conclusively without a proper analysis of the nature of perversion - which, though interesting, is otiose to the main criticism I have advanced above against these two claims by Lewis (*viz.*, the charge of irrelevance). For, to repeat, if having a goal is not a defining feature of revenge, claims that revenge by its very nature loses sight of, or perverts, its own goal are off the point altogether.³⁵

But this is not all. In addition, under the assumption that the former (a) is the explanation for the latter (b), both these claims are vitiated by a conceptual problem as well. For, if revenge loses sight of its own goal by definition, in the way Lewis seems to suggest, it seems incoherent to insist further that revenge has that goal at all. This seems to remain true even if his claim is of a normative flavour, that revenge ought to be goal-oriented in such and such a way but it fails to keep it in view somehow. Going normative certainly seems to offer no way out of the paradox that seems to threaten the very claim that there can be certain things, like revenge is claimed to be, which by their very nature lose sight of their own ends. For, to repeat, if by their very nature they lose sight of their ends, is it not absurd to impute to them these ends in the first place?

³⁵Bob Goodin suggested to me that Lewis's point is unsatisfactorily formulated. According to Goodin, "what is perverted is not the goal but the practice (*viz.*, inflicting suffering on evildoers)." But, beside not being true to what Lewis himself says quite clearly, it is hard to see how this particular formulation is going to stand a better chance of showing that punishing evildoers through revenge is a perversion.

I mentioned earlier that the Lewis/Hobbes definition of revenge has both a consequentialist and a non-consequentialist reading. The foregoing criticisms of the first two claims, that is, (a) and (b), were made in the context of a consequentialist reading, according to which revenge was understood to be necessarily having a *forward-looking* goal. I shall, therefore, discuss here very briefly what difference it would make to these two claims if, instead, they were evaluated on the basis of a non-consequentialist reading. For, with some ingenuity, this forward-looking feature in the definition can be eliminated by incorporating the goal in question to form a proper *part* of the very concept of revenge.³⁶ The result of such a definitional tinkering is that rather than revenge being performed for some particular purpose in mind (viz., the goal of making the wrongdoer recognise his own evil), this goal is regarded to be a constitutive part of revenge - so that revenge cannot be properly said to be complete (or ever taken) until, as a consequence of the inflicted suffering in retaliation for the harm or wrong, the wrongdoer recognises his own evil - in Hobbes words: until he *denies some fact of his own*.

Though such a conceptual exercise is interesting in itself, at the end of the day it makes no difference to what has been urged against Lewis' two contentions so far. The 'lack of explanation' objection certainly remains and the main charges, those of irrelevance and incoherence, can be repeated just as effectively, *mutatis mutandis*, against both the above claims (a & b) on this non-consequentialist reading as well. The problems of irrelevance and incoherence arise for both claims from the fact that the said goal is at all incorporated in the definition of revenge and not from imputing to revenge by means of it a *forward-looking* character. To see this more clearly, the irrelevance problem remains because a non-consequentialist reading will not make the definition any more realistic. That definition is incapable of capturing the majority of what we ordinarily regard as being unambiguous instances of revenge. Clear cases of revenge where suffering is successfully inflicted on a wrongdoer but where the wrongdoer, for

³⁶My attention was drawn to this interpretation by Michael Tooley.

some reason or another, fails to recognise his own evil are excluded definitionally from being cases of revenge. Such an unrealistic definition is useless for any practical or philosophical purpose. It is certainly useless for analysing and evaluating that controversial class of human actions which we ordinarily call and regard to be proper cases of revenge.

The incoherence problem doesn't disappear on this reading of the definition either. It is an outright contradiction to say that the concept of revenge is composed of a sub-action (viz., the infliction of suffering on the evildoer) and the proper goal of this sub-action (viz., the evildoer's education) where the former somehow necessarily has to lose sight of the latter. For, if the sub-action loses sight of its goal by necessity, and by definition, in the way Lewis would have to be taken to suggest, it seems incoherent to further insist that the sub-action in question has that goal at all. If by nature it loses sight of its own end, is it not absurd to impute to it this end in the first place? Thus, regardless of the reading, the paradox remains inextricably tied to Lewis's conception of revenge.

1.2.3 Revenge as an evil vindictive passion

Lewis points out that there is a truth behind "the universal human feeling that bad men ought to suffer" and argues that "it is no use turning up our noses at this feeling, as if it were wholly base"³⁷

[For] On its mildest level it appeals to everyone's sense of justice. ... On a sterner level the same idea appears as "retributive punishment" or "giving a man what he deserves". Some enlightened people would like to banish all conceptions of retribution or desert from their theory of punishment and place its value wholly in the deterrence of others or the reform of the criminal himself. They do not see that by so doing they render all punishment unjust. What can be more immoral than to inflict suffering on me for the sake of deterring others if I do not *deserve* it? And if I do deserve it, you are admitting the claims of "retribution". And what can be more outrageous than to catch me and submit me to a disagreeable process of moral improvement without my consent, unless (once more) I *deserve* it?³⁸

³⁷Lewis 1957, p. 81.

³⁸Lewis 1957, pp. 81 - 82.

Lewis's defence of the retributive theory of punishment is basically correct, and I shall argue for the philosophical defensibility of the retributivist position in the last chapter. However, I disagree with Lewis's attempt to categorise revenge as something distinct from, and opposed to, retributive punishment:

On yet a third level we get vindictive passion - the thirst for revenge. This, of course, is evil and expressly forbidden to Christians.³⁹

Obviously, for Lewis, revenge is intrinsically evil, something barbaric by nature, something which automatically and necessarily destroys the morality of anything that partakes of it. But what defensible reasons could there be for holding such a grim view or revenge, especially if we accept a retributive theory of punishment as Lewis does? To my mind at least, there are none. For there is no doubt that the ideas of retribution and revenge are very closely allied especially in this respect: *that some unpleasant treatment is meted out to someone as desert for what they have done* and regardless of whether or not that treatment will have other desirable consequences (such as reformation of the offender and deterrence). It is well to remember in this regard that Lewis's earlier attempt to impute goal-orientedness to revenge as a defining feature fails precisely for this reason.

Retribution and revenge just aren't different enough in this crucial respect. As I shall argue in more detail in the following two chapters, the only difference between them is that revenge is necessarily personal whereas other forms of retribution, such as judicial and vigilante retribution may not be. All indications are that Lewis makes the mistaken assumption that retribution is definitionally guaranteed to be deserved. This is to honorify, embellish and ennoble retribution by assuming or suggesting that all instances of it are invariably and unfailingly justified.⁴⁰ As I shall argue in more detail in the following chapter, the mistake here lies in supposing that the crucial moral elements of desert and the principles of justice, such as proportionality and not punishing the

³⁹Lewis 1957, p. 82.

⁴⁰Hedenius 1973, p. 13. The concept of honorification is going to be discussed in more detail with regard to punishment further below.

innocent, are built into the very idea of retribution. But even if this were correct and there were a difference between retribution and revenge so that these sanitising elements had to be added onto revenge in order to make it fair, just and morally acceptable, this would warrant only caution in the case of revenge, not the vilifying, grim view Lewis takes of it. In the absence of supporting reasons, Lewis's claims that revenge is a vindictive passion and that it is evil remain unsubstantiated assertions against which we can put Wallace's argument against Schopenhauer's similar assertion that revenge is evidence of our own wickedness. According to Wallace, moderate and proportionate revenge need involve no such thing.

Consider the much put-upon domestic servant. For many years she has been in the service of a domineering and petty employer. On her final working day, she prepares madam's soup, pees in it and serves it for lunch. She feels suitably triumphant. Does revenge require that her martinet of an employer is harmed by the soup? It is unlikely that she will be. ... Schopenhauer is obviously assuming that revenge must involve wickedness, but our servant example show this to be too strong.⁴¹

A similar and equally unsubstantiated assertion to that made by Lewis, is Glover's claim that "hatred and pleasure ... combine ... unpleasantly in revenge."⁴² Glover makes no attempt to justify this claim. Instead, he seems to assume that the claim is self-evident and uncontroversial. But such assumptions are unwarranted. I shall argue in the last main section of this chapter that the typical emotion behind revenge is not hatred but the retributive emotions of resentment, indignation and anger which are occasioned by perceptions of injustice and unfair treatment. Only when a seriously offended person is powerless to remedy the injustice in question are such retributive emotions likely to turn into hatred and bitterness toward those responsible for the unfair treatment.

But even if it were the case that some kind of hatred is typically involved in resentment and revenge, it is questionable whether there is anything wrong with hating an unjust act, or even an unjust person who on top of the wrongdoing refuses to make

⁴¹Wallace 1995, pp. 369 - 70.

⁴²Glover 1970, p. 145.

amends and remains defiant and unapologetic. The other element of Glover's implied allegation, that revenge is characterised by some morally objectionable pleasure in the suffering of another, is considered in conjunction with Nozick's similar claim to this effect in the following main section where I argue that the pleasure in question is most plausibly identified as satisfaction in justice having been done.⁴³

In this sub-section, I have argued against claims by Lewis that revenge perverts and loses sight of its own end, and that it is an evil vindictive passion. I have raised several objections against each of these claims. I argued against the first two claims that, for conceptual reasons and in the interests of relevance and plausibility, goal-orientedness (trying to make an evil man condemn his own evil) cannot be imputed to revenge as a defining feature in the way suggested by the Lewis/Hobbes definition of revengefulness on which those two criticisms are based. This holds true, I argued, under both consequentialist and non-consequentialist readings of that definition. I also pointed out that Lewis fails to clarify how revenge loses sight of the end in the means and that he also fails to explain the manner in which revenge is supposed to pervert its otherwise good end. Furthermore, in response to Lewis's claims that revenge is a vindictive passion and that it is evil, I argued that revenge is not sufficiently different from retributive punishment, of which Lewis is a strong advocate, to merit Lewis's low opinion and vilification. Finally, I raised doubts regarding the accuracy of Glover's unexplained and unsubstantiated assertion that revenge involves an unpleasant mixture of hatred and pleasure. I suggested that hatred and bitterness are likely to develop only when people find themselves too disempowered to address serious injustices committed against them. Besides questioning whether there would be anything wrong with hating unjust acts, or even unjust persons who refused to make amends and remained defiant following their

⁴³With some relevance to the foregoing discussions, Wallace distinguishes between vindictive and non-vindictive revenge. The latter, according to him, consists merely in finding satisfaction in evening the scores with someone, such as when one wins a squash game after losing one. While I find it implausible to regard such competitively, as opposed to morally, grounded acts of payback to be revenge, except in a metaphorical or derived sense, I merely wish to point out that Wallace uses 'vindictive' and 'malicious' to mean about the same thing as 'hurtful' or 'harmful,' and that he accepts that such cases of revenge may, on occasions, be justifiable. (Wallace 1995, pp. 372 - 73.)

wrongdoing, I made a promisory note to explore in detail the nature and role of emotions which typically characterise revenge. These are the tasks to which I now turn.

1.3 The emotional tone of revenge

Similarly to Glover, Robert Nozick claims that “revenge involves a particular emotional tone, pleasure in the suffering of another.”⁴⁴ Nozick makes this claim, so to say, only in passing in the context of a comparison between revenge and retributive justice. He does not say explicitly that taking pleasure in the suffering of another is morally condemnable. However, if his exercise of attributing this feature to revenge and denying it from retribution is to be of any significance, he must at least think that the feature in question might be a morally problematic one - especially given that he makes the comparison as part of the groundwork for an account and eventual defence of retribution.

The notion of taking pleasure in the suffering of another is an ambiguous one. It can be understood to refer to at least two things and my main project in this section is to demonstrate that when it refers to the morally problematic one, it is highly unlikely to characterise revenge and when it does characterise revenge, it is highly unlikely to be morally problematic. However, before discussing this in more detail I want to draw attention to the fact that whatever the exact nature of the emotional tone in question may be, contrary to suggestions by Nozick and Glover, taking pleasure in the suffering of another is not a defining feature of revenge. It is not difficult to find either imaginary or real cases of revenge where this is evidently not the case. For example, one may feel quite awful about taking revenge on a previous friend but do it nevertheless under pressure from the family (for the betrayal or rape of one’s sister, or the wrongful death of one’s father, for instance). In such a case, rather than taking pleasure, one may resent being forced to cause the suffering of the other. But even in a more general way one may

⁴⁴Nozick 1981, p. 367.

have an aversion to seeing and causing the suffering of any living being and it is easy to think of such people taking revenge, again, under pressure from others, perhaps, while not enjoying any, and in fact detesting every bit of it. Once again, and this is the crucial point, taking pleasure in the suffering of another cannot be a *defining* feature of revenge.⁴⁵

This is not to deny that it may be an essential feature of revenge that the other has been harmed. Just like any other form of punishment, revenge may well require some sort of harming of the subject by definition. In ideologies that recommend revenge (say, as the best reply to wrongs, or as a duty) there can indeed be a form of satisfaction - if not exactly pleasure - which can be said to be properly felt when revenge has been successfully accomplished. However, this is something entirely different from taking pleasure in the suffering of another. Rather, it is the taking of a positive attitude of some sort toward something that necessarily entails suffering for the other. So even if we acknowledged that, as a contingent matter of fact, revenge may often be accompanied by a sense of satisfaction which is comparable to some sort of pleasure, there is far more work required than Nozick and Glover seem to realise to make their claims start having some grip.

But even if it were the case that most instances of revenge are accompanied by some sort of pleasure in the suffering of the other - though, for reasons already indicated and for others to follow, it is highly unlikely that this in fact is the case - we ought to be wary of forming hasty judgements as to the moral propriety of such pleasures and the propriety of acts of revenge they happen to accompany. Before forming any such judgements we must first of all enquire as to the exact nature of these pleasures in order to find out which forms of revenge are acceptable and which are not. Then, if any forms of it are found to be contrary to the requirements of morality, we must look at the relevant empirical information in order to decide whether such unacceptable forms of revenge

⁴⁵The possibility that it might be a *common* feature of revenge is considered further below.

taking are at all common or not. As I already indicated, the phrase 'pleasure in the suffering of another' is an ambiguous one. It can either mean (a) pleasure in the suffering of another *simpliciter*, or (b) pleasure in the suffering of another who (is believed to have) harmed one. Although, strictly speaking, Nozick and Glover are committed by their claim only to the second reading, it is quite possible that in making their claims they have in mind the first one. However that may be, it will be worthwhile to discuss the status of revenge with respect to both types of pleasure. This is what I do in the following two sub-sections.

1.3.1 Revenge and the taking of pleasure in the suffering of another *simpliciter*

To identify the emotional tone in question simply as pleasure in the suffering of another *simpliciter* would be to present revenge (intentionally or not) in an artificially bad light. For, to be sure, this sort of pleasure is perverse and morally base. It is the mentality of the torturer who derives pleasure from the suffering of his victims. And to suggest, as the claims made by Nozick and Glover may be taken to suggest, that revenge is even typically accompanied by an emotional tone of this repulsive sort is an idea I definitely want to resist.

If in many, or perhaps in most cases, revenge were accompanied by an emotional tone of this morbid sort, given the fact that quite generally human beings have a strong and natural inclination to be revengeful, their claims would seem to commit them to the view that human nature naturally inclines to take pleasure in the suffering of others. But, besides being very uncharitable to ourselves, maintaining this of human nature would be a serious mistake.⁴⁶ This view is strongly supported by several respectable empirical findings. Experimental studies on aggression consistently show that, with rare exceptions, not only do people not enjoy hurting others under normal circumstances

⁴⁶As John Braithwaite pointed out to me, the suggestion in question should also "cause us to be puzzled about how as a species we have survived this long."

when they could do so with impunity, they will not enjoy doing so even if they are annoyed, irritated or seriously angered. In several studies subjects were given full encouragement to administer as severe or mild electric shocks as they pleased to their “experimental subjects” (who in fact were actors, realistically mimicking the effects of such shocks).⁴⁷ Two important things were found. First, despite the ideal conditions for inflicting pain on others without restraint (and for a good scientific cause), the subjects were not in the least inclined to do this beyond a point where their “subjects” started to find the shocks seriously disagreeable. Second, when the experimenters started to anger, irritate, and frustrate the subjects, in administering the shocks, they would at most move up one or two shock levels, say from shock level 4 to level 6.⁴⁸ Although this represented a genuine increment in aggression, the increase was minimal and negligible. The level of shocks administered voluntarily under both conditions was low and totally reasonable in the context of these experiments. With rare exceptions, the subjects did not display sadistic tendencies, not even when they were annoyed, irritated and angered. In view of the fact that revengefulness is a universal human trait, this is an unambiguous and strong indication that describing the emotional tone in question as pleasure in the suffering of another *simpliciter* would be to misidentify the nature of the emotional tone involved.

This conclusion also receives confirmation from some aspects of Milgram’s famous study on obedience to authority.⁴⁹ In these experiments the aim was to find out the extent to which ordinary people were willing to do certain things under pressure from authority which they would be otherwise unwilling to do. Once again, subjects were required to punish “learners” in a “scientific experiment” by administering increasing levels of electric shocks to them when they made mistakes, the maximum being 450 Volts. They were told (falsely) that the experiment was designed to monitor the effects

⁴⁷Buss 1961; Berkowitz 1962.

⁴⁸Milgram 1974, p. 167.

⁴⁹Milgram 1974.

of punishment on learning. The real aim, of course, was to monitor their obedience to the authority of the experimenter(s) who were presented to them as “experts” and in full charge of the experiment. The rather disturbing findings of these studies are well-known: the vast majority of subjects (who were perfectly ordinary people of diverse backgrounds and representative of all sections of society) obeyed authority and went on to administer what they believed to be potentially lethal levels of electric shocks to their “slow-learning subjects.” This is a very revealing finding about our psychology, no doubt, but what is relevant to present purposes is the following observation by Milgram.

In observing the subjects in the obedience experiment, one could see that, with minor exceptions, these individuals were performing a task that was distasteful and often disagreeable but which they felt obligated to carry out. Many protested shocking the victim even while they were unable to disengage themselves from the experimenter’s authority. Now and then a subject did come along who seemed to relish the task of making the victim scream. But he was the rare exception, and clearly appeared as the queer duck among our subjects.⁵⁰

The hypothesis that it was sadistic tendencies, the lust for inflicting pain on others which found the occasion to raise its ugly head in the context of these experiments was disconfirmed in several different settings. In one of these, the subjects were confronted with two incompatible commands issuing from separate but equally powerful authorities. At 150 Volts the learner started to complain vehemently about the pain but the two experimenters radically disagreed as to whether higher shock levels should be administered. Thus, they left it up to the subject to decide whether to give more shocks to the learner or not. Virtually without exception, all subjects broke off the experiment at this point.

Not a single subject ‘took advantage’ of the instructions to go on; in no instance did individual aggressive motives latch on to the authoritative sanction provided by the malevolent authority. Rather, action was stopped dead in its tracks.⁵¹

⁵⁰Milgram 1974, p. 167.

⁵¹Milgram 1974, p. 107.

In another setting in which the choice was left completely up to them, the vast majority of subjects delivered the very lowest shocks to the victim. “By and large, subjects were not inclined to have the victim suffer [despite the fact that] ... the situation provided a setting in which it was acceptable for the subject to hurt another person.”⁵² The same conclusion follows from the results where the experimenter wanted to halt the experiments but the victim himself insisted that, although the shocks were painful, he definitely wanted to go on. He argued that a friend of his had recently gone through the whole experiment and “it would be an affront to his manliness” to do anything less than that. It was found, however, that “not a single subject complied with the learner’s demand.”⁵³ What is more, the great majority of subjects refused to carry on with the experiments as long as the commands came from non-authority figures.⁵⁴ Indeed, a number of them took physical action to protect the victim from such over-zealous non-authority figures.⁵⁵ While it is true that this picture changed when subjects were commanded to go on with the experiment by figures of authority, there can be little doubt as to how this change should not be interpreted. As Milgram also points out,

Those who argue that aggressive motives or sadistic instincts are unleashed when the command to hurt another person is given must take account of the subjects’ adamant refusal to go on in these experiments.⁵⁶

Once again, revengefulness being a virtually universal human trait, these findings can leave no doubt that the emotional tone in revenge cannot be identified with any plausibility as one of taking pleasure in the suffering of another *simpliciter*. As already indicated, I find implausible the suggestion that revenge is even typically accompanied by pleasure in the suffering of another of any sort. The empirical findings above show that even if this were the case, as a matter of fact, the pleasure in question could not be of the

⁵²Milgram 1974, p. 72.

⁵³Milgram 1974, p. 92.

⁵⁴Milgram 1974, p. 97.

⁵⁵Milgram 1974, pp. 97 - 98.

⁵⁶Milgram 1974, p. 104.

repulsive sort I have just discussed. Rather, it would have to be pleasure in the suffering of another *who (is believed to have) harmed one*. Nozick and Glover are committed at least to this interpretation by their claims and this is what must be considered next.

1.3.2 Revenge and the taking of pleasure in the suffering of an aggressor

If we identified the emotional tone in question in this way, the hypothesis in question is not disconfirmed by empirical findings. As a matter of fact psychological research has consistently found that, while pain cues from victims who have not angered the experimental subjects tend to decrease the aggression of these subjects, pain cues from victims who have beforehand aroused the anger of these subjects have a markedly opposite effect.⁵⁷

In angry aggression, [as opposed to instrumental aggression] some form of attack, insult, or annoyance by another person is the stimulus to act aggressively. This leads to anger that can eventually result in aggressive behavior. In this kind of aggression, the intent is to cause harm to the person who was responsible for angering us, and that person's suffering reinforces our aggression.⁵⁸

But, once again, we must be careful how we interpret these findings. For, although the idea that hurting others in revenge involves pleasure in the suffering of those believed to have harmed one is compatible with these empirical findings, these findings are completely silent over why exactly pain cues in people with whom we are angry tend to increase our aggression toward them. We would be jumping our gun far too soon if we concluded that it is pleasure in the suffering of these people which leads us to increased aggression. Of course, it could be this but equally, if not more plausibly, it could be caused by a whole host of other reasons. I shall cover a number of these shortly but first I want to suggest that the phenomenon of increased aggression at the sight of the suffering of those who angered us can be explained without reference to any

⁵⁷Kahn 1984, p. 186. (Figure 7.3)

⁵⁸Kahn 1984, p. 170. (Table A)

other emotions besides our anger, indignation and resentment towards them for what they did.

However, even if such explanations were inadequate or wrong, and increased aggression was due to some emotional tone, it would still be doubtful whether this tone could be plausibly identified as one of pleasure, except perhaps in a minority of cases. There are several other ways of accounting for the emotional tone in question which are equally, if not more, plausible. One of these I have already mentioned: satisfaction at a job well done, a duty discharged - even if that duty necessarily involved pain for another. This may be applicable in situations where the norms of one's society prescribes revenge as a duty but also in situations where there is injury to one's dignity and self-respect as a consequence of some wrong or affront. In such situations one may consider it owing to oneself not to allow the culprit to get away with it.

The tone in question may also be identified as a sense of relief from conditions arising from having been unfairly treated by others. When we are wronged and feel hard done by we may easily feel hurt, demeaned, humiliated, resentful and angry about the incident and a feeling of relief may well accompany a successful act of retaliation which puts an end to these intense and taxing emotions. Standing up for ourselves through retaliation also boosts our self-confidence that we are in form, that we are safe, that we can cope even if a similar aggression comes our way in the future. Not being able to retaliate or defend oneself against aggressions is likely to make one feel insecure and anxious regarding the possibility or likelihood of similar future aggressions. In other words, being able to discourage similar aggressions reassures oneself of one's ability for self-defence. But what is significant in all this, for present purposes at least, is that as soon as this is done, one's oppressive feelings that originated from being wrongfully treated by others are naturally relieved. This relief from oppressive emotions and feelings may well be part of the characteristic experience of satisfaction we are trying to identify and explain.

A different but compatible suggestion is that the emotion in question is better identified as satisfaction at *regaining* one's identity and status as a worthwhile person. This sort of explanation would be especially plausible where part of the motivation for revenge is a wounded sense of honour, self-esteem or even social and public reputation. And if this satisfaction also has an element of pleasure, it is quite plausible to describe it as a pleasure at having been able to do something about the violation against one's personhood, against one's physical, psychological and moral integrity as a worthwhile, important individual. There is no doubt that being considered significant, at least to the point of being treated with respect and consideration by our fellow human beings, is an essential part of, and a prerequisite for, our psychological well-being in a social setting. It matters to us, for example, whether we are greeted by our friends or colleagues at work with a smile, or with a contemptuous, nasty comment. Contemptuous treatment can hurt, let alone the more serious violations of one's personhood, sense of self-worth and dignity.

Yet another way of accounting for the experience in question is by reference to the fact that retaliation, whether instantaneous or delayed (but especially if it is delayed), is extremely demanding and stressful. When retaliation is completed, this stress is released which undoubtedly feels good and pleasurable in the same way as the mere relief from an intense and prolonged physical discomfort or pain is often experienced as a form of pleasure. Moreover, it is not at all clear why, at least in some cases, the emotional tone in question could not be identified in the same way in which Nozick is happy to identify the possible emotions involved in retribution. He claims that in contrast to revenge (which, according to him, involves pleasure in the suffering of another), "retribution either need involve no emotional tone, or involves another one, namely, pleasure at justice being done."⁵⁹ For if in the end revenge turns out to be not as different from justice as Nozick and Glover seem to believe, if it turns out to be a perfectly possible way

⁵⁹Nozick 1981, p. 367.

of doing justice, then surely this explanation cannot be automatically dismissed on *a priori* grounds where the emotional tone of revenge is concerned.

Finally, the notion of satisfaction, as an alternative to pleasure, cannot be over-emphasised in the case of revenge. Even linguistically there is ample warrant for emphasising the notion of satisfaction (satisfaction that justice has been done, or that duty has been discharged, ...etc). When challenging someone to a duel to avenge an affront, one says: "I demand satisfaction" as one throws down the gauntlet;⁶⁰ and again, according to the OED, revenge is an act done to *satisfy* oneself with retaliation for an offence on the offender.

It should be noted that all these alternative explanations for the characteristic feeling of satisfaction in question are quite compatible with one another. My suggestion is that, with possible exceptions, whenever that characteristic emotion is experienced by the avenger, it is one or a combination of these which are at work and not pleasure in the suffering of others in the way claimed by Nozick. But even if this were all wrong and the emotional tone accompanying most cases of revenge was in fact pleasure in the suffering of those who have harmed and wronged one, it is not immediately obvious that these cases of revenge would be morally reprehensible on that account. It is not at all obvious that, on retributivist principles, there is anything wrong with taking pleasure in the deserved suffering of wrongdoers. For if one accepts the view that wrongdoers should or ought to suffer, then it is hard, if not impossible, to see why taking pleasure in the suffering of a wrongdoer should be objectionable, since one is only taking pleasure in the fact that someone got what they deserved.⁶¹

To summarise the discussions of this second main section, I argued against suggestions by Glover and Nozick that revenge is characterised by a morally objectionable feeling of pleasure in the suffering of others. I pointed out that there are

⁶⁰This was pointed out to me by Bob Goodin.

⁶¹Thanks to Michael Tooley for pointing this out to me.

two ways of interpreting their claims and argued first that revenge is highly unlikely to involve morally distasteful pleasure-taking in the suffering of others *simpliciter*. I supported these arguments with experimental findings from research on the nature and psychology of human aggression. I then went on to consider the possibility that revenge might be characterised by a feeling of pleasure in the suffering of an aggressor. I expressed doubts as to whether revenge might be even typically accompanied by such feelings of pleasure and suggested alternative ways in which the emotional tone referred to by Nozick and Glover may plausibly be identified. I especially emphasised the plausibility of identifying the emotional tone in question as *satisfaction* and pointed out that there is ample warrant for doing so, even linguistically. Finally, I noted that, even if all these alternative identifications should turn out to be unviable or inapplicable, on retributivist principles there seems to be nothing wrong with taking pleasure in the deserved suffering of wrongdoers. This suggestion, of course, must ultimately rely on a satisfactory account and defence of moral retributivism. This is a task to be taken up in the last chapter. In the remainder of the current chapter I consider the attempt to exclude revenge from the domain of the morally permissible by denying that it is, and that it can be, a kind of punishment.

1.4 Punishment and revenge

In this section I consider, and argue against, the view that revenge is not a form of punishment at all, that it is only a personal expression of feeling, with no distinctly moral content. According to such views, while punishment is the justified imposition of penalties on an offender by an appropriate public authority for offences prohibited by law, revenge is just a private expression of hurt, resentment and anger towards those who interfered with one's interests. As such, there is nothing moral about revenge to qualify it as punishment.

There are, in fact, two arguments here, bundled into one. The first argument is that punishment requires a civil authority which determines, by law, what counts as

wrongdoing, what the penalties for the various wrongdoings are going to be and which imposes the appropriate penalties. Since private revenge fails this test, it is not punishment at all, but something quite different. Exponents of this view would include, for instance, Hobbes and Rawls.⁶² The second argument is that revenge is just an individual, private expression of emotion which has no moral basis. As such, revenge falls outside the moral sphere to which punishment properly belongs. Once again, the conclusion is that revenge is not punishment. Among the supporters of this view are Mill and Kleinig.⁶³ I want to show that both of these arguments are flawed. I take them in turn, starting with a discussion of Hobbes.

According to Hobbes, there are certain laws in civil society which are made by appropriate public authority prohibiting certain actions, such as, say, stealing and murder. Corresponding to each prohibited action there are appropriate penalties which are incurred for transgressing the aforementioned laws. However, in Hobbes' view, transgression of such a law is not only sufficient for punishment, but also necessary.

A PUNISHMENT, is an evil inflicted by public authority, on him that hath done, or omitted that which is judged by the same authority to be a transgression of the law; ... harm inflicted for a fact done before there was a law that forbade it, is not punishment, but an act of hostility: ... punishment supposeth a fact judged, to have been a transgression of the law;⁶⁴

As Hobbes also observes, from his definition of punishment it follows that "neither private revenges, nor injuries of private men, can properly be styled punishment; because they proceed not from public authority."⁶⁵

I propose to argue that such a legalistic approach to defining punishment is indefensible. Firstly, the crucial presupposition of punishment is *wrongdoing*, and not merely *an action prohibited by law*. Therefore, the question whether some law has been

⁶²Hobbes 1962, Part 2; Rawls, 1955.

⁶³Mill 1962, p. 307; Kleinig, 1973, p. 39.

⁶⁴Hobbes 1962, pp. 277 - 79.

⁶⁵Hobbes 1962, p. 278.

broken is not the only one that is relevant to ask. An equally legitimate, if not more important, question to ask is whether a wrong has been committed. If this is correct, then, short of giving up his legalistic definition, Hobbes would have to say that it is public authority's task and prerogative to determine what counts as moral wrongdoing. But this move renders the project vulnerable to the familiar criticism that making the state to be the sole arbiter of what counts as wrongdoing leaves no basis from which to evaluate laws made by civil authority. The only way to avoid the view that in matters of punishment might is right, therefore, is to acknowledge the existence of pre-legal, moral notions of right and wrong that can serve as a basis for punishment. Indeed, in a well-governed society, such pre-legal moral principles inform and guide the legislators.

Furthermore, apart from the possibility of unjust laws, the civil authority might not have laws determining right and wrong behaviour in every area of human life, such as family and close relationships and associations. Even further, in communities with no civil authority to determine codes of behaviour, there are nonetheless moral precepts and understandings at play, perhaps upheld by custom, history and tradition, which furnish a basis for punishment. In all three cases - that is, where the law is unjust, in areas where there are no laws, and in communities which do not have modern, complex legal codes - on the legal positivistic view harmful, injurious and socially disruptive behaviour would have to go unpunished if legalistic definitions of punishment, such as that espoused by Hobbes, were taken seriously. This would be unacceptable.

A related point is that, as Kleinig points out in his criticism of Rawls, cases which we ordinarily regard as punishment, such as self-inflicted punishments, divine punishment, punishment of children by those other than their parents or guardians, punishment by private individuals where law and order have broken down, etc., cannot be accounted for on a narrow, legalistic definition of punishment.⁶⁶ The legalistic definition of punishment is inadequate on both practical and theoretical grounds and,

⁶⁶Kleinig 1973, pp. 19 - 20.

therefore, it fails to show that revenge is alien to punishment. The following point by Morris is further confirmation of the foregoing criticisms.

Criminal punishment will normally be authorized and carried out by the state, though this need not be the case; to make it part of the proper characterization of the notion would be to beg the question against anarchists.⁶⁷

This disagreement with Hobbes can be expressed in another way. Hobbes, along with a number of other philosophers, uses ‘punishment’ not in the ordinary, neutral sense, but in, what Ingemar Hedenius has aptly termed, the “honorific” sense.

Used in this honorific sense the word “punishment” always refers to a certain way of intentionally inflicting pain or something normally regarded as unpleasant. By calling this infliction of pain “punishment” it is suggested that it is justified. It is embellished, ennobled.⁶⁸

It is worth observing at this point that, ultimately, the question whether revenge qualifies as punishment is important only if the word ‘punishment’ is used in an honorific sense. If it is not assumed, as I argue in this section that it should not be, that punishment is somehow more moral and justified, nothing much would depend on whether revenge is punishment. For, arguably, systems of revenge and judicial punishment are both capable of serving very similar, if not identical functions. Consequently, if punishment was not honorified by philosophers like Hobbes, they would have to leave it genuinely up for argument as to whether judicial punishment or revenge would be the better response to wrongdoing. Such philosophers would also have to recognise that the answer can go either way, depending on the circumstances and the varying degrees of possible focus which can be adopted towards the interests of the offender, the victim or the general public. This, I suspect, is what they might be trying to avoid but the result of that avoidance is a conceptual sleight-of-hand whereby revenge is definitionally excluded from being punishment which, in turn, is honorified.⁶⁹

⁶⁷Morris 1991, p. 55.

⁶⁸Hedenius 1973, p. 13.

⁶⁹As a worthy aside, whether used consciously or not, this dubious technique of conceptual misappropriation is recognisably similar to the appropriation and reservation of such expressions as ‘justice’ and ‘doing justice’ by judicial systems which try to honorify all their decisions and activities

That Hobbes' concept of punishment is honorific, rather than neutral, is also evidenced by the fact that his definition rules out the possibility that an innocent person might be punished, even if by an unfortunate miscarriage of justice. Quinton defends this feature of the honorific concept by claiming that "punishment of the innocent ... is only logically possible if the word 'punishment' is used in an unnatural way, for example, as meaning any kind of deliberate infliction of suffering."⁷⁰

This is obviously not the case. If I damaged someone's car out of envy and jealousy, or sadistically tortured them, that would certainly qualify as a kind of "deliberate infliction of suffering." Yet, it would be most inappropriate to describe my actions as punishment. Contrast that with someone being tried, found "guilty" in legalistic technical terms, and sentenced to life imprisonment in accordance with law for some crime she did not in fact commit. It would not at all be unnatural to say in such a case that the process of justice failed this person and, as a result, the poor woman was punished in spite of her innocence. Thinking of it in retributive terms, she could complain without contradiction or oddity that she was forced to pay the (retributive) price for someone else's misdeed, that she was forced to endure the imposition of punishment which was properly deserved by someone else.

It would be pointless, if not insulting, to tell such a person that if she is really innocent, she is misunderstanding her situation because, by definition, no one who is innocent can be punished. This would be to use, in Hart's words, "the definitional stop."⁷¹ The pointlessness of using the definitional stop consists not only in its unhelpfulness in addressing the injustice and the tragedy which befalls the individual concerned, but also in its failure to take the problem off the hands of the system or

with such positive labels - including decisions and activities which are blatantly wrong and unjust in a more neutral, ordinary sense of the word.

⁷⁰Quinton 1954, p. 140.

⁷¹Hart 1968, p. 6.

practice of punishment which has to take responsibility for imposing penalties on innocent people by mistakenly or wrongfully finding them guilty before the law. Punishment of the innocent, even if unintentional, is a blemish and a problem for any practice or institution of punishment. This is a genuine moral problem and it cannot be defined out of existence by resorting to the definitional stop, or hiding behind an honorific concept of punishment. As Kleinig points out, “in these cases definitions serve only to shift the problems from one point to another.”⁷²

It was an illuminating suggestion by Hedenius that honorific concepts of punishment are ideal concepts.⁷³ They are concepts of what punishment should ideally be. Ideally, punishment should be just, in that it should not be out of proportion, or imposed on someone innocent. It should also be justified, in that there should be good, morally defensible reasons for its imposition. Ideally, however, punishment should also be imposed with due process, by appropriately appointed authorities. For this is the best way to ensure that important principles of justice, such as not punishing the innocent and not overdoing the punishment, are observed, and also that there is a finality built into the process, thus eliminating the danger of conflict escalation. This, I suggest, is the reason why the requirement that punishment be imposed by public authority figures so prominently in honorific definitions of punishment, such as Hobbes'. But, as with the possibility of punishing the innocent, while there are good reasons why punishment is often better placed in the hands of some authority rather than left to private individuals, as Kleinig also points out against Flew, “this is quite different from saying that an imposition *is* punishment *only* if it is imposed by some authority.”⁷⁴

I turn now to the second argument for the view that revenge is not in the domain of punishment, the argument that, while punishment necessarily has to do with the realm of

⁷²Kleinig 1973, p. 13. For expressions of similar views, see Hart 1968; Armstrong 1969.

⁷³Hedenius 1973, pp. 28 - 29.

⁷⁴Kleinig 1973, p. 39.

the moral, revenge is merely a personal expression of emotion, devoid of moral substance. Mill and Kleinig are representative of this position.

[T]he natural feeling of retaliation or vengeance, ... in itself, has nothing moral in it,⁷⁵

The key to revenge lies in the motives of the aggrieved party. For the ground of revenge, and the feature which distinguishes it from punishment, is the fact that someone has *hurt me* or someone close to me. Revenge is the getting of one's own back; the notion of moral wrong is irrelevant to it. ... It is for this reason that revenge can be taken on a person whose actions are recognized to be morally praiseworthy. If *X* murders *Y* and is gaoled for it, I, being a close friend, relative or admirer of *X*, may without logical oddity be said to seek revenge on the policeman who caught him or the judge who sentenced him.⁷⁶

I want to argue, once again, that this is not a defensible position. For, while it is true that revenge is personal in a way that other forms of punishment are not, and even if it is sometimes possible to feel momentary anger-like, non-moral resentment, these factors do not place revenge beyond the domain of the moral. The emotion of resentment which typically characterises revenge is not momentary but is of the enduring sort and, as Butler points out,

... it is not natural, but moral evil; it is not suffering, but injury, which raises that anger or resentment, which is of any continuance."⁷⁷

Contra Kleinig, there is something very odd about him (or anybody else) seeking revenge on the policeman or the judge while believing at the same time that the policeman and the judge were "morally praiseworthy" for catching and sentencing *X*. The most charitable interpretation of such a person, I suggest, would be to regard him to be very confused. Such a person would have failed the process of socialisation in a rather serious way: he was not taught, or he somehow failed to learn, the appropriate emotional responses required by moral belief.

⁷⁵Mill 1962, p. 307.

⁷⁶Kleinig 1973, p. 39.

⁷⁷Butler 1896, Upon Resentment (Sermon VIII), p. 96.

The failure in this case is the violation of the logic of resentment. For, unless a person is forced to take revenge under the powerful sway of custom or social norm which makes it mandatory in that situation, it would have to be resentment which fuels his desire for revenge; and resentment proper is impossible in the absence of some perceived unfairness, injustice or otherwise demeaning treatment. If we tried to understand a person seeking revenge in a situation such as the above, we would want to find out what feature of X's punishment he considered unfair, unjust or improper. Perhaps he found the sentence too harsh or inappropriate for X's physical and emotional condition, or he thought that the policeman used humiliating or excessive force when arresting X, or that X was not given a fair trial, or that in sentencing X, the judge was reprehensible for ignoring those who depended on X. But if the person in question did not think that there was anything amiss with the arrest and the trial and believed that the police and the judge acted in a praiseworthy manner in all respects, it would be very odd if he still wanted to get back at them.

Similarly, in response to Mill we need to ask *why* "the natural feeling of retaliation or vengeance" arises in the first place. The answer to this question lies in an emotion or sentiment which is more basic than the desire for revenge: this is the feeling of resentment. It is resentment that furnishes and fuels the desire for revenge. Were it not for the remarkable human capacity to resent with intensity over prolonged periods of time, revenge would be unlikely to be still with us today.

Resentment, however, as I already intimated, is a complex emotion. It is different from immediate, spontaneous flashes of anger which may arise in response to almost any kind of hurt. Flashes of anger most often die as quickly as they arise, especially in response to accidental hurt. By contrast, however, resentment is a kind of settled, lasting anger and what settles it in the human heart is perceptions of injustice, unfairness, uncalled for personal violation, wrongful injury, a feeling of having been demeaned or badly done by. For example, one might become momentarily angry at being accidentally hit by a falling tree branch but that is likely to be the end of the matter. The reason why it

would make no sense to resent the tree is because the tree is incapable of performing a moral wrong, and therefore an injustice, to anybody. However, one may well feel resentful towards a neighbour if one perceives that they have deliberately lopped the tree in such a way as to hit one with the falling branch. It is such a perception of an unjustified provocation which gives rise to, and fuels, resentment. As Rashdall argued, indignation or resentment at wrong arises naturally and spontaneously in the human mind without any calculation of the personal or social benefits to be derived from gratifying it, ...⁷⁸

It is well to note here that there is an important difference between enduring resentment on the one hand and other well-known negative emotions on the other. Anger, hatred, envy and jealousy, for example, can be directed against others without the presence of that component which is so crucial to enduring resentment which is typically present in revenge: the belief that some kind of unfairness or injustice has befallen one. This kind of resentment can be directed against people only when they are believed to bear some degree of responsibility for unfair, unbecoming, or otherwise morally improper treatment.⁷⁹ The idea of feeling that one is rightfully resentful of someone, as is typically the case in revenge, who at the same time is perceived to have done nothing wrong (either directly or by association) makes no sense.⁸⁰ But, since such perceptions and feelings of injustice or wrongful conduct already presuppose an awareness of what is improper, unfair, and therefore an awareness of what is acceptable, fair and proper, it cannot be the case that the desire for revenge is devoid of moral content. Notions of fairness, unfairness, justice and injustice, propriety and impropriety, etc. are surely as moral as any notions can be.

⁷⁸Rashdall 1924, p. 304.

⁷⁹Richards 1971, p. 253; Gaus, 1990, pp. 284 - 86.

⁸⁰Note that this view can accommodate even Nietzsche's claim that slave morality, such as Christian morality, originated with the *ressentiment* of the herd - the frustrated and impotent feelings of envy, jealousy, resentment and rancour generated in the ugly, the weak, the untalented and the poor towards the beautiful, the strong, the talented and the well off. For, in so far as they are part of *ressentiment*, feelings of resentment can be attributed to the social injustices and the discrimination to which common people felt exposed at the hands of the master class or the aristocracy. (See Nietzsche's "On the Genealogy of Morals" in Nietzsche 1989.)

Of course, in the case of resentment the perceived insult, injustice or wrongful conduct concerns those perceived to have been directed against oneself, not those directed against unrelated others. When improper, wrongful conduct is perceived by us to be directed towards someone with whom we do not have a personal or special tie, our moral response, to the extent that we do have one, is one of indignation.⁸¹ When such conduct is directed towards us in particular, the equivalent and corresponding emotion is one of resentment. This is because resentment is *self-*, rather than *other-*, related. This feature of self-relatedness, however, neither places resentment outside the moral sphere, nor renders it vulnerable to moral criticisms on that account. As Murphy points out,

If it is proper to feel indignation when I see third parties morally wronged, must it not be equally proper to feel resentment when I experience the moral wrong done to myself?⁸²

Similarly, Adam Smith in his *Theory of the Moral Sentiments*, writes that

The violation of justice is injury ... it is, therefore, the proper object of resentment, and of punishment, which is the natural consequence of resentment.⁸³

Finally, in his sermon *Upon Resentment*, Butler speaks in the following vein:

And this seems to be the whole of this passion, which is, properly speaking, natural to mankind: namely, a resentment against injury and wickedness in general; and in a higher degree when towards ourselves, in proportion to the greater regard which men naturally have for themselves, than for others. From hence (*sic*) it appears, that it is not natural, but moral evil; it is not suffering, but injury, which raises that anger or resentment, which is of any continuance.⁸⁴

Although Butler does not insist on a sharp distinction between resentment and indignation, as Murphy does, he nevertheless concurs with the general point that resentment is a response to a perceived moral wrong.

⁸¹Murphy and Hampton 1988, p. 16.

⁸²Murphy and Hampton 1988, p. 18.

⁸³Adam Smith, as quoted in Solomon 1990, p 293.

⁸⁴Butler 1896, *Upon Resentment* (Sermon VIII), p. 96.

This much should suffice by way of showing that resentment is a cognitively based moral emotion,⁸⁵ and therefore that revenge and the desire for revenge are as squarely in the domain of the moral as any (other) form of punishment can be. But there seems to be another way of strengthening this case. Hampton and Murphy point out that another essential part of resentment, besides the perception of a moral wrong, is self-respect.

... the primary value defended by the passion of resentment is *self-respect*, ... proper self-respect is essentially tied to the passion of resentment ...⁸⁶

... the ability to feel resentment following a wrongdoing depends upon one's having enough sense of one's own worth to believe that the treatment is inappropriate and worthy of protest. ... We will also criticize people for not feeling resentment ... if we believe this shows they have too low an evaluation of themselves.⁸⁷

I find this a plausible line of argument. If correct, it furnishes another reason for rejecting the claims made by Kleinig and Mill that the natural feeling of retaliation or vengeance has nothing moral in it. This particular argument against their claims goes as follows: Self-respect is an essential part of resentment and "resentment ... is a good thing, for it is essentially tied to a non-controversially good thing - self-respect."⁸⁸ Furthermore, since resentment in turn is an essential part of the desire for revenge, it cannot be the case that revenge lacks a place within the domain of the moral.

While this argument leaves completely open the question whether any particular instance of revenge is morally justified, it might be objected that only the self-respect of respectable people is good. But even if this were true, it would not affect the argument where it matters most: the moral appropriateness of a victim feeling indignant and resentful in the face of injustice and unfair treatment by others. Moreover, it is questionable whether the undermining of an offender's self-esteem and self-respect could

⁸⁵For an interesting and detailed account of such a cognitivist account of the retributivist emotions, see Mackie, 1985.

⁸⁶Murphy and Hampton 1988, p. 16.

⁸⁷Hampton in Murphy and Hampton 1988, p. 55.

⁸⁸Murphy and Hampton 1988, p. 16.

be desirable. It is at least arguable that low self-esteem is a significant contributing factor to juvenile delinquency, domestic violence, rape and similar forms of criminal behaviour. Rehabilitation from such anti-social behaviour patterns has little chance of succeeding unless there are ways to improve self-esteem and self-respect in those concerned. Indeed, one of the most useful ways to mark the difference between constructive and destructive shaming of wrongdoers is by reference to whether their self-esteem and self-respect is protected (and possibly boosted), rather than being undermined, in the process.

Returning to the question of relating revenge to the moral domain, many thinkers in the history of philosophy have shown a readiness to acknowledge the moral element in revenge, even if some of them were against revenge as a practice.

Revenge, my friends, revenge and the natural hatred of scoundrels, and the ineradicable tendency to revancher oneself upon them, and pay them what they have merited; this is for evermore intrinsically a correct, and even a divine feeling in the mind of every man.⁸⁹

Another example is Bacon who has provided us, perhaps, with one of the most elegant and insightful descriptions of revenge by calling it “a kind of wild justice.”⁹⁰ More recently, in their book *Vengeance: The Fight Against Injustice*, Marongiu and Newman have argued that

All acts of vengeance arise from an elementary sense of injustice, a primitive feeling that one has been arbitrarily subjected to a tyrannical power against which one is powerless to act.⁹¹

Although these authors are by no means favourably disposed towards revenge, they do not hesitate to recognise the fact that the avenger’s motivation consists of a desire for equality, justice and reciprocity. Russell also, in his *History of Western Philosophy*, points out that

... [while] revenge is a very dangerous motive, ... [it cannot be] wholly condemned, for it is one of the forces generating punishment, and punishment is sometimes necessary.”⁹²

⁸⁹Carlyle quoted in Rashdall 1924, p. 304.

⁹⁰Bacon 1985, p. 72.

⁹¹Marongiu and Newman 1987, p. 9.

⁹²Russell 1961 p. 561.

There is also a distinct group of contemporary philosophers, however, who are willing to throw their full support behind revenge. Robert Solomon is representative of this group:

Sometimes vengeance is wholly called for, even obligatory, and revenge is both legitimate and justified. Sometimes it is not, notably when one is mistaken about the offender or the offense. But to seek vengeance for a grievous wrong, to revenge oneself against evil - that seems to lie at the very foundation of our sense of justice, indeed, of our very sense of ourselves, our dignity and our sense of right and wrong.⁹³

Similarly, Susan Jacoby argues that

The personal and social price we pay for the pretense that revenge and justice have nothing to do with each other is as high as the one paid by the Victorians for their conviction that lust was totally alien to the marital love sanctioned by church and state.⁹⁴

Finally, it is worth mentioning Andrew Oldenquist who argues that

... personal accountability makes no sense unless it means that transgressors deserve punishment - that is, they are owed retribution; and there is no doubt that retribution is revenge, both historically and conceptually.⁹⁵

Given the boldness with which Solomon, Jacoby and Oldenquist declare their allegiance to revenge, it hardly comes as a surprise that they are retributivists who identify retribution (including judicial retribution) with revenge. I find the identification of retribution with revenge problematic. This is an issue which I will discuss in the next section.

It is worth noting here that while it is right to claim that revenge is accompanied by powerful sentiments and feelings, the presence of these emotions is quite compatible with there being a moral content in revenge. The perception of an injustice or wrong which gives rise to and keeps fuelling the emotions in question provides, at the same time, as good a moral basis as there can be.

⁹³Solomon 1990, pp. 292 - 93.

⁹⁴Jacoby 1983, pp. 12 - 13.

⁹⁵Oldenquist 1986, p. 76.

To summarise this section, I have argued against the view that revenge falls outside the domain of punishment. In support of this position I argued, first, that the state is not the sole arbiter of what counts as a punishable moral wrong. I also showed that revenge is normally taken in response to perceived offensive, wrongful or unfair conduct and, therefore, that contrary to the suggestions of Mill and Kleinig, revenge does not fall outside the moral domain where punishment is more readily recognised to belong. I also argued that the element of self-respect which is normally present in resentment, and therefore in revenge, provides additional ground for not excluding revenge from that sphere of human concern and activity which we call the moral. I also pointed out that there is a strong philosophical tradition which does not attempt to write revenge out of the moral script, although it is often ambivalent, if not critical, of it. Finally, I mentioned that philosophers who support revenge most strongly are those retributivists who identify retribution with revenge. Their views will be examined in more critical detail in the following chapter.

1.5 Conclusion

In this chapter I critically examined what appear to be the principal culturally and philosophically embedded beliefs and rationalisations in contemporary Western societies for the continued mistrust and ostracism of anything perceived to do with revenge. I argued that there are no philosophically defensible foundations to support these dogmas and ideas. In particular, I argued in the first main section that the charges levelled against revenge by Lewis, that revenge loses sight of the end in the means, that it is a perversion, that it is a nasty vindictive passion and an evil, are unjustified. Following this, in the next main section I challenged suggestions by Nozick and Glover that revenge is characterised by a morally repugnant emotional tone: pleasure in the suffering of another. Apart from questioning whether there would be anything objectionable about taking pleasure in the deserved suffering of wrongdoers, I argued that there are more plausible ways to identify emotional tones which may typically accompany revenge and that the notion of satisfaction would especially be appropriate in this regard. Finally, in the last

main section I explored the question of whether revenge can be properly regarded as a form of punishment and argued in support of an affirmative answer. I argued that revenge has all the morally relevant features to qualify as punishment since it is a punitive response to injustices and wrongs perceived to have been committed against us and those close to us. I also argued that the dominant emotions which typically accompany revenge are the retributive emotions of anger, indignation, but especially resentment, which quite properly arise in us in response to injustices and wrongs. Having examined the most familiar ideas and rationales for the suspicion and low esteem in which revenge is held in contemporary Western cultures, I now move on to the next chapter where I consider the similarly distrusted notion of retribution and its relationship to revenge and the principles of offender justice in punishment contexts.

2

Retribution, revenge and justice

2.1 Introduction

This chapter continues the critical survey of the philosophical literature on revenge in relation to retribution and the principles of justice. I show that, by and large, issues in this area of the philosophy of punishment have received unsatisfactory treatment. Apart from continuing to place later chapters in a wider context, I motivate through these discussions a departure from all previous accounts of revenge. To be successful, such a departure requires first that the crucial notion of retribution, which is very closely related to the concept of revenge, be conceptually disentangled and clarified.

This is no simple task. The nature and the degree of difficulty is indicated by the fact that there is no consensus on what retributivism as a seriously proposed philosophical thesis really is. There is a multitude of proposed definitions, rationales and justifications on offer. John Cottingham, for example, identifies not less than nine senses in which the words 'retribution' and 'retributive', are being used in the literature. The remarkable diversity of these senses is testimony to Cottingham's claim that "the term 'retributive' as used in philosophy has become so imprecise and multi-vocal that it is doubtful whether it any longer serves a useful purpose."⁹⁶

⁹⁶Cottingham 1979, p. 238.

The main reason for this unsatisfactory state of affairs, in my view, is that the philosophical literature on retribution contains at least three different clusters of ideas which are often confused with each other and are indiscriminately jumbled together. The ideas concerned are as follows: First, the definition of retributive punishment, commonly referred to simply as 'retribution.' Second, the principles of justice appropriate to punishment contexts, such as proportionality and not punishing the innocent. Third, moral retributivism, which is a distinctive philosophical position on what justifies punishment.

I clarify the notion of retribution and the principles of justice in the following two sections, while moral retributivism and its justification is covered in Chapter 6. In the remainder of this chapter, I consider the relationship between retribution and revenge. This is an issue over which the contemporary literature in the philosophy of punishment is polarised. According to one view, retribution and revenge are one and the same thing. According to the other view, retribution and revenge hardly have anything in common. In the final main section of this chapter I show that both these views are implausible, and I shall argue in the next chapter that revenge is a species of retribution.

2.2 The definition of retribution

It is unproblematic to agree with Cottingham that the most appropriate and basic sense of 'retribution' is the *repayment* sense. This is the primary meaning that most dictionaries give of the term and, apart from its etymology which is the Latin *retribuo* = *to pay back*, there is an ordinary way of talking about punishment that fits this sense. This is also acknowledged in the literature.⁹⁷ In this sense, retributive punishment is some kind of negative repayment, a kind of TIT FOR TAT. It is the *repayment* of some

⁹⁷Wolgast 1987, pp. 147ff; Cottingham 1979, pp. 238 - 39.

wrong or offence with the imposition of some kind of penalty, pain or unpleasantness.⁹⁸

However, according to Cottingham,

our basic (repayment) sense of 'retribution' ... is not so much a theory as a *metaphor*; a metaphor which ... is central to the basic signification of 'retribution', but which ... cuts remarkably little ice as a justificatory device.⁹⁹

I agree with Cottingham that a simple-minded appeal to the notion of repayment is rather unsatisfactory as a justificatory device. However, such appeals could be regarded as only shorthand justifications which can be given fuller, more satisfying explanations. I present an explication of such shorthand retributive justificatory claims in the last chapter.

A different but popular way of construing retribution is via the idea of *desert*. As with the repayment sense, the kind of desert we are dealing with here is a negative one. According to this formulation, retributive punishment is imposed as a matter of desert for wrongdoing. However, there are difficulties with using the notion of negative desert to *define* retribution. It is not entirely clear, for example, why the desert concept of punishment should be labelled 'retributive.' The idea of imposing punishment on someone because they deserve it does not seem to entail that they are being paid back for what they have done. Even if we adopted the view that desert is a function of the free actions of persons, the entailment from negative desert to repayment is still not straightforward, and requires further explanation and argumentation.¹⁰⁰ All the same, it is not rare to find whole discussions of retribution exclusively in terms of desert as if this were the only correct way to define and understand it.¹⁰¹

⁹⁸The fact that the imagery of the repayment metaphor runs in the other direction as well, in that the wrongdoer is also said to pay with suffering for the wrong committed, does not undermine this way of understanding what retributive punishment is.

⁹⁹Cottingham 1979, p. 245.

¹⁰⁰This suggestion by Michael Tooley could be developed, for example, along similar lines to the justification of retribution I defend in the last chapter by reference to the notions of moral community and responsibility.

¹⁰¹For example, Nozick 1981, p. 36; Braithwaite and Pettit 1990, p. 2.

A more serious problem, however, is that the notion of desert is too restrictive to allow the definition to range over cases of retribution where certain principles of justice are violated. For, just like any other kind of punishment, retribution can also get the wrong person, it can be inflicted for offences which are not moral wrongs, and it certainly can be overdone. The natural thing to say in such cases is that the retribution is unfair, unjust. But in order to allow for this possibility the definition must be left unbound by the principles of justice. This is where, unlike the notion of negative repayment, the notion of negative desert cannot deliver. For example, if five massive TITS are returned for just one little TAT, we have Draconian, disproportionate retribution in the repayment sense. It becomes rather problematic, however, to express the same judgement under a desert conception of retribution because one little TAT is hardly deserving of five massive TITS.¹⁰²

At least a partial honorification of the concept of retribution, therefore, is inevitable, if it is defined in terms of desert. In the honorific sense, retribution cannot, by definition, get the wrong person, it cannot inflict more than what is deserved and, by definition, it is retributively deserved. This kind of definition of 'retribution', however, leaves what we would normally call unfair, unjust cases of retribution without a proper name or a natural category of their own.¹⁰³ Standard construals of retributive punishment, such as the ones put forward by Nozick and Ten, for example, suffer from this kind of malady.¹⁰⁴ The reason for this, I suggest, is that they understand retribution very much in terms of negative desert, which inevitably commits them to some degree of honorification of the concept. As a result, they are unable to accommodate, for example, perfectly ordinary newspaper reports according to which whistle-blowers who complain

¹⁰²A somewhat provocative example is the notion of eternal suffering in hell as punishment for a brief, sinful life.

¹⁰³But see Braithwaite's and Pettit's use of the expression 'just deserts' to denote retribution. (Braithwaite and Pettit 1990.) It is suggestive of the idea that there can be unjust deserts. It is a possibility, of course, that the word 'just' is not meant by them to generate such an impression.

¹⁰⁴The views of these authors will be discussed in detail further below in Section 2.4.2.

about misconduct and corruption in their respective organisations justifiably fear retribution from some of their superiors and colleagues.

Still, the notion of negative desert plays a very important role in discussions of retribution and is not without its proper domain. Its role, however, is not a definitional one, since retribution *per se* cannot be defined by reference to it without running into difficulties of the sort I just pointed out. The fact that it does not range beyond the limits set on punishment by the various principles of justice indicates, in my view, that the notion of negative desert is best regarded as a justificatory device for what is properly termed as ‘just retribution’, retribution within the limits set on punishment by various principles of justice. I shall give an analysis of such justificatory negative desert claims in the last chapter. In the section below I examine the principles of justice with which retribution is often confused.

2.3 The principles of offender justice

There are at least two senses in which punishment can be claimed to be just or unjust. In one sense, the claim is a retributive justification for the imposition of the punishment and manifests in such phrases as ‘the punishment served the offender right’, ‘it is only just, fair, or right, that offenders get the punishment they deserve’ and that ‘punishing serves, or is imposed in, the interests of justice.’ These are shorthand retributive justifications for punishment and I shall give a fuller account of them in the last chapter.

In this section I consider a different sense in which punishment can be just or unjust, the sense which depends on whether certain principles of justice are honoured or violated. In this sense, the justice or injustice of punishment imposed on a wrongdoer consists essentially in the observance or non-violation of certain basic principles, such as not punishing the innocent and not overdoing the punishment. These principles can be conveniently called ‘principles of justice.’ Unlike in the former sense, when punishment is said to be just in virtue of such principles, it remains a further question as to whether

the punishment (which is imposed on the right person, is not out of proportion, ...etc.) is morally justified. Punishment, being generally painful or unwelcome to the person receiving it, has to have good reasons justifying its imposition, and such principles of justice provide none. Their role or function is to set limits to punishment, not to furnish reasons for its imposition. The burden of justification for imposing punishment falls on theories of punishment justification, such as preventionism and retributivism. I argue in this section that the principles of justice are not in the business of punishment justification where retributivism properly belongs and, therefore, that it is mistaken to regard these principles as retributive in nature. They belong to a separate set of their own and nothing but confusion is served by placing them in the 'retributive' category.

One principle of justice which has often been labelled 'retributive' forbids punishing the innocent. J. L. Mackie calls this "negative retributivism,"¹⁰⁵ while H. L. A. Hart calls it "Retribution in the Distribution of punishment."¹⁰⁶ But, as the following quotes illustrate, Mackie and Hart are not alone in regarding this principle as retributive.

[An] important version of the retributive point of view can be called 'distributive'. It insists merely that a penalty should not be inflicted on a person who has not culpably broken a rule.¹⁰⁷

Most contemporary retributivists ... maintain a *minimalist* position. This holds only that no one should be punished *unless* he is guilty of a crime and culpable.¹⁰⁸

Essentially, then, retributivism is the view that only the guilty are to be punished.¹⁰⁹

Another principle of justice which has similarly elicited the 'retributive' label forbids overdoing the punishment. This is the principle of proportionality which sets an upper limit, beyond which punishment is too harsh, in that it is disproportionate to the seriousness of the offence committed. Thus, according to Braithwaite and Pettit, "the negative retributivist is concerned that punishment should not be imposed on the wrong

¹⁰⁵Mackie 1985, p. 207.

¹⁰⁶Hart 1968, p. 12.

¹⁰⁷Walker 1980, p. 26.

¹⁰⁸Golding 1975, p. 85.

¹⁰⁹Quinton 1954, p. 136.

people or be imposed too harshly.”¹¹⁰ Similarly, Mackie regards the latter principle to be one of the “quantitative variants” of retributivism,¹¹¹ while Walker calls it ‘limiting retributivism’:

Most penal codes are also constructed on lines consistent with limiting retributivism, providing *maximum* sentences which set the upper limit to severity without obliging the court to impose the maximum.¹¹²

The expression ‘distributive retribution’ is odd, especially that it stands for the simple idea that the innocent should not be punished. But, more importantly, why use the label ‘retributive’ for basic principles of justice, such as not punishing the innocent and not overdoing the punishment? There is nothing retributive about limiting punishment in these ways. Retribution and moral retributivism are about the reasons one gives *for* imposing punishment - be those reasons explanatory, such as repayment, or justificatory, such as negative moral desert. By contrast, principles of justice merely place constraints on punishment, and even Walker recognises this concerning so-called ‘limiting’ and ‘distributive’ retributivists by saying that

all they offer is principles for restricting punishment, not reasons for imposing it. Only the pure retributivist, who argues that penalties should be imposed because they are deserved, is offering a justification of them.¹¹³

It is true, of course, that moral retributivism looks to the principles of justice to demarcate the boundaries of its claims, for it does not want to justify punishment beyond those limits. But this provides grounds neither for calling these principles retributive, nor for retributivists claiming monopoly over them. Observance of these or similar constraining principles is not unique to retributivism. They are a must in any theory or system of punishment, if it is to be immune of the charge that it perpetuates unfairness and injustice. Braithwaite and Pettit, for example, are quite aware of this point.

¹¹⁰Braithwaite and Pettit 1990, p. 35.

¹¹¹Mackie 1985, p. 207.

¹¹²Walker 1980, p. 26.

¹¹³Walker 1980, p. 26.

It is clearly vital for a consequentialist theory to be able to give constraints ... [a] derived status, for no one can be attracted by the prospect of a system in which agents are unconstrained.¹¹⁴

However, repeating an argument by C. S. Lewis,¹¹⁵ K. G. Armstrong does not think that such a non-retributivist project is feasible. He claims that retributivism is an inescapable part of a just system of punishment.

In the area of the moral justification of the practice [of punishment] a retributive theory is essential, because it is the only theory which connects punishment with desert, and so with justice, for only as a punishment is deserved or undeserved can it be just or unjust. What would a just *deterrent* be? The only sense we could give to it would be a punishment which was just from a retributive point of view and which also, as a matter of fact, deterred other people. 'But' it may be objected, 'you are only talking about *retributive* justice.' To this I can only reply: What other sort of justice is there?¹¹⁶

There are two ways of interpreting the above argument. As I have already pointed out, there are two senses in which punishment can be said to be just or unjust. Accordingly, one interpretation of the claim that the justice of punishment consists in the person's desert is that a person's desert is what justifies his punishment. Under this interpretation, the claim is that only the (negative) desert of retribution justifies the imposition of punishment. The other interpretation of the argument that the justice of punishment consists in the person's desert is that a person's desert is what sets the proper limits of his punishment. Under this interpretation, the claim is that only the (negative) desert of retribution can furnish grounds for principles of justice by reference to which punishment may be termed just or unjust.

Although it is not clear which of these interpretations Lewis and Armstrong had in mind, with all likelihood they would endorse both. But however that may be, their

¹¹⁴Braithwaite and Pettit 1990, p. 33.

¹¹⁵While Armstrong refers to Lewis's article only in other contexts, in my opinion it was C. S. Lewis who first made the crucial point at the heart of the quoted passage. Lewis writes: "The Humanitarian theory removes from Punishment the concept of Desert. But the concept of Desert is the only connecting link between punishment and justice. It is only as deserved or undeserved that a sentence can be just or unjust. ... There is no sense in talking about a 'just deterrent' or a 'just cure.'" (Lewis 1949, p. 6.)

¹¹⁶Armstrong 1969, p. 155.

argument is indefensible under both interpretations. The first interpretation suffers because there are good consequentialist reasons for imposing punishment. To take a familiar example, if other measures fail to serve the important goal of crime control, deterrence is a weighty reason for imposing punishment. Protection of the community from crime is essential and it cannot be dismissed as not being a legitimate reason for imposing punishment within generally accepted limits.

The argument is similarly untenable under the second interpretation because a retributive theory of punishment is not the only theory which can limit punishment in line with the principles of justice. There are at least three different non-retributive ways of approaching these principles and I shall discuss them shortly. Before that, however, it will be worthwhile to examine the grounding of these principles in some kind of notion of desert. For, if there is an initial ring of plausibility to the claim in question, I suggest that this is largely due to the idea that the principles of justice are somehow desert-based. Lewis and Armstrong make a strong appeal to this idea when they claim that only as a punishment/sentence is deserved or undeserved can it be just or unjust.

While non-desert-based groundings of the principles of justice render the Lewis-Armstrong claim ultimately untenable, their argument nevertheless retains an appearance of plausibility for two reasons. One reason is that appealing to a notion of desert is a familiar and simple way of grounding the principles in question. The second reason is that this appeal to desert is easily confused with the concept of *negative* desert which is frequently used to justify the thesis of moral retributivism. The distinction between the two kinds of desert is subtle, but significant. Insofar as the notion of desert is appealed to by an account of moral retributivism as a justificatory reason for imposing punishment, the appeal is to a *negative* notion of desert. But, while this notion may be used to justify imposing punishment on the guilty, it is hard to see how it could be taken to forbid the punishment of the innocent. The claim that the guilty should be punished because they deserve it does not entail that, therefore, the innocent should not be punished because they do not deserve it. For there may be good reasons for imposing

punishment on people other than their negative deserts. Denying this would be implausible. Indeed, consequentialists in general would consider people's negative deserts to be the least of the reasons justifying punishment.

If this is right, then desert claims embodied in the principles of justice, such as the claim that the innocent should not be punished because they do not deserve to be punished, must be grounded in a somewhat different notion of desert from that of the justificatory negative desert claim of moral retributivism. This notion, I suggest, is a *positive* notion of desert, for the reasoning here is no longer about what the guilty deserve in a negative way, but rather about what the innocent and the guilty may be said to deserve in a more positive way in matters of punishment: To the extent that they have been law-abiding in the sense that they have not broken the law, they deserve not be framed and punished, or overpunished, either as a way to serve the interests of utility (deterrence), or as a way to impose on them the negative deserts of others.

But, even though the grounding of the principles of justice in a notion of positive desert is possible, it is only in a very old fashioned sense that this sense of 'desert' may still be called 'retributive' - a sense akin to the one used in the phrase "never did a charitable act go away without the retribution of a blessing."¹¹⁷ There may be other ways, of course, to link the principles of justice to the notion of desert. It may be possible, for example, to construe retribution in such a way that both its justification and the limiting principles of justice issue from a common, more basic principle, such as the principle that *Happiness and unhappiness should be proportioned to virtue and vice (or right and wrong, or good and bad conduct), respectively.*¹¹⁸ While such appeals to a cosmic sense of fairness and justice may be workable, and even plausible, the pertinent point remains: Non-desert-based groundings of the principles of justice render ultimately

¹¹⁷Seventeenth century quotation from OED, as quoted by Cottingham, 1979, p. 239.

¹¹⁸This was suggested to me by Michael Tooley.

untenable the second interpretation of the Lewis-Armstrong claim that only as a punishment/sentence is deserved or undeserved can it be just or unjust.

This is because, as indicated above, there are at least three non-retributive ways of approaching the relevant principles, each of which constitutes a counterexample to the second interpretation of Armstrong's claim, according to which a retributive, desert-based approach to these principles is inevitable. The first counterexample is provided by the fact that it is quite plausible to regard these principles as having an immediate, underived moral authority, independently of any notions of repayment, desert or, for that matter, utility. This much at least has been correctly recognised by Mackie.

This is undoubtedly true of negative retributivism, including its quantitative variant ... [that it has] an immediate appeal and underived authority.¹¹⁹

The second counterexample is provided by the fact that it is also possible to adopt these principles as utilitarian secondary rules.¹²⁰ Braithwaite and Pettit, for example, derive these principles in their comprehensive, republican theory of criminal justice from the consequentialist goal of promoting dominion (republican freedom).

The right of the innocent not to be punished, and the upper limit we put on the punishment of the guilty, are both derived within our theory as measures required for the promotion of dominion. ... unless such measures are firmly in place the dominion of people in the society at large will be seriously compromised. Although consequentialist in structure, our theory ... is rights- and limits-respecting.¹²¹

A third counterexample can be provided by grounding these principles in a Rawlsian concept of 'fairness' along lines suggested by Walker.¹²² Suppose that the choice is to be made from a Rawlsian-type original position between two kinds of societies. In societies of type A, the principles of justice, such as not punishing the

¹¹⁹Mackie 1985 p. 208.

¹²⁰Cottingham 1979, p. 241.

¹²¹Braithwaite and Pettit 1990, p. 207.

¹²²Walker 1991, pp. 92 - 95. Walker develops this idea only in relation to the principle that the innocent should not be punished.

innocent and not overpunishing the guilty, are honoured as inviolable limiting principles of punishment. By contrast, in societies of type *B*, there is no provision for an upper limit on punishment and innocent people may be framed and punished in cases where the police find it either impossible, or too costly, to find and punish the real offenders. This might be done (covertly, of course), in the interests of deterrence and political image management.

Placed in a Rawlsian-type original position, having no way of knowing whether one would in fact be an offender and whether one would in fact fall victim to the above mentioned type of political expediency, which kind of society would a rational person choose to live in? The choice seems fairly clear. A rational, self-interested chooser would opt to live in a society of type *A*. There would be nothing to be gained by choosing otherwise. Since there is no guarantee that a *B*-type society would have a lower crime rate than an *A*-type society, one would only expose oneself to heightened risks of unwanted and undeserved punishments for no discernible gain.¹²³ Once again, this account is non-retributive and it makes no appeal to any notions of desert. Furthermore, unlike the second account, it is not instrumental either. In relation to his Rawlsian account of the principle that the innocent should not be punished, Walker rightly argues that

Rawls's non-altruistic man is not a utilitarian. The society which he would choose for its fairness is not necessarily the one which would have the lower crime-rate. Nor does Rawls's kind of 'fairness' involve the notion of desert: simply rational choice of a self-interested sort.¹²⁴

¹²³ Arguably, the ultimate viability of the Rawlsian approach may be suspect. Michael Tooley suggested to me that it might be reasonable to expect that *B*-type societies would have a lower crime rate than would *A*-type societies, which could make rational choice in this instance much harder than it would first appear.

¹²⁴ Walker 1991, p. 93. Walker makes out a good case for distinguishing between two versions of the principle that the innocent should not be punished. One is what he calls 'the wrong-person rule', the other one 'the blameless-doer rule'. (pp. 88 - 95) But then, for some unexplained reason, he unnecessarily weakens the force of his Rawlsian account of these rules by construing type *B* societies to be ones in which law abiding citizens do not run the risk of being framed for things they did not do. Only if a person "was incarnated as a law-breaker [would he] be at risk of being penalized not only for his own offences but also for those which the authorities found it difficult to pin on anyone else." (p. 93) A world in which all citizens run the risk of being penalised for things they did not do would be even less preferable than the one in which this is true only of law-breakers.

None of these three approaches - namely, where the principles of justice are regarded as morally basic, self-evident and autonomous, where they are derived as secondary rules as a means of maximising promotion of some value in a consequentialist framework, and where they are grounded in a Rawlsian conception of fairness - concedes one iota to ideas of retributivism, such as the idea that punishment is due as repayment or desert for an offence or wrong. Moreover, none of them appeals to any ideas of desert either. The Lewis-Armstrong argument, therefore, is mistaken under the second interpretation as well. Neither a retributive theory, nor a desert theory is essential in a moral justification of the practice of punishment. The principles of justice which set the limits of fair punishment certainly need not be based on any notion of desert, let alone a retributive one.

If this is right, then it is clear that we have an alternative answer to Armstrong's question "What would a just deterrent be?": Just deterrent punishment is punishment imposed on someone for purposes of deterrence while fully observing the principles of justice which, as I have argued, can be non-retributively grounded. As to Armstrong's rhetorical remark "What other sort of justice is there (besides *retributive* justice)?," the answer is similar: There is *criminal justice* consisting of punishment imposed for non-retributive reasons, such as deterrence and the promotion of dominion, punishment which, at the same time, fully honours non-retributively grounded principles of justice, such as not punishing too harshly and not punishing the innocent.

The above arguments are sufficient to establish that principles which prohibit punishment of the innocent and the overpunishment of the guilty are independent of any notion of retribution. Along with other principles of justice, such as the ones listed below, they have an immediate, underived but powerful moral authority all on their own. Their violations automatically constitute serious moral wrongs which cannot be easily justified. The appropriateness of these principles is virtually unquestionable. I suggest that, far from belonging to the retributive category, these principles properly belong to,

and are a natural part of, a set of principles to which some (if not all) of the following may be claimed also to belong.

- 1) Punishment must be for a moral wrong, or the transgression of some defensible rule, standard or law.
- 2) Punishment must not be imposed on the wrong person, an innocent person who is not appropriately responsible for the offence in question.
- 3) Punishment must not be overdone; it should not be harsher than the gravity of the offence warrants. Punishment is unjust to the extent that it is out of proportion.
- 4) Punishment must be equitable and comparable to other punishment for similar offences (at least within the same community). To exaggerate, it would be blatantly unjust if a person were to be sentenced to ten years in jail for an offence for which other offenders have received only minor fines.¹²⁵
- 5) Punishment should not be imposed without due process, without giving the accused a fair chance to argue and defend his or her case.

This list may not, of course, be complete. However, the interesting thing is that retributivists can also (decide to) violate the principles of justice (though not with a retributivist justification, or without being reprehensible, perhaps) without ceasing to be retributivists on that account. Resorting to somewhat provocative examples, with respect to the principle of proportionality, we only need to remember Biblical and Islamic notions that Divine Retribution for a short, unrepentant sinful life is Eternal Hell.¹²⁶ With respect to the principle of not punishing the innocent, it is well to remind ourselves of the

¹²⁵At the same time, in the interests of restoration, it is highly desirable that there be provisions for flexibility and plurality in sentencing in order to take advantage of diverse cultural and community values. Such flexibility is necessary, for instance, for models of justice based on reintegrative shaming. As Braithwaite puts it in the context of youth justice, "because the whole idea of the process is to empower local communities to come up with their own approach to dealing with the life problems of a particular young person and their victims, plurality and unpredictability is inherent in the strategy." (Braithwaite 1993, p. 40.) Also, in response to criticisms that community- and family-based diversionary conferences (to be described in more detail in Chapter 6) may impose harsher or lighter penalties on offenders than a court would, McElrea made the telling point that "sentencing is not an exact science and there can be considerable disparity between the sentences imposed by different judges in similar cases." (McElrea 1993, p. 4.)

¹²⁶It must be recognised, however, that many believers from both faiths have rejected this view.

scapegoat of the Old Testament, and of Christ's innocent (even if voluntary) suffering on the cross paying the (retributive) price for our sins in accordance with the demands of (God's) Divine Retribution, as it is often interpreted in Christian theology.¹²⁷ Susan Jacoby makes an observant remark about the independence of the core ideas of retribution and the principles of justice when she points out that: "Retribution per se is an integral component of just as well as unjust legal systems."¹²⁸

Furthermore, as I already indicated, non-retributivist moral theories of punishment can also adopt these principles (along with similar ones) either as utilitarian secondary rules, because their adoption maximises overall utility,¹²⁹ or as limiting principles, recognising their independent moral authority as principles of natural justice which should not be violated.¹³⁰ This last possibility also disposes of another misconception about non-retributive moral theories of punishment, that they sanction injustice, such as overpunishing and the punishment of the innocent. I hope to have shown by now that while non-retributivism has the potential to sanction violations of the principles of justice, it does not have to, just as retributivism's lack of any considerable potential to justify violations of those principles does not amount to an ability to prohibit their violation.

2.4 The relationship between retribution and revenge

In this section I consider the relationship between retribution and revenge. This is an issue over which the contemporary literature in the philosophy of punishment is polarised. According to one view, retribution and revenge are one and the same thing. According to the other view, retribution and revenge hardly have anything in common. I

¹²⁷A secular interpretation of the notion of Divine Retribution is the normative, if sometimes wishful, notion of Cosmic Justice, that the good (ought to) prosper while the wicked (ought to) suffer in some way in proportion to their respective deserts. This is what Lewis refers to when talking about "the universal human feeling that bad men ought to suffer." (Lewis 1957, p. 81.)

¹²⁸Jacoby 1983, p. 115.

¹²⁹A point also acknowledged by Golding in a footnote. Golding, 1975, p. 85.

¹³⁰Similar point made by Cottingham, 1979, p. 241.

show that both these views are untenable. As I shall argue in more detail in the next chapter, revenge is a species of retribution.

2.4.1 Revenge and retribution as identical

The identification of retribution with revenge is particularly implausible in cases where the agent of punishment bears no special tie to the victim of the crime. Nozick also supports this view:

Revenge is personal: "this is because of what you did to my -----" (self, father group, and so on). Whereas the agent of retribution need have no special or personal tie to the victim of the wrong for which he exacts retribution.¹³¹

This point, simple as it may be, is a critical one. Yet, it seems to be lost on advocates of the identity thesis between retribution and revenge. As I already mentioned, one such advocate is Jacoby:

The very word "revenge" has pejorative connotations. Advocates of draconian punishment for crime invariably prefer "retribution" - a word that affords the comfort of euphemism although it is virtually synonymous with "revenge." ... The death penalty is certainly a form of legalized vengeance - revenge writ large - but so is any lesser punishment if a crime is unlikely to be repeated.¹³²

Whilst I would agree with Jacoby that the social ostracism of revenge has reached a height where "the very word 'revenge' has pejorative connotations," she is mistaken in claiming that 'retribution' is merely a "euphemism" for 'revenge.' Retributive punishment through a judicial system, no matter how harsh, does not qualify as revenge. The requisite personal elements having been eliminated from such punishment, on account of its retributive character, it still qualifies as retribution - institutionalised, judicial retribution - but not as revenge.

¹³¹Nozick 1981, p. 367. See also Kleinig 1973; van den Haag 1975; Wallace 1995.

¹³²Jacoby 1983, pp. 4 & 6.

Robert Solomon seems to be aware of the personal nature of revenge. He points out, for example, that “vengeance retains the virtue of being personal.”¹³³ Yet he also claims that “retribution and revenge are one and the same”¹³⁴; and again that “vengeance and retributive justice are in the end identical”¹³⁵ - apparently failing to realise that the personal involvement in question is not always present in retributive punishment. To the degree that Solomon is willing to admit that a personal involvement may not be present in retributive punishment, his position in identifying retribution with revenge becomes inconsistent. This comes out quite clearly when he attempts to reconcile the identity thesis with the highly compelling distinction between the two pointed out by Nozick:

One might, adopting a distinction from Robert Nozick, say that retribution is justified revenge, where revenge is strictly personal but retribution is not.¹³⁶

But how could retribution be identified with revenge, even if the latter is justified? For even if an act of revenge is justified, it still remains personal in a way that an act of retribution might not. Therefore, the conclusion must remain that the revenge-retribution identity thesis is indefensible. To repeat Nozick’s point, while acts of revenge are necessarily personal, acts of retribution need not be. Cases in point are when criminals are retributively punished through the legal institutions specifically set up for the purpose. More often than not, these are cases that involve people whose only reason for participating in the process of punishment is that they have to do it as part of their job in that institution.

An interesting case has been put forward for the view under criticism by Oldenquist who, in addition to the conceptual identification, makes an historical claim about identifying retribution with revenge.

¹³³Solomon 1990, p. 302.

¹³⁴Solomon 1990, p. 300.

¹³⁵Solomon 1990, p. 301.

¹³⁶Solomon 1990, p. 300.

Can decent people accept the idea that retributive justice is revenge? ... there is no doubt that retribution is revenge, both historically and conceptually."¹³⁷

While it is true that historically revenge has been the most prevalent form of retribution, it has not been the only one. Retributive punishment has, from time to time, been executed on wrongdoers by vigilantes who did not have the requisite personal relationship to the victims.¹³⁸ Furthermore, judicial retribution, where the victim and those closely related are deliberately prevented from taking part in the punishment, has been part of our history for a long time now. In cases like these, where the personal element is absent, it is inappropriate and misleading to identify the retribution with revenge. Contra Oldenquist, there is a clear difference both historically and conceptually between purely institutionalised retribution and vigilante retribution on the one hand and personal retribution on the other. This difference is marked by the presence or absence of a personal or special tie between the victim of the injustice or offence and the agent of retribution.

Oldenquist's arguments for conceptually identifying retribution with revenge are subtle, however:

... judicial retribution is not mere revenge but 'sanitized' revenge, revenge over which a moral community assumes stewardship ...[and which is] administered predictably, impartially, publicly, and relatively effectively by police and judges who are not the criminal's victims or relatives.¹³⁹

I agree that there is a sanitising process involved in moving from simple revenge (which is set and carried out by the victims or their relatives) to judicial retribution. But this move, or evolution, is not necessarily from one form of revenge to another. For, to repeat the earlier point, if there is no personal involvement, it is no longer appropriate to talk about revenge. My claim is that the move from revenge to judicial, institutionalised retribution, at least as we currently know it, is from one form of retribution to another.

¹³⁷Oldenquist 1986, p. 76.

¹³⁸Dimsdale 1866.

¹³⁹Oldenquist 1986, pp. 76 - 77. See also his 1988 about some formally set out conditions of this sanitising process (p. 474).

What is common in both is not revenge, but retribution pure and simple, whether construed as repayment, desert, fair play, etc.¹⁴⁰

If the personal element were retained through this process of institutionalisation, if the victims and their relatives retained their involvement, if they had a say or a hand in the punishment of their wrongdoers, the imposed punishment would quite appropriately be called institutionalised or sanitised revenge. But where all personal involvement of this nature is excluded and the institution takes the retribution completely out of the hands of the victims concerned, it is most inappropriate still to regard such punishment as revenge.

A slightly different way of understanding Oldenquist's claim, one which is in fact present in his own work but is run together with the above ideas, is to think of institutionalised retribution in our institutions where no personal motives are involved as being "society's revenge on its own members who harm and betray it."¹⁴¹ But this is to over-extend the meaning of the word 'revenge.' For there is no personal involvement here in the relevant sense of the word that would make it appropriate for purely institutionalised punishment to count as revenge.

There are two reasons for this. The first is that society, and those involved in enforcing the law and administering punishment, are equally closely related, or not related at all, to the victim or the offender, so the tie in question is the same to both. The second, related point has to do with the motivation of those who determine and mete out the punishment in the institutionalised legal system. They are not motivated by a personal tie to the victims, whom they customarily would not know or identify with. They are appointed to serve the public interest rather than their own private interests.

¹⁴⁰On the various construals of retribution see Cottingham 1979.

¹⁴¹Oldenquist 1986, p. 77.

Both these points contrast with the case of revenge. First, as a general rule, the avenger has a closer tie to the victim than the offender.¹⁴² Second, the avenger becomes involved in the punishment of the wrongdoer at least in part because of this personal tie they have to the victim. Hence, judicial retribution visited on the offender does not qualify as revenge in any serious sense of that term. It is rather institutionalised retribution which is distinct and clearly distinguishable from both vigilantism and revenge. The claim that institutionalised retribution in our institutions is society's revenge on its own members, therefore, can hardly be more than a metaphor. Taking it as literal truth would require an unacceptable dilution, if not trivialisation, of our notion of a personal or special tie.

We already have three familiar forms of retribution, namely, institutionalised, vigilante and personal retributive punishment. Recognisably, this last one is what normally goes by the name *revenge*. A less familiar form of retribution, however, is a combination of the first and last of these: institutionalised personal retribution or *institutionalised revenge*.

Apart from ignoring the fact that the requisite personal involvement is not always present in retributive punishment, a conceptual identification of retribution with revenge not only disregards, but also allows no satisfactory account for, the differences between the various forms of retribution which are clearly there at the pre-theoretical level. A situation where we have both the words, 'retribution' and 'revenge,' standing for one and the same thing, with the added result that certain compelling pre-theoretical differences are rendered hard to account for, is far from satisfactory. Taking seriously the presence or absence of a personal involvement as described above provides an immediate and straightforward solution to these theoretical difficulties. The presence or absence of a personal or special tie seems to provide the best basis all around for a

¹⁴²However, closeness need not always be along biological or family lines. One may feel much closer to a good friend, for instance, than to some members of one's family. I consider in more detail the nature of the relevant personal or special tie further below in Sub-section 2.4.2.

satisfactory explanation of the highly compelling distinction between revenge and other forms of retribution.

A more subtle version of the identity thesis has been expressed by Mill.

The sentiment of justice, in that one of its elements which consists of the desire to punish, is thus, I conceive, the natural feeling of retaliation or vengeance, rendered by intellect and sympathy applicable to those injuries, that is, to those hurts, which wound us through, or in common with, society at large.¹⁴³

The crucial claim that Mill makes in this passage is that people's desire to punish for the sake of retributive justice is but a desire for revenge. In addition to the problems already raised, and which must be faced by any version of the identity thesis between retribution and revenge, there is one more objection to which Mill's claim is vulnerable. For, in claiming that people's desire to punish for the sake of justice is but a desire for revenge, Mill has put the cart before the horse. As I have already indicated, some notion of fairness or justice is necessary for explaining resentment and resentment in turn is necessary for explaining the desire for revenge. Therefore, it is circuitous and mistaken to appeal to the desire for revenge to explain the desire to punish for the sake of justice.

Resentment is a feeling that is aroused and maintained by perceptions of injustice and unfair treatment. But such perceptions *presuppose* some kind of sense or idea of what is fair, what is just and unjust. Hence, the sentiment of justice can hardly have as its basis the natural feeling of vengeance. If anything, it must be the other way around.

In conclusion, Mill's attempt to identify the desire for retributive justice with a desire for revenge is untenable. Hence, this particular version of the identity thesis fails as well.

Having shown that an identification of retribution and revenge is untenable, the question remains: how exactly are the two related? In the section below I consider the

¹⁴³Mill, 1962, p. 307.

suggestion that retribution and revenge are different species of a common genus at best and, therefore, that they are in different moral categories.

2.4.2 Revenge and retribution as different species

Being a defender of retribution, Robert Nozick seems to be troubled by the fact that retributivism - which he defines as *the view that people deserve punishment for their wrongful acts even if such punishment fails to act as a deterrent* - “strikes some people as a primitive view, expressive only of the thirst for revenge.”¹⁴⁴ In response, unlike Solomon and Oldenquist, Nozick attempts to distance retribution from revenge as much as possible by arguing that there are several key differences between the two. These, according to Nozick, are that:

- (1) Retribution is done for a wrong, while revenge may be done for an injury or harm or slight and need not be for a wrong.
- (2) Retribution sets an internal limit to the amount of the punishment, according to the seriousness of the wrong, whereas revenge internally need set no limit to what is inflicted. Revenge by its nature need set no limits, although the revenger may limit what he inflicts for external reasons.
- (3) Revenge is personal: “this is because of what you did to my -----” (self, father group, and so on). Whereas the agent of retribution need have no special or personal tie to the victim of the wrong for which he exacts retribution.
- (4) Revenge involves a particular emotional tone, pleasure in the suffering of another, while retribution either need involve no emotional tone, or involves another one, namely, pleasure at justice being done.
- (5) There need be no generality in revenge. Not only is the revenger not committed to revenging any similar act done to anyone; he is not committed to avenging all done to himself. Whether he seeks vengeance, or thinks it appropriate to do so, will depend upon how he feels at the time about the act of injury. Whereas the imposer of retribution, inflicting deserved punishment for a

¹⁴⁴Nozick 1981, p. 36. Note: With Nozick, and also generally, one must distinguish between retribution as an act, or practice, of punishment and retributivism which is a philosophical thesis that explains and possibly justifies the practice. A defence of moral retributivism is provided in Chapter 6.

wrong, is committed to (the existence of some) general principles (prima facie) mandating punishment in other similar circumstances.¹⁴⁵

Commenting on Nozick's account of retributivism, C. L. Ten is of the opinion that "Nozick usefully distinguishes between retributive punishment and revenge with which it is sometimes confused" and claims that there is in addition "one other important difference between revenge and retribution, which could be added to Nozick's account."¹⁴⁶ This is the following:

(6) Retributive punishment is only inflicted [knowingly] on the wrongdoer, whereas revenge is sometimes inflicted [knowingly] on an innocent person close to the revengee, either because this is an easier target, or because it is thought that this would hurt the revengee more.¹⁴⁷

With the exception of the third one, with which I agree, all of these claims are flawed. Though neither author spells this out, judging from the context, and adopting a familiar classificatory terminology from biology, their view is that revenge and retribution are different species of some common genus at the very best. With the venerable exception of the third point, these comparisons are incompatible with the view that retribution and revenge might be more closely related. For want of a better name, I shall refer to the position held by Nozick and Ten as the (*no closer than*) *species-species categorisation* of retribution and revenge - henceforth 'S-S categorisation.'

The first reason for distrusting the S-S categorisation arises from the fact that under such a schema miscarried acts of retribution under, say, an entirely institutionalised system of punishment, have no natural category of their own. Suppose, for example, that punishment is imposed on someone for retributivist reasons through an institutionalised system for something that is not a wrong or for something of which the person is not guilty, or that the punishment is out of proportion to what is morally just and justified in the circumstances. Such punishment is clearly unjust. Normally, if it

¹⁴⁵Nozick 1981, pp. 366 - 68.

¹⁴⁶Ten 1987, pp. 42 - 43.

¹⁴⁷Ten 1987, p. 43.

was imposed for retributivist reasons, we would call such punishment misplaced retribution or unjust retributive punishment.

Yet the S-S categorisation does not allow us to do this, at least not in the way constructed by Nozick and Ten. In their sense, punishment of this kind cannot be either retribution or revenge. It is not retribution because it is incompatible with what they claim about retribution as part of their attempt to ground the S-S categorisation, namely, that retribution is always in accordance with the principles of justice. It cannot be revenge either, since such punishment is not personal in the requisite sense of the word. But, if such an unjust case of punishment is neither retribution, nor revenge, there seems to be no straightforward answer to the question as to what in particular it could be.

Difficulties of this kind are only the tip of the iceberg of problems for the above comparisons. Before dealing with them individually, I start with a general criticism that has to be levelled against comparisons 1, 2, 5 and 6. In these Nozick and Ten compare revenge pure and simple not with retribution pure and simple but with *just* retribution, that is, retributive punishment which is in accordance with principles of justice, such as that punishment must always be for a wrong, that it must get the right person, that it must be proportional to the seriousness of the wrong, etc. But, as I have already argued above, these principles are quite independent of the notion of retribution as such, regardless of whether retribution is construed as repayment, desert, or whatever.¹⁴⁸ Wittingly or not, Nozick and Ten choose to build into *their* notion of retribution these principles of justice and it is in virtue of this that for them retribution is necessarily for a wrong, that it is necessarily proportional, that it necessarily gets the right person, and so on. Since they do not build the same principles into their notion of revenge, it is hardly surprising that for them these are not defining features of revenge. But, in order to get a proper and fair contrast between retribution and revenge in these respects, the principles of justice will have to be built either into both of them, or into neither of them.

¹⁴⁸See Section 2.2 above.

I move on now to consider each comparison individually:

(1) Retribution is done for a wrong, while revenge may be done for an injury or harm or slight and need not be for a wrong.¹⁴⁹

Nozick is mistaken in claiming that retribution must always be for a moral wrong and he is mistaken in thinking that retribution differs from revenge in this regard. Retribution *per se* is a perfectly possible response to morally good actions. A good example is provided by the *Stewart Inquiry* into the administration of Australian soccer. Soccer personalities were reluctant to testify about corruption in the administration of the game for “fear of retribution in terms of promotion within the sport” and “fear of actual physical violence.” In Justice Stewart’s words,

There were grave misgivings, many of which were justified, that if material of this nature were to find its way into the hands of vengeful people ... retribution would follow.¹⁵⁰

Justice Stewart and Nozick can’t both be right about the proper use of the term ‘retribution.’ My claim that the mistake lies with Nozick is further confirmed in the case of the NSW Police Royal Commission conducted by Justice Wood. According to reports, police officers (whistle-blowers) who report police misconduct fear and face “retribution” in the form of “a series of old or trivial complaints against them” and “disciplinary matters.”¹⁵¹

Moreover, just like revenge, even retribution which is institutionalised through the legal system need not be for a moral wrong. As I pointed out in Section 1.2, it is possible that the legislators create morally unjust laws through bad advice, poor judgement, error, corruption, or undue pressure from small but powerful interest groups. The laws being unjust in the first place, transgressing them may well not count as a moral wrong. Nevertheless, such a transgression still amounts to an offence and may be

¹⁴⁹Nozick 1981, p. 366.

¹⁵⁰Cockerill, M. Soccer Scandal: Judge Urges Police Investigation/Involvement. *The Sydney Morning Herald*, Wednesday, January 11, 1995, p.1.

¹⁵¹Brown, M. Informers Face Retribution. *The Sydney Morning Herald*, Wednesday, December 14, 1994, p. 9.

retributively punished through the standard channels. The retributive character of such punishment is not changed by the injustice of the laws concerned. In such cases we would say that the punishment was unjust because the person being punished did no moral wrong.

Consider the following example: The ruler decrees for self-serving reasons that all boys under two are to be slaughtered, and that anybody who hinders this order be put to death as retribution for their defiance. The parents who hide their son when approached by the authorities are later discovered in their deceit. They are put to death. Even though these parents committed no moral wrong, indeed they acted in a morally praiseworthy way, the retributive nature of their punishment remains. We may quite coherently suppose here that their punishment was not instrumental with a view to deter, but genuinely retributive, as desert or repayment for daring to disobey the ruler's decree.

Such instances of punishment, retributive or otherwise, where there were no moral wrongs committed, are unjust. I shall argue in Chapter 2, that for punishment to be just, it must not violate certain basic principles of justice. One such principle is that punishment may not be imposed unless a moral wrong has been committed. Other principles of justice concern such issues as proportionality in punishment and not punishing the innocent.

In claiming that "retribution is done for a wrong," Nozick is assuming that retribution has built into it this principle of justice. That is, he assumes that retribution is always *just* retribution. In fact, as we have seen, retribution may be just or unjust depending on whether the principles of justice are observed. A similar situation applies in the case of revenge, indeed all forms of punishment. Nozick is correct when he says that "revenge may be done for an injury or harm or slight and need not be for a wrong." Revenge can be just or unjust, depending, again, on whether the principles of justice are observed, in this case whether the offence for which the revenge is taken is a moral wrong or not. This brings us back to the earlier point that Nozick's contrast between retribution and revenge is only apparent, obtained by building the principles of justice

into their notion of retribution while leaving the same principles out of their notion of revenge. In conclusion, this first comparison made by Nozick between retribution and revenge is indefensible.

But even if this contrast were defensible, nothing significant would follow from it regarding the moral status or justifiability of revenge. It certainly does not follow from it that there is something intrinsically wrong with revenge. The mere fact that it is *possible* to exact revenge for slights, harms and injuries which are not moral wrongs, need not prejudice in any way the possibility of morally acceptable and just forms of revenge. Moral condemnations of revenge for injuries, and harms which are not wrongs, because they were unintentional, perhaps, are quite compatible with revenge having moral approval in cases where it is a response to unambiguous and unjustifiable wrongs.

(2) Retribution sets an internal limit to the amount of the punishment, according to the seriousness of the wrong, whereas revenge internally need set no limit to what is inflicted. Revenge by its nature need set no limits, although the revenger may limit what he inflicts for external reasons.¹⁵²

Inflicting a penalty which is disproportionately greater than what is fair and appropriate is to commit an injustice. The notion of proportionality, the idea that the severity of the punishment should not exceed the seriousness of the offence, as I have already observed, is a principle of justice. While this notion is necessarily built into the very concept of just retribution, it is not, I should think, built into the concept of retribution as such. As with revenge, these limits are external to the notion of retribution. They are added onto it, rather than built into it, and it is here that Nozick goes wrong by departing from ordinary usage. For, to be sure, it is possible to overdo the punishment, not only through revenge, but also through vigilantism and even through institutionalised systems of punishment. More punishment can be inflicted than what, in all fairness, would be just in the circumstances. For example, many people would find unacceptable the stoning of women for adultery and the severing of hands for stealing, as has been the

¹⁵²Nozick 1981, p. 367

practice in some Islamic societies. Other, potentially more controversial, examples demonstrating the point are the criminalisation of alcohol during the prohibition years in the United States and the criminalisation of soft drugs, such as marijuana which, on the balance of evidence, do not seem to be more harmful than are tobacco and alcohol.

These considerations show that the concept of retribution is a wider one than Nozick is willing to allow, for both the concepts of just retribution and unjust retribution fall under it. My analysis which allows for both unjust and just retribution still to be regarded as retribution concurs more closely with the ordinary understanding of the term than does the explication put forward by Nozick and Ten. We may still regard, for example, the state's infliction of the death penalty upon a criminal as excessive punishment even if it were imposed for purely retributive reasons, such as repayment or desert for what the criminal has done. We would simply say that it was excessive and, therefore, to that degree unjust, retribution. Its retributive nature is by no means nullified by its excess. The explication put forward by Nozick and Ten is not only unable to account for the possibility that an excessive amount of punishment may be inflicted out of retribution, but leaves no room for discussion or dispute about the amount of punishment which morally can be exacted in the name of retribution. Their definition simply legislates that retribution is by nature fair and that it cannot be overdone, even by mistake. This is clearly not the case.

While on the topic of proportionality in punishment, it is worth mentioning that quite generally, in societies with revenge-based systems of punishment, there are clearly defined limits as to what is appropriate to inflict as punishment in retaliation for injuries and wrongs. This limit may vary from society to society, just as there is variation between the limits set by the institutionalised systems of various states and countries. Some of them sanction the death penalty for murder, for instance, whereas others do not. But what is significant here is that excessive acts of vengeance or acts which are considered unusually cruel or dehumanising are not acceptable. Avengers can meet with stiff moral and social disapproval from their community and "social life could become

painfully uncomfortable” if they inflict too much in retaliation, as judged by the established standards of their community.¹⁵³

These considerations keep paving the way for the thesis that revenge can also be just and that *just revenge* is nothing other than one form of just retribution, a form of retributive justice. This form of retributive justice is one in which the executor of the punishment is someone with a special tie to the victim of the original offence and who becomes involved because of personal retributive reasons but who, notwithstanding, does observe all the principles of justice in punishing the offender.

(3) Revenge is personal: “this is because of what you did to my -----” (self, father group, and so on). Whereas the agent of retribution need have no special or personal tie to the victim of the wrong for which he exacts retribution.¹⁵⁴

This comparison captures much of the essence of the difference between revenge proper and its genus, retribution. But, although I have already expressed agreement with this comparison, it will be useful at this point to consider in more detail the question of what constitutes a personal or special tie. In this regard two observations need to be made. First, this tie is relative in nature, in that, for some reason, such as familial, friendship or group membership ties, the avenger identifies more closely with the victim than the offender. For example, on one level of identification I can take revenge on someone, say P, for what P has done to me or my family, while it is also a possibility on another level of identification that I take revenge on someone else, say Q, for what Q has done to P, where I identify more closely with P than Q in virtue of the fact that P is a compatriot of mine, while Q is not.

The second observation is that this closer identification with the victim forms part of the reason for the avenger becoming involved in the punishment. This identification is

¹⁵³Boehm 1984, p. 70. See also the rest of Boehm’s Chapter 5.

¹⁵⁴Nozick 1981, p. 367.

part of what provides the necessary reason for the avenger making it his or her business to become involved in the imposition of the retribution in question.

(4) Revenge involves a particular emotional tone, pleasure in the suffering of another, while retribution either need involve no emotional tone, or involves another one, namely, pleasure at justice being done.¹⁵⁵

Before evaluating this comparison, a distinction needs to be made between two types of satisfactions that may be present when punishment is meted out. One is the sense of satisfaction which is felt by victims and those close to them when due punishment is imposed on their wrongdoers. This feeling of relief and satisfaction is not tied to any particular form of punishment. It can be experienced by victims regardless of whether due punishment is imposed on their wrongdoers privately by themselves or someone close to them (that is, revenge), by independent, non-legal third parties (that is, vigilantism), or by legally and institutionally appointed persons for the purpose (that is, the judicial system). The other type of satisfaction that may be present when punishment is meted out is a more impersonal, abstract satisfaction at justice being done which may be felt by anyone who cares about retributive justice, whether they be the victim, the executor of the punishment, or the community.

The feeling described by Nozick in the case of revenge as a “particular emotional tone, pleasure in the suffering of another” can be explained by reference to the fact that the victim of the crime is also the person (or someone close to the person) who is imposing the punishment. I argued earlier that the feeling in question is largely the satisfaction victims experience upon seeing justice being done by way of due punishment being imposed on the offender at whose hands they suffered. If this is right, Nozick’s comparison becomes vulnerable to the following two criticisms: The first is that he wrongly attributes to the avenger as the *executor* of the punishment a satisfaction the avenger feels as the *victim* of the wrong in question. The second criticism is that to the degree that the emotional tone he is referring to is indeed victim satisfaction, he is

¹⁵⁵Nozick 1981, p. 367.

mistaken in his claim that the emotion in question cannot be felt in retribution. Victim satisfaction can be as much present in judicial retribution as it can be in revenge.

The avenger may, of course, also feel some closely related satisfaction as the executor of the punishment on account of having been *personally* able to punish the wrongs done to him. But, once again, this satisfaction is felt by him because he is the *victim* (or someone close to the victim) of the offence or wrong in question. What is important to note here is that neither this particular feeling of satisfaction, nor the previously mentioned victim satisfaction, is perverse sadism to which revenge gives vent. This is another point on which Nozick's simplistic, misleading description of these satisfactions simply as "pleasure in the suffering of another" runs afoul.

All the same, Nozick is right in observing that administrators of, say, judicial retribution are not going to experience exactly the same feelings and satisfactions as avengers might do. The reason for this lies in the fact that while revenge is highly personal in the way already explained, judicial retribution is not. Since imposers of judicial retribution bear no personal or special tie to the victim, theirs can only be the abstract satisfaction of knowing that they have successfully imposed due retribution on a wrongdoer. They can also feel, of course, an even more abstract and impersonal form of this satisfaction, that of knowing that retributive justice was successfully imposed on a wrongdoer. But, as I already mentioned, this satisfaction can be experienced by anybody who values retributive justice in general.

By contrast, since executors of revenge are also the victims (or are closely related to the victims) of the wrongs concerned, they can, in addition, experience more personalised forms of this satisfaction: One of these is victim satisfaction, the satisfaction of knowing that retributive justice has been imposed on *their* own wrongdoers. The other one is the satisfaction of knowing that they themselves have successfully imposed due retribution on their own wrongdoer.

(5) There need be no generality in revenge. Not only is the revenger not committed to revenging any similar act done to anyone; he is not committed to avenging all done to himself. Whether he

seeks vengeance, or thinks it appropriate to do so, will depend upon how he feels at the time about the act of injury. Whereas the imposer of retribution, inflicting deserved punishment for a wrong, is committed to (the existence of some) general principles (*prima facie*) mandating punishment in other similar circumstances.¹⁵⁶

This comparison is both ambiguous and problematic. One reason for the ambiguity is that, strictly speaking, it does not live up to the requirements of a proper contrast. A basic requirement for contrasting two things is that they be compared in the same respect. If we want to contrast, say, apples and pears, then it is inappropriate to say that they are different because apples are red but pears are soft. To make a proper contrast between the two we would need to compare them either with respect to their hardness or to their colour. In the above comparison Nozick seems to abandon this basic rule. Revenge is claimed to lack generality because the revenger is not committed to *avenge* similar acts done to himself or to others while retribution is claimed to have generality because the executor of retribution is committed to the *existence of principles* which only give a *prima facie* mandate to punish similar acts. To get a proper contrast, the two must be compared either with respect to their commitment to *actions* of the appropriate sorts or with respect to their commitment to the *existence of principles* of the relevant sorts, but not in the way done by Nozick.

The ambiguity is made worse by some of Nozick's wording. In particular, the claim that the imposer of retribution is committed to the existence of certain principles makes no proper sense. If we are to talk sensibly about principles (or, for that matter, about anything in particular), then we either have to talk about their existence (and, by extension, about our belief or non-belief in their defensibility, validity, existence, etc), or about our commitment to (live by) them, but not about our commitment to their existence. What exactly am I committed to if I am committed to, say, the existence of the principle of truth telling? Am I committed to live by that principle and tell the truth in all things, or am I merely acknowledging that there is such a principle - a principle to which I do not

¹⁵⁶Nozick 1981, p. 368.

give significance in my life if it doesn't suit me, a principle to which I am not committed? As with Nozick's claim, it is not clear. Nozick's claims, therefore, need to be interpreted.

The most generous and plausible interpretation seems to be the following: The difference between retribution and revenge with respect to generality consists in the grounds on which their respective executors act. The avenger acts merely on inconstant, variable and unprincipled feelings about the act of injury at the time, whereas the executor of retribution acts on general principles which render permissible, or else obligatory, the imposition of retribution, not only in this instance, but also in all other relevantly similar circumstances.

This comparison, however, is problematic. For, although Nozick is right in pointing out that generally revenge has strong emotions or feelings associated with it, he is mistaken both in suggesting that these emotions are whimsical and in claiming that these emotions are the basis on which the avenger's actions must rest. I argued in section 1.2 above that the predominant emotion associated with revenge is the cognitively based feeling of resentment. As such, resentment does not simply rise and die in an unstructured, whimsical fashion. It arises in response to perceptions of injustice, offences against ourselves or those to whom we feel close. But if the feeling of resentment is anchored in the moral idea of injustice and if it arises in response to perceptions of injustice, then it is at least arguable that, therefore, (the existence or presence of) resentment can provide a general, consistent and legitimate moral ground for punishment in the form of revenge, (*prima facie*) mandating revenge in other similar circumstances.

This is at least arguable, but the avenger need not rely on it because he can do better. For, while a strong feeling of resentment might be associated with revenge, this feeling is no more of a justification for taking revenge than is the feeling of indignation experienced by imposers of judicial retribution in the case of legalistic, institutionalised punishment. The crucial principle which mandates punishment in relevantly similar

circumstances is the same for both revenge and judicial retribution. This is the general principle which renders permissible, or mandates, the retributive punishment of unjust and wrongful conduct in a way which does not violate the principles of justice.

What justifies retributive punishment is the fact that the person has committed a punishable moral wrong, a moral wrong for which she is responsible. Punishment is rendered permissible by this fact alone and not the presence of emotions of resentment and indignation which arise, in addition, in response to the wrongdoing. In other words, perceptions of moral wrongdoing give occasion to responses on two levels. On the practical level, the (*prima facie*) appropriate response is punishment, while on the emotional level the appropriate responses are resentment and indignation. I examine the link between wrongdoing and punishment in Chapter 6.

Even though the principle which permits the imposition of punishment is the same in both judicial retribution and revenge, its range of application is different in the two cases. In the case of judicial retribution the principle ranges over all members of the community. Revenge being personal in nature, the principle has a more limited scope, being restricted to offences against oneself and those closely connected. But that is quite different from the claim that there is no general principle that mandates revenge in other, relevantly similar circumstances.

What could be more plausibly argued, however, is that the imposer of, say, judicial retribution is morally *obliged*, as opposed to merely *permitted*, to abide by the principle of equality in justice, the principle that *like cases must be treated alike* and, therefore, that unlike the avenger, he is morally obliged always to give due weight to general principles which, *prima facie* mandate retribution in all relevantly similar circumstances.

But does this distinction really ground a genuine difference between retribution and revenge in the way required by the categorisation of the two suggested by Nozick? I don't think it does. For one thing, retributive punishment can be imposed not only through the judicial system but also through vigilantism. Similar to lone avengers, vigilante imposers of retribution are not required or obligated to impose fair retribution

for all relevantly similar wrongdoing. Not being part of an institution of punishment, their decision to punish wrongdoing in particular cases is voluntary, beyond the call of duty. Unlike imposers of judicial retribution, they are under no obligation to do so. If Batman or Superman bring a couple of criminals to justice today, their action is supererogatory and is a bonus for the community. But tomorrow, they can hang up their costumes and go fishing without being answerable to anyone for not doing the same as today.¹⁵⁷

Furthermore, with institutionalised forms of revenge, such as when social norms tightly regulate all aspects of revenge behaviour, or in cases where codes of revenge are legally institutionalised, revenge is often as mandatory in all relevantly similar circumstances as can be the imposition of retribution in the case of an institutionalised, judicial system. This point is illustrated by the following passage.

A man slow to kill his enemy was thought “disgraced” and was described as “low-class” and “bad.” Among the Highlanders he risked finding that other men had contemptuously come to sleep with his wife, his daughter could not marry into a “good” family and his son must marry a “bad” girl. ... He paid visits at his peril; his coffee cup was only half-filled, and before being handed to him it was passed under the host’s left arm, or even under his left leg, to remind him of his disgrace. He was often mocked openly.¹⁵⁸

Clearly, in communities with a revenge-based system of punishment, the social ostracism which can await those who fail to carry out their duty of revenge can be crippling. The pressure is enough to push into line even the most reluctant and it can be justified by reference to these communities’ reliance on everybody carrying their own share of the burden of punishment for purposes of retributive justice and deterrence.

In the light of these observations and discussions, it becomes evident that the contrast furnished here by the principle of like treatment of like cases is not between

¹⁵⁷I cover vigilante punishment in more detail in Chapter 3.

¹⁵⁸Hasluck 1954, pp. 231 - 32.

revenge and retribution as such, but between institutionalised and non-institutionalised forms of retribution.¹⁵⁹

(6) Retributive punishment is only inflicted [knowingly] on the wrongdoer, whereas revenge is sometimes inflicted [knowingly] on an innocent person close to the revengee, either because this is an easier target, or because it is thought that this would hurt the revengee more.¹⁶⁰

Revenge can be, and sometimes is, exacted by inflicting suffering on those close to the wrongdoer, although this is not a defining feature of it. The avenger may choose to inflict suffering on someone close to the offender, say, a relative, for two possible reasons. In the first instance, the avenger makes the offender suffer by making the relative suffer. It is interesting to note that in these cases the relative may not be *punished*, as such; their suffering is merely used as a *means* to an end, namely, the end of punishing the offender. Nonetheless, inflicting suffering on an innocent person in this way can be objected to both because it treats that person merely as a means to an end and because it is morally wrong to make an innocent person suffer as a means of punishing the guilty.

The other reason why the avenger may choose to inflict suffering on someone close to the offender, say, a relative, is not simply because, as Ten suggests, they are easier targets. Rather, it is because, given the prevailing norms regulating revenge, in many societies the avenger may not be required to distinguish sharply between the offender and their family or group. Hurting any member of the offender's group may be considered to be either as good as hurting the actual offender, or at least good enough. Consequently, the target of the punishment in these cases may not be just the individual who committed the offence, but the whole group (usually the family) with which the offender is identified. In many cultures all members of the group are held responsible for an offence committed by a member of that group because, as with most things, when it comes to punishment, people are identified with their larger groups. Their identities (in the

¹⁵⁹The institutionalisation of punishment will be covered in more detail in Chapter 3.

¹⁶⁰Ten 1987, p. 43.

epistemic, rather than the metaphysical sense) are very closely tied up with, and are largely derived from, those close to them - usually their families, friends, clan or tribe.¹⁶¹ Their moral responsibilities, that is, their liabilities for praise and blame, reward and punishment are inextricably tied to this extended sense of identity. This basically means that it is quite legitimate and proper for any member of the wronged family to direct the retaliation towards any member of the wrongdoer's family. Feuds, many of which are completely bloodless, are good examples where cycles of revenge and counter revenge take place between families, not individual people. A feud is made possible primarily in contexts of shared identity and responsibility between family members. Wrongs and their punishment were matters between families and clans, not merely between individuals.

Insofar as it is this kind of context that lends initial plausibility to Ten's observation about revenge, it should be easy to see that Ten's comparison depends on an individualistic paradigm of responsibility which cannot explain, let alone do justice to, the phenomenon under consideration. So-called revenge on the innocent is not explained by a lack of concern for ethical principle, but by a wider notion of responsibility according to which any member of the offending group constitutes a legitimate target for revenge. Patterson explains this beautifully.

Whereas Pakeha [Europeans] tend to concentrate upon an individual actor, Maori tend to see the actor as a family or other tribal group. In Pakeha terms, I insult you and so your uncle obtains utu [repayment] by insulting my elder brother. In Maori terms, my family insults yours and so your family obtains utu by insulting my family. So, seeing the situation in Maori terms, it turns out that it is an ethical requirement in the case of utu, just as it is in the case of punishment, that the recipient be the offender. It would definitely be a mistake, according to Maori values, to inflict utu upon the wrong tribal grouping, just as it is a mistake, according to Pakeha values, to punish scapegoats.¹⁶²

¹⁶¹The notion of epistemic identity is discussed in more detail in Chapter 5 in connection with the importance of honour in such cultures.

¹⁶²Patterson 1992a, p. 130.

Ten's corresponding claim about retributive punishment, that it is knowingly inflicted only on the wrongdoer, may be true of some judicial systems, but this is not a necessary feature of retributive punishment itself. Judicial retributive punishment can be as unjust in this respect as revenge can be. Unknowing company directors or shareholders, for example, can be punished for the wrongdoings of employees.¹⁶³ We can also imagine situations where the legislators determine that, once a person is found guilty of an offence, the punishment which is imposed for retributive reasons must be shared among immediate family members. As a matter of fact, something like this was the practice when it came to punishing disgraced samurai in medieval Japan,¹⁶⁴ and also among the Maori of New Zealand in the form of *muru* (plunder) and *utu* (repayment).

A standard method of resolving a dispute in traditional Maori law is the practice of *muru*. In this practice, an aggrieved family may directly confront the group that has caused their trouble, and confiscate goods in payment for the injury.¹⁶⁵

If, for example, I insult you, that calls for *utu*. ... The responsibility for taking *utu* falls not on the individual but on the whole family. Likewise, the responsibility for the insulting behaviour is borne not only by the individual who delivers the insult but by a whole family group. A suitable retaliatory insult could be directed against any or all members of the family of the offender.¹⁶⁶

Again, from an individualistic mode of thinking it might be tempting to say that such punishment is unjust, since one person cannot be held responsible for the actions of another, no matter how closely they might be related. But, apart from being questionable,¹⁶⁷ an attack along these lines would be beside the point in the current

¹⁶³John Braithwaite drew this to my attention.

¹⁶⁴Benedict 1977.

¹⁶⁵Patterson 1992a, p. 140.

¹⁶⁶Patterson 1992b, p. 15.

¹⁶⁷For a defence of collective responsibility in Maori tribal contexts, see Perrett 1992. Perrett shows that the Maori concept of collective responsibility is not unlike the collective responsibility borne by corporations - an account of which has been articulated and defended, for example, by French (French 1984). It is also important to note that "The Maori view recognises both the tribe and its individual members as moral subjects. Thus, in the collective's external relations (as in intertribal matters) *utu* is exacted at the collective level, rather than simply directed at the individual offender. However, in the collective's internal relations the individual offender is not treated as if he had no individual

context. For even if it were unjust, that would not change the retributive character of the punishment itself, which consists in the reasons proffered for its imposition, that is, that it is imposed as desert or repayment. Like Nozick, Ten in his comparison of retribution and revenge has tried, fallaciously, to build principles of justice into the concept of retribution but not into the concept of revenge. To get a real contrast between retribution and revenge, the principles of justice will have to be built either into both of them, or into neither of them.

It has been suggested to me that the disagreement I have with Nozick and Ten is mostly verbal, rather than substantial.¹⁶⁸ This, however, is not the case. I disagree with them on just about every important point that concerns retribution and revenge. First, I disagree with them over what the correct analysis of revenge is, as well as over the question of how revenge is related to retribution. Through their comparisons, they exclude the possibility that revenge might be a kind of retribution. They also exclude the possibility that retribution and revenge might be more closely related than are two species of some genus. In my view these are mistakes. Revenge is essentially retributive in character. The core justificatory reason behind revenge is the essentially retributive reasoning of payback, repayment and ill-desert. But since, in addition, revenge is personal in a way that other forms of retribution need not be, the correct way to analyse the relation of the two is in the form of a genus-species categorisation. A more thorough analysis and discussion of these points will be undertaken in the following chapter.

Furthermore, Nozick and Ten give a distorted analysis of retribution by failing to distinguish it from what are properly regarded to be principles of justice. Nozick and

responsibility: hence the practice of *muru*." (Perrett 1992, p. 37.) Obviously, the Maori used *muru* (plunder) to punish groups, as well as individuals.)

Note also that one of the central criticisms made by Moana Jackson of the criminal justice system of New Zealand is its refusal to accept the concept of collective responsibility. It is hard to disagree with Jackson that the resultant alienation of offenders from their families and victims has disastrous results in terms of rehabilitation and reintegration. (Jackson 1988.)

¹⁶⁸This suggestion was made to me, for instance, by Philip Pettit.

Ten unacceptably restrict the scope of the term 'retribution' by conceptually tying it to principles of justice. The result is an implausible account according to which retribution is necessarily for a wrong, it is necessarily proportional, it necessarily gets the right person, and so on. This is mistaken because it is possible to impose unjust retribution. Retribution can be overdone, it can get the wrong person and, as the example of whistleblowers illustrates, it can be imposed for offences which are not moral wrongs. The correct analysis of retribution, therefore, must be one in which retribution is independent of the principles of justice. I offer such an analysis in the next chapter.

2.5 Conclusion

In this chapter I clarified the notion of retribution and considered its relationship to the principles of justice and revenge. First I considered the concept of retribution and argued that the best way to define it is by reference to the notion of negative repayment. I then showed that principles of justice which set the limits of just punishment are not retributive in character. The thesis of moral retributivism, the view that retribution within limits set by the principles of justice provides a morally defensible ground for punishment, is considered in Chapter 6. In the final main section of this chapter I focused on the relationship between retribution and revenge. I considered the two rival views on this issue, namely the view that revenge and retribution are identical and the view that the two fall into quite different moral categories. I argued that both views are unsatisfactory. Yet, the two concepts are significantly related, as is evidenced by their frequent, albeit mistaken, identification. Therefore, an analysis of revenge would be incomplete if its relationship to retribution was not clarified. This is one of the main tasks of the following chapter, in which I present a map of the logical space of punishment with a view to locating revenge on it.

3

Revenge: definition and domain

3.1 Introduction

In this chapter I consider the logical space of punishment with a view to defining revenge and to giving a visual representation of its relation to other key punishment concepts. I define revenge as personal retribution which is typically accompanied by feelings of indignation, anger and resentment for wrongs suffered in one's personal domain of concern. The map I intend to give consists of two parts: a *typology* of the various reasons for punishing and, secondly, a categorisation of the various *forms* of punishment on the basis of two clearly identified criteria. But it is important that neither of these criteria be confused with what I call respectively the *modalities* and the *process* of punishment.

In the following two sections I discuss briefly these modalities and the process of punishment. I then proceed in the fourth section to draw a map of the logical space of punishment. In the fifth section I locate revenge on that map.

3.2 The modalities of punishment

There are as many modalities of punishment as there are ways of imposing penalties and deprivations. Imprisonment, fines, shame, embarrassment, deprivations of various sorts such as a negative vote in the next departmental meeting, or arriving late for dinner or some other arrangement, are all ways to impose punishment on someone.

These can be conveniently labelled as the “modalities” or the “manner” of punishment, as they are modes and manners in which punishments occur.

There are two points worth mentioning about the modalities of punishment. The first is that there is room for constructing a rich typology of the possible modalities. For example, some broad categories are suggested by the fact that punishment can be physical, deprivatory of freedom, economical, psychological, sexual, etc. Possible categories under, say, punishment by deprivation of freedom might be concentration camps, labour camps, standard imprisonment, solitary confinement, house arrest, exile, etc. Lists of the categories and sub-categories could go on almost indefinitely, since there are at least as many ways of imposing punishment on someone as there are ways of penalising them.

The second point is that the modality of punishment is a legitimate moral concern. There are two dimensions to this, one of which relates to the offender, the other to the victim. With regard to the offender, there is a principle of justice the violation of which seems automatically to constitute a serious moral wrong. This principle is that, regardless of the offence, punishment should not be imposed in grotesque, dehumanising modes. For example, intentional physical and psychological torture, mutilation and disfigurement are dehumanising to those punished and are, on that account alone, serious moral wrongs. Not being punished in dehumanising ways could be regarded quite plausibly as a basic human right. A difficulty here might lie in the fact that there can be radical disagreement about what kinds of punishment would count as dehumanising. Indigenous people such as Aboriginal Australians and the Maori of New Zealand, for example, consider imprisonment to be dehumanising and cruel, whereas Europeans perceive the spearing of the offender’s thigh, for instance, to be cruel and barbaric. Western culture also condemns the cutting off of people’s hands for stealing, to use another example, as evidenced by strong reactions in the West to reports that the practice

has been revived in at least one Islamic country.¹⁶⁹ Such culture-specific differences may be irreconcilable because they are often a reflection of the prevailing values and the way of life of a people. Interfering with these, in turn, creates moral problems of its own because it raises questions about political autonomy and about rights to national, or cultural, self-determination. However, discussion of such difficult but important problems falls beyond the scope of the current project.

With regard to the dimension relating to the victim, there is a good moral case for choosing the manner of the punishment in such a way that it not only compensates the victims for their loss but also maximises their restoration from the trauma of having been victimised. This might mean that, whenever possible, the offender ought to be given a chance to express regret and sincere apology directly to the victim for their behaviour and pay the victim compensation as a constructive way to make amends. In turn, this requires that those punished should be kept economically productive whenever possible. Part of their earnings might also be used to cover the costs associated with their punishment.

3.3 The process of punishment

A somewhat different dimension of punishment from the modality is the process the punishment takes. Invariably, punishment in any of its modes is preceded by certain deliberative processes. These can be very short and simple, or long and involved. Using the judicial model as a benchmark, four distinctive stages of the process of punishment seem to be the following: The *gathering of evidence*, leading to the *plea* and/or verdict, which is the determination of whether the accused is guilty or not of the alleged offence, the *sentencing*, which is the determination of both the appropriate mode and the amount of punishment, if the accused is found guilty and, finally, the *execution of the sentence*, which is the punishment itself.

¹⁶⁹Cockburn, P: Mutilation for Iraqi thieves, deserters. *The Sydney Morning Herald*, Saturday, January 14, 1995. (p. 15.)

The process of punishment is also a moral concern, again, for both the criminal and the victim. It would be wrong to accuse and punish someone even if they happened to be guilty without giving them a fair hearing, without giving them a decent opportunity to explain and defend themselves. Insisting on extremely elaborate processes regardless of the circumstances may well be unnecessary, if not undesirable. But the absolute minimum surely has to be that, insofar as possible, the process must be such that it safeguards against abuses and violations of what may be regarded as basic human rights and operates according to the principles of justice outlined in Chapter 2. As with dehumanising modes of punishment, there seems to be a good case here for not being punished without due process. But the process of punishment can be a moral concern for the victims as well. Their rehabilitation could often be greatly helped if they were given some sort of place and say in the process of punishment. It has a great potential to empower them to feel again in control of their disrupted lives. I return to this point again in a later chapter on the morality of revenge.

3.4 The logical space of punishment

3.4.1 A motive-based typology of punishment

This typology relies on motive-based justifications behind punishment. A fine and a prison sentence are just that, a fine and a prison sentence, and whether they constitute retribution, or deterrence, or simply the community's way of denouncing the wrong and expressing its disapproval for it, depends on internalist, motive-related reasons, justifications and rationales for their imposition. Whether some particular punishment falls into one or the other of the above categories cannot depend on externalist considerations such as the time, place or circumstance of the offence, or indeed the punishment. Thus, when punishment is imposed in order to reform, or to deter future would-be offenders, the punishment is clearly instrumental. By contrast, when punishment is imposed as repayment (payback) or a kind of negative desert for some wrong or offence, the punishment in question is recognisably retributive in nature.

However, when punishment is imposed as an expression of disapproval, or as a symbolic affirmation of society's fundamental values, it seems to have a strongly functionalist character. The most familiar reasons for punishment can be listed and broadly categorised as indicated below.

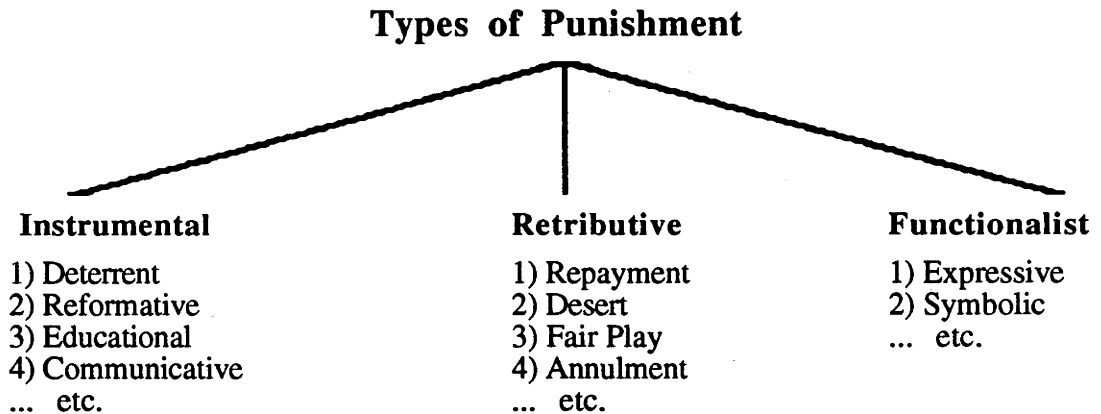


Figure 1

I take it to be uncontroversial that the various reasons for punishment, the natures of which underlie this typology, are not incompatible with one another and several of them can be appealed to in explanations and justifications of punishment. Thus, there is no difficulty in imposing punishment for the combined reasons that it is due as repayment or desert, that it deters future would-be offenders and that it expresses disapproval towards the offensive conduct.

It is also of some importance to note the following contrast between instrumental and retributive punishment. Instrumental aims, such as deterrence and reformation, are linked to punishment only contingently and circumstantially. It is possible, if not probable, that deterrence and reformation could be better achieved quite often through other means than punishing. Deterrence, for example, might be better achieved through a reward system for good behaviour, while reformation would certainly be better served by therapy and friendly education. This does not hold true, however, for retribution. In the case of retribution, there is a necessary link between the retributive aim of the punishment and the punishment itself. Any response to criminal activity would necessarily fail to achieve or actualise such retributive aims as giving the offenders their negative deserts

and repaying them in kind for the wrong they have done, if the response in question was not punishment of some sort. Retribution is impossible without the imposition of some penalty, pain or unpleasantness for the wrong or offence in question.

It should be further noted that not all of the reasons for punishment invoked by the above list have to function as justifying reasons. For example, to say that some punishment was imposed as payback for a prior offence can easily be a mere indication or explanation of the motive, rather than an exercise in justification. A full justification of punishment as negative repayment certainly requires a more detailed account of *why* this kind of repayment should be morally desirable or acceptable. Even though I argued in the previous chapter that a more detailed account of such justificatory claims is possible, it remains important to base typologies of punishment, such as the above, solely on the explanatory functions of the reasons concerned. This is important for at least two reasons, both of which have to do with retribution. One of these reasons is that a person imposing punishment, say, as repayment prescribed by law, social norms and customs, might not be able to formulate the proper justifying reasons for the imposition of punishment, but that should not present an obstacle to identifying the punishment in question for what it recognisably is: law-, or norm-driven retribution. The other reason why such a typology is preferable has to do with a particular philosophical position about the possibility of justifying retribution. According to J. L. Mackie, for example, it is misguided to look for a justification of retribution. There is no justification for it, but only a socio-biological explanation.¹⁷⁰ While I shall argue against Mackie on this point in the last chapter, his position on retribution is coherent. Yet, a typology of punishment based on justificatory reasons alone could not properly acknowledge the existence of retribution, while allowing at the same time that a position such as Mackie's is coherent.

¹⁷⁰Mackie's position with regard to the justification of retribution will be discussed in the last chapter.

3.4.2 The basic forms of punishment

Once it is established that there are good reasons for punishing someone, it becomes important to decide two questions. One is the question of who has the right (and, possibly, the duty) to be involved in the various major stages of the process of punishment. The other question concerns the form such involvement should take. I will now provide a useful framework for later discussions of these two questions, in Chapter 6. A more immediate role of the framework in question is to aid in locating revenge in the logical space of punishment. The framework is based on two criteria. The first criterion is whether or not the process of punishment is institutionalised. By an institution I mean an established law, custom or practice. An obvious example of institutionalised punishment is our current judicial system where the process is undertaken by neutral third parties. An example of non-institutionalised punishment could be, say, where a wife divorces her husband upon discovering his infidelities. To the extent that an institution punishes people retributively - for, obviously, it does not have to - to that degree the punishment exacted by the institution in question is *institutionalised retributive* punishment. To the degree that the institution punishes people instrumentally, the punishment is *institutionalised instrumental* punishment.

The second important criterion is whether punishment is imposed for a personal reason or not. A necessary, though not a sufficient, condition for this is the existence of a personal tie between the victim of the offence for which punishment is being imposed and the person who is involved in the process of punishment. A personal tie is conceptually presupposed by the notion of a personal reason and, therefore, it is necessary for it. Very briefly, someone has a personal reason for getting involved in the process of punishment if that person gets involved *because* of the personal tie s/he has with the victim. Typically, the involvement of such a person is going to be accompanied

by a feeling of resentment on account of being personally affected by the wrong or offence in question.¹⁷¹

These two criteria, that is, the institutionalisation of the process of punishment and the existence of a personal reason for one's involvement in that process, provide the basis for four basic forms of punishment. The way in which they are individuated can be easily seen in the following Figure (where 'I' stands for 'Institutionalised', 'P' for 'Personal' and '~' for 'Not'):

	Institutionalised	Non-institutionalised
Personal	I & P	~I & P
Non-Personal	I & ~P	~I & ~P

Figure 2

I shall run through these in order of familiarity. *Institutionalised non-personal punishment* is punishment imposed by a person or a body of persons who are specifically entrusted (with or without pay) with the task of ensuring fair, equitable, and systematic ways of punishing people for their wrongdoings. In this form of punishment the punishment is completely taken out of the hands of the wronged party. The victim(s) has no power in deciding over any of the major stages of the process of punishment. Clearly, it is this form of punishment that is exemplified in our courts within the framework of the judicial system.

Non-institutionalised personal punishment is punishment imposed by a person (or a body of persons) who has a personal reason for being involved in the process of punishment. This personal reason is that they have a personal or special tie to the victim of the offence or wrong. In this form of punishment the punishment is completely in the

¹⁷¹I have already discussed the moral status of resentment, arguing that it is a kind of moral indignation over perceived injustice or unfairness done to oneself or those close to oneself.

hands of the wronged person and those who are personally related to him/her rather than in the hands of an institution set up for the purpose. The offended side has all the power to decide over guilt and degree of responsibility, the amount and type of punishment, as well as the execution of the punishment.

Non-institutionalised non-personal punishment stands in contradistinction to both institutionalised and personal punishment, as well as to their combination. It is both useful and correct to refer to this form of punishment as *vigilantism*. Vigilante punishment is imposed by a person or a body of persons who are neither specifically appointed to punish some wrongdoing nor do they have the above mentioned personal reason for taking on the task themselves. Typically, in vigilante punishment the infliction of suffering is carried out by persons who are personally unrelated to either party. Vigilantes, in a sense, are self-appointed agents who take up the task, most often in the name of fairness, justice and deterrence. Most often these agents take it on themselves in times of need to see to it that wrongdoing is appropriately punished and that life and property are protected against criminal activity. They take on the responsibility of supplying the deterrence and the retribution if in the circumstances these cannot be ensured either privately by the wronged parties or institutionally because of the inefficiency, or the total absence, of such a system. Sometimes, however, vigilante involvement may only be partial. Fictitious vigilante figures like Batman and Superman are well known. Real life examples are the Vigilantes of Montana¹⁷² and the Guardian Angels of the New York City subways.¹⁷³

Finally, *institutionalised personal punishment* is a cross between the first two basic forms. In this form of punishment the victim(s) of the alleged wrongs and offences become involved in the process of punishment within an institutionalised framework. Viking law, revenge practices of Medieval Japan, and the earlier mentioned Maori

¹⁷²Dimsdale 1866.

¹⁷³Kenney 1987.

practices of *murū* (plunder) and *utu* (repayment) are good historical examples.¹⁷⁴ Notwithstanding the abundance of historical precedents, until very recently, substantial victim involvement has been virtually unknown in contemporary Western systems of institutionalised criminal justice. However, there are good reasons for giving it serious consideration and I shall return to consider personal forms of punishment in the last chapter where I argue the merits of institutionalised revenge. I move on now to represent the above mentioned four basic forms of punishment in logical space. One way of doing this is as follows:

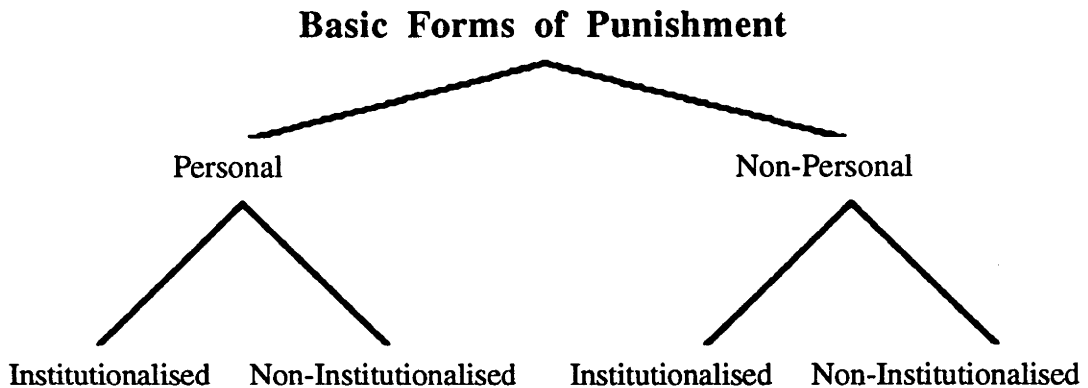


Figure 3

These four basic forms of punishment are pure forms and there is scope for mixed cases. There certainly is scope for degrees of institutionalisation with respect to the process of punishment. There may also be scope, perhaps, for ambiguity in certain cases as to the presence of a personal reason for becoming involved in the process of punishment and, for better or worse, it is for guarding against this reason being present in judicial proceedings that current institutionalised systems are set up in such a way as to exclude the victims and those close to them from having any role of significance in the process of punishment. This much said, the stage is ready to identify revenge on a combined map of the logical space of punishment.

¹⁷⁴ Friedman 1979; Foote 1970; Miller 1990; Benedict 1977; Patterson 1992a; Patterson 1992b.

3.5 Revenge: definition and domain

The taxonomy of punishment I presented in the previous section is conceptually and terminologically independent from any ideas of revenge. In this section I provide reasons for identifying certain kinds of punishment as revenge. The strategy here, of course, will have to consist in pointing out the conditions which seem to be individually necessary and jointly sufficient for revenge. If revenge is to be found within the framework I have set up so far, these conditions and features should already figure on the relevant maps of the logical space of punishment that was drawn up earlier. Indeed, these features must be there, if the arguments I presented in the first chapter for the claim that revenge is a form of punishment were correct.

In order to locate revenge on that map, however, we need to take into consideration both the typology of the various specific concepts of punishment and the typology of the forms of punishment. Starting with the typology of the concepts, the first clue is the fact that revenge is essentially retributive in character. Revenge is motivated by the desire to pay back the other, to even the scores, to impose on the other what they deserve in the retributive sense described in Chapter 2. If an act of punishing someone is not done at least in part out of such retributive reasons, be that by reference to repayment, desert or whatever, then it cannot be a case of revenge. We cannot properly talk of revenge in cases where the punishment is purely instrumental. Take the example of the parent who punishes the child purely to discourage the child from playing a dangerous game, even if that game has already hurt the parent. That sort of punishment cannot be taken to be a case of revenge by any stretch of the imagination and I suggest that this is because the punishment is purely instrumental and not at all retributive in character.

But, as I have already argued against Solomon and Oldenquist, not all cases of retribution are cases of revenge and, therefore, we need to look further. More specifically, what needs to be done here is to capture or accommodate the personal nature of revenge in the sense already explained. This feature figures quite prominently, of course, in the typology of the basic forms of punishment and so locating revenge on it is

a straightforward matter. Combining the two typologies, the domain of revenge can be located in the logical space of punishment in the following way:

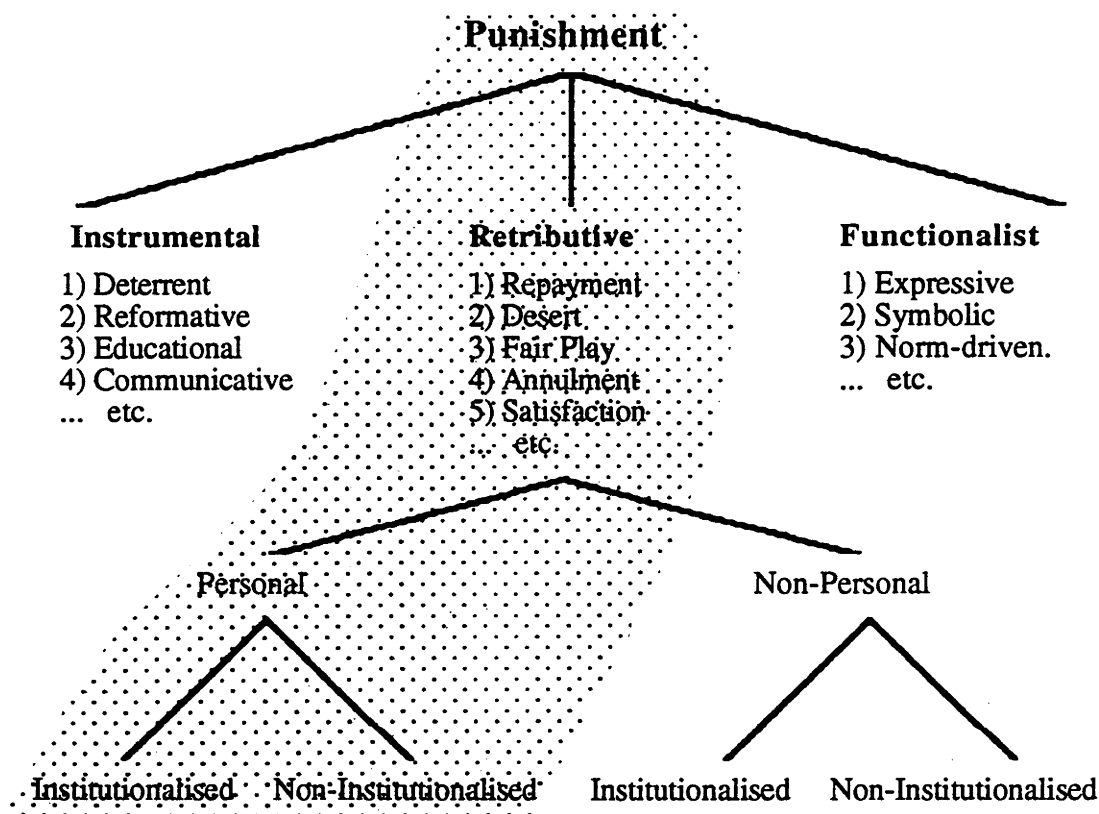


Figure 4

The necessity of a personal reason for revenge selects two and excludes the other two of the four basic forms of retributive punishment as being possible forms of revenge - save by way of being possible instruments for someone's revenge. *Non-Institutionalised Personal Retributive Punishment* is what I take to be the thing we commonly refer to by the word 'revenge.' But, clearly, this is not the only form of retribution to which the term 'revenge' properly applies. For, although the phenomenon had been until recent years virtually absent in industrialised Western cultures, *Institutionalised Personal Retributive Punishment* has not been totally unknown as a practice and this is also a form of revenge. In medieval Iceland and Japan, for example, individual intentions to take revenge often had to be first registered with, and normally

approved by, higher authorities. State sanctions applied for failing to do so.¹⁷⁵ This is what might be called *Institutionalised Revenge*. The remaining two categories, of course, are the familiar judicial punishment and vigilantism. In conclusion, the place occupied by revenge in the logical space of punishment corresponds to the following definition of revenge: *Revenge is personal retributive punishment*. That is, revenge is a form of punishment which is imposed for retributivist personal reasons to satisfy the requirements of, and a corresponding personal need for, retributive justice for wrongs suffered in one's sphere of care and concern.

While this definition captures quite well the necessary and sufficient conditions for revenge, it is important to bear in mind that the retributive element in revenge remains compatible with consequentialist considerations of one's interests. For example, someone could adopt a policy of personally paying back for offences only when the costs of doing so were not too high. Indeed, if one's decision to take revenge is to be rational, its expected overall costs cannot be allowed to outweigh its expected benefits. The status of revenge from the viewpoint of rationality will be argued in detail in the next two chapters. What is important to note here is that the compatibility between retributive and instrumental considerations goes so far as there being many cases of revenge which seem to be largely instrumental, their retributive element fading considerably to the point of becoming secondary in the immediate, direct reasoning behind the decision to take revenge.

Many pre-industrial societies have, or had, an honour-based system of morality and norms of revenge frequently occupy an important place in their respective codes of propriety and honour. As I shall argue in more detail in the following chapter, for an individual living in such an honour-centred culture, honour, good reputation, and a good name were precious, if not vital. Importantly, honour and good name were in jeopardy if one failed to live up to society's norms and customs - the relevant ones here being those

¹⁷⁵Footnote 1970; Gragas 1980; Benedict 1977.

which regulated the appropriate occasions for, and other aspects of, revenge. Although the codes of honour and the norms of revenge would generally allow a limited number of alternatives, such as accepting an apology or a negotiated settlement, in special circumstances they could conspire to leave the individual with no viable alternative to revenge, but only the loss of face, disgrace, becoming the subject of ridicule, of gossip and social ostracism. Especially in such cases, but also more generally where honour and the individual's, as well as the family's, good names were on the line, there was a considerable shift in focus and emphasis away from the importance of retribution pure and simple towards the importance of maintaining the honour and good reputation of the individual and those of the whole family. It is not difficult to see how in certain circumstances this shift in focus could be so marked that the individual seeking revenge hardly considered the underlying point and rationale for revenge (which is retribution) and made his decision by reference to the relevant retributive norms and codes alone.

The blood feud is a good example of revenge where individuals and families are primarily moved by concerns to maintain honour, and avoid the disapproval of the community which, as I show in the next chapter, can be crippling. The blood feud being one of the severest forms of revenge, it might be wondered whether retribution really does form part of the concept of revenge. The answer to this question is affirmative. The rationale of revenge is retribution even in cases where instrumental considerations of honour and social standing figure as prominently in the avenger's reasoning as they do in the feud. The retributive idea of negative reciprocity, or repayment (payback) is an essential element in any satisfactory analysis of the feud, and so is the idea that retaliation is due, and is justified, on the basis of the responsibility the targeted person bears for the offence to be avenged. In other words, what justifies any particular case of revenge in the course of a feud is not the preservation of honour and social standing, but the retributive liability of the targeted person (or family) for some prior offence which is based solely on their moral responsibility (sometimes simply by association) for the same.

While considerations of honour and disgrace, of social standing and sanctioning form a very important part of the feud, their role is not a justificatory one. Though the concept of honour is essential to understanding fully the feud, the desire to preserve honour and social standing justify neither the feud as a whole, nor the individual cases of revenge within it. Rather, these instrumental considerations are part of a powerful mechanism of social control which reinforces the value and the importance of retribution in moral communities where, often by choice, there is no centralised system of punishment. These communities evolved alternative, largely revenge-based, systems of punishment by requiring that everybody pull his or her own weight within his or her own spheres of concern and responsibility. Thus, in moral communities with no centralised system of punishment, but which, nevertheless, want to hold their own members responsible for improper, wrongful conduct, the two critical elements of revenge are explained in the following way: The fact that they are moral communities explains their insistence that offenders be retributively punished for their wrongful behaviour, while the absence of a centralised system of punishment explains their insistence that this be done personally by the offended parties concerned. Those who fail to carry out their moral duties in these matters are punished by the community as a whole through gossip, open ridicule and other sanctions.

The earlier quoted passage from Hasluck illustrates clearly the unbearable social ostracism which can await a person who refuses to take revenge.¹⁷⁶ Such sanctions make it individually rational even for the faint-hearted to carry out the duty of revenge - a consideration which I will discuss in more detail in Chapter 4. What is most important to keep in mind at this point is that it is retribution, based on moral responsibility, which provides the justification, the rationale not only for revenge, but also for the reinforcing social mechanisms encompassing the relevant codes of honour, norms of revenge, the gossip and the real threat of severe social sanctioning of those reluctant to discharge their

¹⁷⁶See quote in Section 2.4.2. (Hasluck 1954, pp. 231 - 32.)

duties. These mechanisms of social control have evolved in cultural contexts so that it becomes individually rational for everybody to live by the law of retribution even if, and when, their other interests would tempt them to do otherwise.

Such mechanisms of social control are not unique to the area of punishment. Analogous ones which control other aspects of people's behaviour often operate in traditional, conservative societies. For example, in traditional Montenegro "in choices such as hospitality versus stinginess, honesty versus overuse of deceit in casual transactions between people, or sexual transgressions versus sexual propriety, gossip was the primary sanctioning mechanism."¹⁷⁷ Similarly, "in the segmentary society of the Bedouins of the Western Desert, which have only recently been submitted to a strong governmental authority, honour has played, and still plays, a most vital role in the field of social control, besides providing an effective code of morality."¹⁷⁸

The existence of such mechanisms of social control also explains how overdetermination of certain behaviour can take place. To the marital difficulties which would be brought about by extramarital sex, for example, people are given an additional, equally powerful reason to resist temptation by the gossip and the scorn with which society would regard them. Similarly with revenge: in addition to internal feelings of resentment and considerations of retribution which often suffice for revenge, immense social pressure of the kind described above can give at the same time an additional, often weightier, reason to carry out one's duty of revenge. I will return to this point again in Chapter 4.

3.6 Conclusion

I began this chapter by discussing the modalities and process of punishment, pointing out the moral issues in both areas. I then went on to prepare the ground further

¹⁷⁷Boehm 1984, p. 84.

¹⁷⁸Zeid 1965, p. 259.

for a definition of revenge, beginning with a categorisation of the reasons for punishment. I showed that, whatever the reasons for punishment, the process of punishment can occur in institutionalised as well as non-institutionalised settings and that punishment can be administered with personal or non-personal involvement of the wronged persons in question. I argued in the fifth section that, in relation to reasons for punishment, revenge is retributive and that, with respect to the forms of punishment, revenge is personal. I concluded that revenge is personal retributive punishment and that it can take institutionalised, as well as non-institutionalised forms. My defence of revenge relies on this distinction in that it involves a significant shift in focus from non-institutionalised to institutionalised forms of personal retributive punishment.

Following this, I argued that the retributive element in revenge remains compatible with consequentialist considerations of one's interests. I drew attention to the fact that in certain cases, such as the feud, instrumental considerations of honour and social standing can become the primary considerations in the direct reasoning leading to a decision to take revenge. However, I also pointed out that this phenomenon of strong instrumental focus does not disconfirm my analysis of revenge in terms of retribution because the rationale of revenge is retribution even in such cases. The retributive idea of negative reciprocity, or repayment (payback) is an essential element in any satisfactory analysis of the feud because the most basic point of justification for retaliation is not the need to maintain honour and social standing but the retrospective responsibility the targeted person bears for the offence in question and, correlatively, a deontologically based moral duty of the offended party to materialise such retrospective responsibility through appropriate retaliation. Honour acts merely as an additional stake and serves to reinforce the importance of such responsibilities and duties.

With the nature and definition of revenge now clearly formulated and its relationship to other important elements of punishment clarified, it is time to consider its normative defensibility from rational and moral perspectives.

4

Rationality and the collective rationality of revenge

4.1 Introduction

Notwithstanding the strong anti-revenge climate in many contemporary societies, no one seems to have argued against the rationality of revenge in any serious detail, except for Jon Elster who asks: “Why would a rational person ever carry out an act of revenge? Why would anyone believe a rational person if he threatened to do so?”¹⁷⁹ Elster argues that there is no defensible answer to these questions, that revenge cannot be given a satisfactory rational choice explanation and that it cannot be individually rational. In this chapter, and the following one, I propose to argue the opposite. But, since no one besides Elster seems to have argued specifically against the rationality of revenge, I will be mostly using Elster’s comments and arguments as a springboard for a thorough treatment of the topic. This will often require going beyond his narrow focus on rationality as understood solely in terms of what is known as ‘rational choice theory.’

To be sure, I do not intend to defend the view that revenge is somehow inherently rational, that it is rational by its very nature. There would be very few human endeavours, if any, that would have that character. My aim is to argue for the position that, as with most other undertakings, revenge can be rational or irrational, depending on the circumstances, and that sometimes the best way of giving a full account of some act

¹⁷⁹Elster 1990, p. 872.

of revenge may well be by means of an appeal to rational choice explanations. By implication, as far as its rationality is concerned, each act of vengeance has to be evaluated on its own merits.

I proceed as follows. In the next section, I consider why it is important for human behaviour, including revenge behaviour, to meet certain standards of rationality. I also distinguish between different forms of rationality: direct and indirect; formal and substantive, and show how these distinctions bear on the rationality of revenge behaviour to be discussed in the rest of this chapter and the next. In the last main section of this chapter, I discuss the collective rationality of group revenge and the collective rationality of revenge practices as institutions. In the course of these arguments I present and discuss group revenge among the Maori, as practiced in pre-colonial New Zealand. The individual rationality of revenge is argued for in the next chapter.

4.2 Rationality

4.2.1 The significance of rationality

In this sub-section I consider the importance of rationality for human behaviour in general and for revenge in particular. I also consider limitations on the applicability of rational choice theory, as well as the morally desirable limits on the rational pursuit of individual interests.

According to the standard (OED) definition, a rational person is one endowed with reason, is agreeable to reason, is sane, intelligent and judicious. Accordingly, a rational person makes reasoned assessment a basis for action. This sounds plausible and accurate in so far as the ordinary use of the term 'rational' goes. Accordingly, the basic sense embodied in this reportive definition should be taken into account in assessing the proper scope and merit of any theory of rational choice attempting to give a more detailed account of what this attribute called 'rationality' is. Such a theory can be criticised for being too narrow if it fails to explain actions which are rational in this pre-theoretical, but

plausible and intuitively correct, sense. It can be similarly criticised for being too wide if it deems rational those actions which are not rational in this pre-theoretical and intuitively correct sense. But a wider theory of rational *behaviour* may have less stringent conditions of rationality than would a theory of direct rational choice. I discuss the details and significance of the differences between direct rationality in rational choice actions and indirect rationality in rational behaviour in the sub-section after the next. Presently, I want to concentrate on the importance of rationality as an attribute or quality.

Rationality - understood as the ability to make rational decisions - is of obvious instrumental value, particularly when the stakes are high. As a tool it can help us considerably in achieving specific short and long-term goals and, generally, in our pursuit of the good life, as well as in our pursuit of the common good. Those who are rational in their choices and conduct are more likely to be successful both in achieving specific goals and in leading the sort of life they want to live than those who are irrational, inconsistent and erratic. Exceptions have to be allowed for, of course, as sometimes it can happen that an irrational choice turns out to be the best one for furthering one's interests. For example, on the balance of probabilities, regularly gambling large sums of money is hardly the rational strategy for the average person in the pursuit of long term financial security. Yet it can happen that through a series of remarkable coincidences one ends up making a fortune. But, far from being the norm, cases like this are the exception. By and large, one's success rate in achieving goals is more or less proportional to the extent that one thinks and acts rationally.

This point has also been argued, and in considerably more detail, by Max Black¹⁸⁰ who challenges the views of Geach, Singer and Ayer according to whom the question *Why should I be rational?* is somehow illegitimate, or absurd, or that it can be dismissed

¹⁸⁰Black 1982.

because “it questions something which is normally presupposed.”¹⁸¹ In his own answer to the question, Black points out that

The price of poor reasoning, whether through misperception or erroneous anticipation, can be frustration, hurt or extinction. ... You should respect “basic reasons” (be basically rational) because not to do so is to expose yourself to almost certain frustration, pain and death.¹⁸²

The rational pursuit of one’s interests (especially self-interest and interests arising out of partisanship), however, can be carried too far and an important function of morality is to curtail rational pursuits of individual interests which would unacceptably infringe on the interests of others. Sen’s following little story is an illustration of just how undesirable and nasty uncurtailed, self-seeking rationality would be.

‘Where is the railway station?’ he asks me. ‘There,’ I say, pointing at the post office, ‘and would you please post this letter for me on the way?’ ‘Yes,’ he says, determined to open the envelope and check whether it contains something valuable.¹⁸³

Such a relentlessly calculating attitude towards others is as undesirable as it is unrealistic to suppose it of most of us. It would be very hard to relate to someone who constantly calculated the costs and benefits of doing every small thing, especially at close personal levels, such as in friendships, marriages and relevantly similar relationships. Also, as Johansen points out, “No society would be viable without some norms and rules of conduct. Such norms and rules are necessary for viability exactly in fields where strictly economic incentives are absent and cannot be created.”¹⁸⁴

But, notwithstanding the necessary moral limits, prudential arguments lend a strong normative dimension to rationality - a dimension within which theories of rationality can be used not only to describe, explain and predict, but also to prescribe, evaluate and criticise individual human behaviour. Moreover, even social and moral norms must be open to broad rational evaluation and criticism from the point of view of

¹⁸¹Singer 1993, p. 316.

¹⁸²Black 1982, p. 162.

¹⁸³Sen 1977, p. 332.

¹⁸⁴Johansen 1987, p. 664.

the collective, overall interests of the community in question, or from the point of view of other communities which are unduly and detrimentally affected by that community's practices and norms. Especially in a fast changing world, norms and customs can become obsolete, inapplicable, cumbersome or even destructive - and not only to the individual but also to the overall interests of those affected by them. Having outlived their usefulness, when norms become no more than the dead hand of the past clutching and strangling the living present through sheer inertia, a burden and an unnecessary liability for the community and its individual members, they ought to be questioned, undermined and abandoned. As with counterproductive and destructive habits and dispositions, continued adherence to such norms is (collectively) irrational and becomes a legitimate target for rational criticism. However, a word of caution is in order here, for it may still be *individually* rational to adhere to a collectively irrational norm. This may be the case, for example, when social sanctioning of defectors is severe enough. I argue this point in more detail below.

From here, for purposes of a satisfactory defence of revenge, the significance of demonstrating the possibility of rational revenge is obvious. For, if arguments were successful in establishing that revenge behaviour is inherently irrational, that it is incompatible with and defies all attempts at rational explanations, the case against revenge would be serious. A satisfactory defence of revenge must be able to resist Elster's claim that "the most plausible explanation for revenge behaviour lies in psychological propensities that do not appear to serve any individual, social, or genetic purpose."¹⁸⁵

It has been suggested to me, perhaps on the basis of lines such as the one just quoted, that Elster's claim is not that rational revenge is impossible, that no fully rational person would (or could) ever decide to take revenge, but merely that, as a matter of fact, actual acts and practices of revenge known to us fail to measure up to the demands of

¹⁸⁵Elster 1990, p. 863.

rationality.¹⁸⁶ This suggestion is not right, however. Elster's overall project and argumentation is philosophical, not anthropological. For example, he explicitly claims that, "In a population of individuals known by each other to be fully rational, nobody would ever exact revenge."¹⁸⁷ Again, after searching and finding no satisfactory answer to the already mentioned questions: "Why would a rational person ever carry out an act of revenge? Why would anyone believe a rational person if he threatened to do so?", Elster moves on to consider the functions of revenge with the opening sentence, "Even if revenge behavior cannot be shown to be individually rational, it might be socially beneficial."¹⁸⁸

Indeed, Elster's concluding claim that "the most plausible explanation for revenge behaviour lies in psychological propensities that do not appear to serve any individual, social, or genetic purpose" is made precisely on the strength of previous arguments which aimed to show that, whichever way we look at it, revenge behaviour can be neither individually rational, nor functionally beneficial for the individual or his group. This interpretation of Elster is also shared by Hamlin in "Rational Revenge"¹⁸⁹ and is going to be further borne out in later discussions of Elster's arguments.

In the light of the quoted passages and arguments, it is hardly defensible that Elster stops short of claiming that rational revenge is impossible. At the same time it must be pointed out that Elster is very selective and one-sided in his use and interpretation of anthropological material in an attempt to lend credibility to his general argument for the complete irrationality and uselessness of revenge. This one-sided treatment by Elster of the anthropological literature might have been another reason for the perception I have just argued against. But, even if Elster's irrationality claim extended only to actual cases of revenge, the anthropological evidence concerning real life revenge practices leaves no

¹⁸⁶In conversation, by Philip Pettit.

¹⁸⁷Elster 1990, p. 875.

¹⁸⁸Elster 1990, p. 876.

¹⁸⁹Hamlin 1991.

doubt that norm guided revenge behaviour in societies with strong conceptions of shame and honour, for example, has been to a very large extent individually rational, and often the result of very careful rational choice considerations. Once again, these points will be borne out in the course of the arguments and discussions which follow later in this chapter and the next.

Notwithstanding the importance of rationality for human behaviour in general, and for revenge in particular, it must be noted that there can be significant limits both to the applicability of particular theories of rationality, and to the moral and social desirability of unfettered rational pursuit of individual interests. There is a clear distinction between rationality as an attribute of individual people on the one hand, and theories of rationality which attempt to explain what rationality is, on the other. It is important to keep these two apart. Criticisms of a theory on account of its theoretical and practical shortcomings need not undermine the importance of the *attribute* the theory intends to explain. Conversely, showing that an attribute is not all-important from, say, a moral perspective, does not affect how good the theory which explains that attribute is - although it may affect the degree to which the theory can be considered important.

Thus, a claim such as Elster's, that revenge behaviour cannot be explained by recourse to a particular ends-means maximising rational choice theory,¹⁹⁰ need not, in itself, amount to a reason or objection against revenge, or even against its rationality, which may be demonstrated and accounted for by recourse to a different theory. This in fact is the case where revenge behaviour can be shown to have a principally deontological basis. The immediate motivation behind norm-guided revenge behaviour, for instance, is mostly deontological: *If a man should rape your mother, wife, or daughter, you must kill him*. Such deontological rules may not be above criticism by any means. But criticising someone who acted on such a deontological maxim by simply claiming that the person

¹⁹⁰Elster's claims and arguments will be considered in detail in Chapter 5.

looked to a rule or social norm, instead of his/her future interests to consider the costs and benefits of killing the rapist, would be misplaced.

This does not mean, of course, that consequences are not important. What it does mean, however, is that direct considerations of consequences in every single instance is the wrong way to make one's decisions. As Hare argued the point in relation to moral thinking and action,¹⁹¹ it is often more appropriate and better to act on the basis of established rules and norms, rather than direct cost-benefit analyses in every situation. Such norm-guided behaviour is not different from meritorious promise keeping. If I keep my promises only if, and when, it suits me in strictly forward-looking terms, I have a questionable claim to merit. If, on the other hand, I keep my promises as a matter of principle, because I accept the fairness and rightness of the injunction 'Whoever makes a promise, must keep it,' I am living up to a moral ideal. Any criticism that in keeping my promises I am looking backwards, rather than conducting a quick cost-benefit analysis for my future interests, would be repugnant.

But while a direct ends-means maximising rational choice theory is inappropriate and is unable to explain deontologically based revenge and promise keeping behaviour, both of them can be shown to be rational by recourse to a theory of indirect rationality. This distinction between direct and indirect rationality is important and will be discussed in more detail in a separate section further below. Now I move on to consider the formal and substantive requirements of rationality.

4.2.2 Formal and substantive rationality

I argued earlier that there is a normative dimension to rationality within which rational choice theory can be used to prescribe, evaluate and criticise human behaviour. It is not difficult to see that the domain, or extent, of this normative dimension is determined by how robust one's theory of rationality is. A good starting point and

¹⁹¹Hare 1981.

framework for discussing this aspect of rational choice explanations is provided by Elster's earlier studies in the subversion of rationality in his book *Sour Grapes*.¹⁹² Elster points out that "the connotations of the word 'rational' range from the formal notions of efficiency and consistency to the substantive notions of autonomy or self-determination."¹⁹³ Accordingly, he distinguishes between 'thin' and 'thick' senses of 'rational.' In this section I shall argue that, although there is a legitimate case for drawing a distinction between formal and substantive rationality, Elster's account of the latter is seriously flawed. I start by considering the relevance of the 'thin - thick' distinction to strategies aiming to establish, or demolish, the rationality of revenge.

On a purely formal level, rationality is merely a matter of efficiency and consistency. To be rational in this thin sense, one only has to choose an option which seems to be the best means in the light of one's beliefs in order to achieve one's proposed goal(s), whatever that goal might be. The thin theory of rationality is thin in that

it leaves unexamined the beliefs and the desires that form the reasons for the action whose rationality we are assessing, with the exception that they are stipulated not to be logically inconsistent. Consistency, in fact, is what rationality in the thin sense is all about: consistency within the belief system; consistency within the system of desires; and consistency between beliefs and desires on the one hand and the action for which they are reasons on the other hand.¹⁹⁴

For most intents and purposes this sense of 'rational' is too thin. A requirement or injunction to be rational in this thin sense can be all too easily satisfied. In fact, it can be satisfied even by people whom we would be reluctant to consider to be rational. For example, a person who has a strong desire to die a most spectacular and violent death by, say, blowing him/herself up in a nuclear explosion, will count as rational in doing so if s/he satisfies the formal requirements of efficiency and consistency. Therefore, if we are to exclude such persons and actions from the domain of the substantively rational, we need to go beyond the formal requirements embodied in the thin sense.

¹⁹²Elster 1983.

¹⁹³Elster 1983, p. 1.

¹⁹⁴Elster 1983, p. 1.

All the same, if it could be shown that revenge behaviour in general (or certain instances of it) cannot meet the formal requirements of rationality in the thin sense, no further arguments would be needed to establish the irrationality of revenge. It would be unnecessary to consider further the rationality of the beliefs and the desires behind revenge on a more substantive level because the formal requirements of rationality embodied in the thin sense are necessary prerequisites for more robust, substantive accounts of rationality in the thick sense. Relying on the fact that thick rationality subsumes thin rationality, Elster tries to establish the irrationality of revenge behaviour by reference to the thin conception alone. For example, in arguing against the individual rationality of revenge as a means of saving one's honour, his attack is not focused on honour as a rationally indefensible goal or value - a view with which Elster clearly concurs - but on the informativeness of the explanation:

While true enough, such statements are uninformative. In particular, they contain no guide to how the agent could rationally choose between alternative means to the stipulated end.¹⁹⁵

I shall endeavour to show in the course of the next chapter that Elster fails to demonstrate the individual irrationality of revenge behaviour in the thin sense. I shall also argue that revenge can, in principle, be rational, and that there are in fact many instances of revenge in real life which are rational not only in the thin sense, but in a proper, substantive sense of the word. For these purposes, however, in addition to the formal requirements, we must specify more demanding conditions for rationality on a substantive level, requiring that the agent's beliefs and desires relevant to the action in question themselves be rational, or at least that there should be no good reasons for questioning their rationality. For an account of substantive rationality I start by considering Elster's earlier work in this area.

We need a broader theory of rationality that goes beyond the exclusively formal considerations ... and that allows a scrutiny of the substantive nature of the desires and beliefs involved in action.

¹⁹⁵Elster 1990, p. 872. Although, as I shall discuss it later, Elster says some unflattering things about honour in his concluding remarks (pp. 883 - 85), he never relies on them in his rejection of the individual rationality of revenge behaviour.

We want to be able to say that acting rationally means acting consistently on beliefs and desires that are not only consistent, but also rational.¹⁹⁶

On the other hand, we do not want to dilute the notion of rationality so that it comes to encompass all the good properties that we might want our beliefs and desires to have. I suggest that *between the thin theory of the rational and the full theory of the true and the good there is room and need for a broad theory of the rational*. To say that truth is necessary for rational beliefs clearly is to require too much; to say that consistency is sufficient, to demand too little. Similarly, although more controversially, for rational desires: the requirement of consistency is too weak, that of ethical goodness too strong.¹⁹⁷

Elster's solution to specifying the broad theory of the rational without collapsing it into the moral is to look at the way in which the relevant beliefs and desires are formed.

A belief may be consistent and even true, a desire consistent and even conformable to morals - and yet we may hesitate to call them rational if they have been shaped by irrelevant causal factors, by a blind psychic causality operating 'behind the back' of the person. The stress here should be on 'irrelevant' and 'blind', not on causality as such. ... All desires and beliefs have a (sufficient) causal origin, but some of them have the wrong sort of causal history and hence are irrational. Since it is so hard to say exactly what would qualify as the right sort of history, I shall have relatively little to say about this (crucial) problem. I shall have more to say ... about all the wrong sorts.¹⁹⁸

Accordingly, throughout the book Elster remains imprecise on what counts as rational belief and desire formation, although it is hard to disagree with John Bishop that Elster "does engender an appreciation of the difficulty of defining substantive rationality."¹⁹⁹ I shall consider Elster's position on the substantive rationality of beliefs first and then move on to discuss his account of the rationality of desires. I shall argue that, with regard to both, Elster has set out in the wrong direction to 'track down his quarry.'²⁰⁰

¹⁹⁶Elster 1983, p. 15.

¹⁹⁷Elster 1983, p. 15.

¹⁹⁸Elster 1983, pp. 15 - 16.

¹⁹⁹Bishop, 1985, p. 246.

²⁰⁰This metaphor is Bishop's. (Bishop 1985, p. 245.)

According to Elster, a belief is rational if it issues from an exercise of judgement.

The positive characterization of rational beliefs can be made in terms of the notion of judgement, defined as the capacity to synthesize vast and diffuse information that more or less clearly bears on the problem at hand, in such a way that no element or set of elements is given undue importance. Clearly this is not a very helpful definition but there is little doubt about the reality of the phenomenon. We all know persons who have this quality and others who lack it. In some walks of life it is indispensable to have it, and those who lack it are soon eliminated.²⁰¹

Elster's characterisation of rational belief formation by reference to judgement is plausible. His requirement that beliefs must not be distorted by cognitive faults or irrelevant affective drives is also plausible. Admittedly, this is a difficult criterion to develop and the issue has to be covered in more detail.

A good starting point for discussion is provided by the observation that one's true beliefs and interests can often be in conflict in the sense that having a different belief, even if false, would serve one's interests much better than some true beliefs would. For example, the belief that one's boss at work is objectionable, narrow-minded and unworthy of one's respect can make one's working environment very stressful, if not unbearable, which in turn can negatively affect other areas of one's life, not to speak of one's promotion and career prospects. It is at least arguable that in some of these situations it would be much better from the point of view of one's interests to believe, even falsely, that one's boss is not an objectionable person. Yet, if someone formed this latter belief in the face of overwhelming evidence to the contrary precisely because the belief served their interests, we would say that the belief in question was irrational because it was distorted by their interests. What this suggests is that the immediate aim or motive of rational belief formation must always be the truth of the matter and that one's non-epistemic interests, desires and wants must be irrelevant in the process.

Yet, this view is too simplistic. For, while we may agree with Pears that "truth always retains some magnetism when a person is forming a belief" and that "it is a goal

²⁰¹Elster 1983, p. 16.

with an attractiveness that cannot fall to zero,”²⁰² it is far from clear that non-epistemic considerations must always be prevented from exerting influence over the formation of our beliefs. To begin, consider the following argument by Heil.

If I am the sort who can compartmentalize attitudes, then an epistemically unwarranted belief or collection of beliefs may be reasonable for me to hold because my so doing affords, on balance, sufficient benefit. A hard-nosed scientist who harbors strong but evidentially substandard religious or political convictions may nevertheless do so reasonably if these make his life worth living and if they remain functionally isolated from his scientifically mediated doxastic corpus.²⁰³

It would be a mistake to accept such interest-generated, epistemically unwarranted beliefs as rational. It is significant to note here that Heil himself stops short of making such a claim. He only claims that someone may *reasonably hold* certain interest-induced but epistemically unwarranted beliefs, provided that such beliefs are compartmentalised so that they are not allowed to corrupt the epistemically warranted beliefs of that person. Such beliefs may be reasonably held because they serve a vital function for the individual holding them without doing any collateral damage. If they are harmless and useful at the same time, there is no reason for ostracising them and good reasons for retaining them - a familiar argument defending the comparable status of white lies.

It is possible to push the boundary further, however, and argue that non-epistemic considerations may exert considerable influence over the formation of our beliefs without rendering those beliefs irrational. Indeed, it is plausible to suppose that non-epistemic interests often play a role in the formation of rational beliefs. It is worth quoting Heil again,

Some cases are clear-cut. When the evidence is overwhelming or when it is inconclusive, we have little difficulty in assessing warrant [to form a belief]. Often, however, matters are less straightforward. When that is so, when there is an epistemic grey area, an agent may be epistemically permitted to hold a particular belief and permitted to refrain from holding it. In tuning doxastic thresholds ... we exploit this grey area so as to reflect pertinent non-epistemic interests. We encourage in ourselves and in others a tendency to be tactful in beliefs formed about

²⁰²Pears 1984, pp. 41 - 42.

²⁰³Heil 1992, p. 56.

friends, colleagues, and loved ones, for instance, a tendency that plays a diminished role in the formation of comparable beliefs about those to whom we enjoy no special attachment. In this respect, at least, non-epistemic factors might - reasonably and naturally - influence agents' doxastic mechanisms without thereby corrupting their epistemic standards.²⁰⁴

One may interpret certain signs in a stranger, for example, to be stinginess while the same signs may be interpreted in a loved one as frugality. Although both judgements are based on the same evidence, both may be regarded as rational in spite of the fact that our tendency to be charitable in our judgement of significant people in our lives is likely to be motivated quite often by non-epistemic interests.

What this shows is that non-epistemic interests may be allowed to exert an influence on the formation of a rational belief, provided that their influence is not too strong. However, determining what counts as too strong in this regard can be difficult and it is of no help simply to insist that this influence is too strong when the truth of the belief in question is placed in jeopardy by it. For one thing, truth and falsehood do not divide neatly along the rational-irrational distinction. It is hardly controversial that a true belief may be irrational and a rational belief may be false. Moreover, the grey area in question is one in which truth may plausibly be seen as displaying a kind of elasticity because a lot depends on interpretation. One and the same set of data can be given a generous, as well as a less generous interpretation in judgement and the result may well be that, in spite of their incompatibility, neither belief can be rejected as false. If this apparently violates the law of excluded middle, then the claim can be reformulated to say that in this grey area in question, judging the truth of the matter is virtually impossible between individually plausible interpretations even if those interpretations are incompatible with each other.

But, in spite of the ambiguities concerning the degree to which desires and interests may influence beliefs without thereby making them irrational, the stress has to remain throughout on the causal factors in question being appropriate and relevant to the

²⁰⁴Heil 1992, p. 58.

resultant beliefs. It is a mistake to give weight, as Elster seems to do, to a requirement that the causal factors in question not be “blind” or “operating ‘behind the back’ of the person,” where ‘blind’ and ‘behind the back’ are interpreted to mean that the process in question takes place without the person’s conscious awareness. The reason for this is that the manner of a person’s conduct in their speech and body language, for example, often tells us more about them than the actual content of what they say. Much of this information is received and correctly interpreted at an unconscious level by us and influences our beliefs and attitudes towards them. As Taylor et. al. point out,

People are seldom consciously aware they receive, interpret, infer, and judge [information] based on nonverbal symbols, but they do. ... Estimates vary, but a conservative conclusion is that 60 to 80 percent of the meaning in a speech communication situation is derived from nonverbal symbols.²⁰⁵

A recent study reported in *Nature* also shows that people can be highly successful in receiving, decoding and judging such information. In one experiment, for example, radio listeners are reported to “have detected the lies 73.4% of the time.”²⁰⁶ This is a result which is highly significant. But if reliable beliefs can be formed on the basis of subconsciously gathered, sorted and analysed data, the resultant beliefs should have no difficulty in passing as substantively rational - though, of course, they can still be mistaken in the same way in which beliefs formed on the basis of consciously gathered, sorted and analysed information can be mistaken.

Indeed, if 60 to 80 percent of the meaning in a speech communication situation is derived from nonverbal symbols, much of which happens outside the sphere of the agent’s conscious awareness, then Elster’s notion of judgement in the rational belief formation process must be extended to allow for such highly relevant and often reliable causal processes which are, however, ‘blind’ and take place ‘behind the back’ of the person in the sense that the person in question is not fully aware of them at a conscious

²⁰⁵Taylor, Rosegrant, Meyer and Samples 1977, p. 77.

²⁰⁶Wiseman 1995, p. 391.

level. Otherwise the conclusion would have to be that, therefore, most, if not all, beliefs formed on the basis of information received in speech communication situations are tainted with unconscious causal influences and must, therefore, be irrational. To say the least, this would be wildly implausible. The conclusion is that the conscious-unconscious distinction is not very helpful in this context. In so far as rationality of belief is concerned, what matters in judgement is that crucial and relevant information is given due weight while irrelevant information is appropriately barred from influencing the belief in question. Whether the relevant information is received and interpreted at a conscious level or not, is neither here nor there.

The same holds true for Elster's account of the substantive rationality of desires and preference change. With regard to desires, or preference formation, Elster's key notion is "autonomy, or self-determination." To be rational, a desire must be autonomous. But what is it for a desire, or for a preference change, to be autonomous? To this question Elster offers no satisfactory answer. To his credit, he once again openly admits that he "can offer no satisfactory definition of autonomy."²⁰⁷ This simply means that he does not offer a *positive* definition of it. He characterises it negatively: A desire is autonomous, and hence rational, if it is not distorted by cognitive defects or affective drives.²⁰⁸

But, whereas in discussing the rationality of beliefs Elster was only running the danger of losing sight of the ball by giving undue attention to the requirement that the processes in question not be 'blind' or take place 'behind the back' of the person's awareness, in the case of desire rationality he often seems to lose sight of the ball altogether. This is because he shifts all the weight from the crucial consideration of relevant and irrelevant causal factors to the mistaken requirement that the desire forming processes be a conscious process, rather than 'blind' and taking place 'behind the back'

²⁰⁷Elster 1983, p. 21.

²⁰⁸Elster 1983, p. 24.

of the person. At the same time, Elster seems to be aware to some degree that the conscious-unconscious dichotomy is unhelpful in his quest to give an account of substantive rationality by reference to rational beliefs and desires. The result is that he ends up saying contradictory things about desire autonomy. He does this, for example, when he contrasts ‘sour grapes,’²⁰⁹ or adaptive preference formation, which he claims to be irrational, with intentional character planning, which he is happy to accept as being rational. Both mechanisms serve the same purpose or function: they reduce the agent’s level of frustration arising from unattainable wants, and the only discernible difference between the two mechanisms is that while the latter is conscious and intentional, the former takes place without the agent’s conscious awareness.

Adaptive preference formation is the adjustment of wants to possibilities - not the deliberate adaptation favoured by character planners, but a causal process occurring non-consciously. Behind this adaptation there is the drive to reduce the tension or frustration that one feels in having wants that one cannot possibly satisfy. ... Sour grapes is a purely causal process of adaptation, taking place ‘behind the back’ of the person concerned. Very different indeed is the intentional shaping of desires advocated by the Stoic, Buddhist or Spinozistic philosophies, by psychological theories of self-control or the economic theory of ‘egonomics.’²¹⁰

Yet, elsewhere in the book, Elster himself argues that the idea that “autonomous desires are desires that have been deliberately chosen, acquired or modified - either by an act of will or by a process of character planning ... gives us both too little and too much.”²¹¹ This, as I argue shortly, is correct. But if awareness and intentionality are neither necessary, nor sufficient, for the autonomy of desires, and by extension to their rationality, then Elster’s heavy, if not singular, reliance on these concepts in order to explain the difference between the substantive rationality of character planning and the irrationality of sour grapes must be wide of the mark. It is not difficult to see that the agent’s awareness and conscious intentions to form a particular desire are not necessary

²⁰⁹From the fable of the fox who, because he could not reach the grapes, decided that he no longer wanted them because they were sour anyway.

²¹⁰Elster 1983, pp. 25 & 117.

²¹¹Elster 1983, p. 21.

for the rationality of that desire. Elster himself warns us that, “From the plausible proposition that the *capacity* for second-order evaluation is a condition for personhood, we should not conclude that the actual exercise of this capacity is a condition for autonomy. Through sheer moral luck people may achieve autonomy without striving for it.”²¹²

We need not go as far as moral luck, however, in order to show that rationality is not subverted just because some mechanism of dissonance reduction operates on an unconscious level, ‘behind the back’ of the person’s awareness. For, as with the indirect rationality of actions I have discussed in the preceding section, there can be unconscious processes of rational desire formation and preference change. Consider the following example. A gambler may lose the desire to continue gambling because of a justified, but unconscious fear of impending financial ruin and the resultant loss of things, and people, most dear to him. He may rationalise his preference change on a conscious level by claiming to have found more interesting and exciting ways of spending his time and money. He might even deceive himself as to his real reasons and needs and come to hold an irrational belief about his preference change - an irrational belief which, while it is *about* his preference change, does not form part of the *causal* process which brought about that change. But does it follow, therefore, that his preference change itself, and the resultant desire not to gamble any more, are irrational?

There is a compelling case here, I believe, for answering this crucial question in the negative. The change is clearly in his self-interest and in the interests of those near and dear to him. Also, the causal factor (the high risk of disaster, or the realistic fear of it) is most relevant to the formation of the new desire. Given these two considerations, the outcome would not have been any different if the process took place ‘in front’ of the person’s awareness, as opposed to having taken place ‘behind his back’ and, therefore, it seems perfectly rational. More generally, desire which is rational to have from the point

²¹²Elster 1983, p. 22.

of view of one's (self-) interests and one forms that desire for that very reason (because it is rational to have it from the point of view of one's interests), it is not a good argument against the rationality of that desire, or against the rationality of forming it, that the process was 'blind' because it took place on an unconscious level, 'behind the back' of the person's conscious awareness.

This is not to say, of course, that there is no difference at all between the two cases. The difference, however, is not between rationally and irrationally formed desires, but between conscious and unconscious processes of desire formation. Both conscious and unconscious ways of forming the desire not to gamble are rational - though, in the case of unconscious processes, the resultant desire might be better described as being rational in an indirect sense. In the following sub-section I explore in detail the analogous difference between directly and indirectly rational actions and I argue that, for most intents and purposes, being indirectly rational in one's actions is as good, or pretty nearly as good, as being directly rational. The same holds true, I suggest, for the direct and indirect rationality of desires. What counts is that one acquires the right desires and preferences, the desires and preferences which it is in one's interests to have. It is also important that those preferences are acquired for, or are caused by, the right, or relevant, reasons. But when both these considerations hold, then the fact that a certain desire has been acquired through unconscious mental processes 'behind the back' of one's direct awareness, rather than through conscious, intentional ones, seems to be, as in the case of rational belief formation, neither here, nor there.

This conclusion is further supported by the fact that awareness and intentionality are not sufficient to guarantee desire rationality. They are unable to screen out the formation of clearly irrational desires. Someone could quite consciously and consistently develop the aforementioned desire to die a spectacular death by being incinerated in a nuclear explosion. In so far as intentionality is concerned, this desire is autonomous, yet crazy and eminently irrational. Indeed, Elster himself argues that

a desire stemming from intentional character planning can be no more autonomous than the intention from which it stemmed, and so we at once get into a regress. There is no reason, moreover, to believe that second-order desires are always immune to irrelevant causal influences. ... The very activity of character planning may bring about a rigidity of character which is incompatible with the 'tolerance of ambiguity' often said to be characteristic of ego-strength or autonomy."²¹³

The preceding discussions and arguments show quite conclusively that intentionality and conscious awareness in the desire forming process, is neither necessary, nor is it sufficient, for the formation of substantively rational desires. This does not imply, of course, that intentionality is never relevant to rational desire formation. As the example of rational character planning shows, it sometimes forms an important part of the rational desire formation process. This, however, is the most that can be said about it in the current context and, therefore, if the notion of autonomy is to be retained, it will have to be defined or explicated, not by reference to awareness and intentionality, but in terms of other concepts.

Before making concrete suggestions in this regard, however, I want to point out that it is doubtful whether Elster's focus on unconscious drives can be of much help here. Elster has an almost mythical conception of drives as being "non-conscious psychic forces that are geared to the search for short-term pleasure, as opposed to the conscious desires that may forego short-term pleasure to achieve some longer-term gain."²¹⁴ According to him, "Drives shape desires (and beliefs), but are not themselves desires, because they are not conscious and known to the person who has them."²¹⁵ Moreover, Elster pleads "ignorance of the substantive character of drives."²¹⁶ This

²¹³Elster 1983, pp. 21 - 22.

²¹⁴Elster 1983, p. 25.

²¹⁵Elster 1983, p. 24.

²¹⁶Elster 1983, p. 25.

ignorance, however, is so deep that the very existence of drives “in the present state of the arts must be inferred from behaviour rather than studied directly.”²¹⁷

To say the least, this is unsatisfactory. It is hard to resist the analogy between these mysterious drives postulated by Elster and early scientists’ postulation of the *funiculus*.²¹⁸ The *ad hoc* nature of Elster’s appeal to drives is not disguised by his implied suggestion that one day we may be in a better position to study these mysterious entities. It is hard to see what kind of philosophical or psychological progress would enable us to study Elster’s drives in a direct way. Apart from being inherently mysterious and *ad hoc*, Elster’s reliance on the distinction between non-conscious drives and conscious desires is further undermined by the fact that there are many consciously held desires which are geared towards short term satisfactions, as opposed to long-term ones. Preferring to have one’s meal in pleasant surroundings is as good an example as any other.

To be sure, I do not wish to deny that, ordinarily conceived, conscious or unconscious drives, such as the drive (urge) for food, sex and self-preservation, or a strong desire for riches, can distort one’s beliefs and desires. To use Elster’s example, “A desire, say, for promotion may bring about a belief that promotion is imminent.”²¹⁹ Such drives or desires, however, are neither necessary, nor is their presence sufficient, for the formation of clearly irrational beliefs and desires. For example, irrational beliefs and desires may result from poor interpretive skills, such as when one mistakenly interprets someone’s friendliness to be a romantic interest. Similarly, a person stranded in the desert may have a powerful drive to relieve his thirst and to survive, but, say, upon seeing a mirage, his drives do not necessarily cause an irrational belief in him that there is water.

²¹⁷Elster 1983, p. 25.

²¹⁸The *funiculus* was believed to be a very thin, and eventually invisible, thread. Its existence was postulated in order to explain why water (and liquids in general) remained “suspended” in an upside-down container with its mouth or rim submerged.

²¹⁹Elster 1983, p. 26.

To sum up these discussions on affective drives, distinguishing between drives and desires in the way done by Elster, is both *ad hoc* and dubious. Neither an appeal to the conscious-unconscious distinction, nor an appeal to the short- and long-term distinction can ensure a plausible criterion for distinguishing the drives in question either from what are commonly regarded as desires, or from what are ordinarily regarded as drives, such as the urge for food and self-preservation. Therefore, a better way must be found to account for the substantive rationality and irrationality of desires than this kind of appeal to obscure, if not mythical entities.

I suggest that a much more promising place to start looking for the substantive rationality of desires is the intersection of the notion of *relevance* which I have stressed earlier, the accompanying notion of *proportional weight*, or *significance*, and the notion of *substantive agent interests*, typically self-interest. It is typically self-interest, but not always, because rationality can be as useful to the altruist as it can be to the egoist. Some desires, of course, such as a desire for ice cream, for example, may be too trivial to make much of a difference to the agent's more substantial, long-term interests and such desires may just as well be regarded either as arational, or just not worthy of rational evaluation. But when it comes to desires which have important implications for the agent's overall interests, they are not going to escape substantive rational criticism, if they are destructive, if they are against the interests of someone without at the same time being in the interests of anyone, not even in the interests of the agent. Thus, it is necessary that rationally formed, or held, desires should not be against the agent's (self- or other-regarding) overall interests.

But this condition of agent interests is not quite sufficient on its own. Conditions of relevance and proportional significance are also necessary for a proper account of the substantive rationality of desires. Regardless of their sphere of operation with regard to the agent's awareness, the causally efficient reasons, beliefs, desires and needs must have a proportional bearing on the resultant desire. In other words, they must be relevant to the desire in question and even then they must contribute to the formation of the desire

to an appropriate degree relative to their proper significance or weight in the situation. Suppose, for example, that the gambler lost his desire for further gambling not because of realistic perceptions of the nature of the game, but as a result of coming to believe that he is facing imminent financial ruin because a black cat has crossed his path. Although his loss of desire for gambling is still due to the fear of losing, and it still is in his interests to change his desire to gamble, we would be reluctant to call the change in his preferences to have been directly rational. This is because the crucial belief behind his reasoning, namely, that black cats bring you bad luck, itself is highly irrational. In this example the irrelevance appears as part of the belief and is ultimately responsible for that belief's irrationality. Black cats crossing one's path have nothing to do with one's good or bad luck in general, or gambling in particular.

Notwithstanding, it is at least arguable that even if the change in desire were irrational in a direct sense, the resultant desire not to gamble any more would be a rational one in a less direct sense of the word. This is the sense in which the desire is in the agent's interests to have. That one acquires a highly beneficial desire through what seems to be quirky and otherwise irrational thinking, need not divest that desire from its rationality in an important, if indirect, sense of 'rational.' The same point could be argued regarding accidental preference changes. Suppose that the gambler's preference change comes about as a result of a bump to his head. While, in one sense, accidental knocks to the head are hardly relevant to preference change, it is not difficult to imagine a relieved wife's commenting tongue in cheek: 'I wish I had known much earlier how to cure him of his irrationality, how to bring him to his senses.' I explore the distinction and significance between direct and indirect rationality which is suggested by these examples in the following sub-section.

4.2.3 Direct and indirect rationality

There are at least two ways in which human behaviour can satisfy the demands of a normative rational choice theory: directly and indirectly. In the case of direct rationality

people consciously decide on certain courses of action because of the recognition that in the circumstances they are the rational ones to take, because they are the best (or just good enough) options from among alternatives for achieving some proposed goal(s). A good example would be a student's decision to stay home to study and sleep before an exam, instead of giving in to peer pressure to have a late night out. In cases like this a person acts rationally by *choosing rationally* and the rationality involved in this kind of decision making which is based on direct cost-benefit analysis is the proper subject of what is commonly known as 'rational choice theory.' Some of my arguments will aim to show the rationality of revenge behaviour in this direct sense.

But people can, and often do, act rationally in an indirect way, that is, without making rational choice calculations. For example, we often act rationally in this indirect way when we keep our promises, follow social norms, brush our teeth, and display good manners, such as the proper greeting of friends and acquaintances, holding the lift for someone, or telling them the time because we are politely asked to do so. In indirect rationality the normative requirement on the agent is that her action must form part of a rational habit, disposition, custom or rule-following. The rationality of habits, customs and dispositions, in turn, can be evaluated by reference to how well they serve or hinder the agent's particular and overall interests, such as dental health, getting on well with others, material prosperity, and whatever else forms part of the agent's flourishing and, what may be referred to as 'the good life.'²²⁰

Satz and Ferejohn have recently defended a version of rational choice theory which is relevantly similar to the account of indirect rationality I have just presented. Although they are critical of rational choice theory when it is interpreted as an internalist,

²²⁰As Roy Perrett pointed out to me, it may be possible to explicate the contrast between direct and indirect rationality by reference to Hare's distinction between two levels of moral thinking: the critical and the intuitive. "The intuitive and critical levels of thinking are both, unlike the meta-ethical, concerned with moral questions of substance; but they handle them in different ways, each appropriate to the different circumstances in which, and purposes for which, the thinking is done. (Hare 1981, p. 26.) According to Hare, while theoretical moral thinking may require us to think always at the critical level, everyday moral thinking often has to be more intuitive, relying on what are often referred to as 'rules of thumb' but for which Hare prefers to refer as 'prima-facie principles.' (p. 38.)

individualist psychological theory, they defend it under a different interpretation, as an externalist, structuralist social theory which explains human behaviour “in terms of irreducible relational or structural properties.”²²¹ They argue that behaviour conforming to a rational choice pattern is multiply realisable and, therefore, that in many situations a rational choice explanation is appropriate by reference to the relational and structural properties presented by the particular circumstances. In such cases, the question whether the theory is psychologically accurate of the individual actor(s) or not, is secondary, if not altogether redundant. Thus, in certain circumstances, rational choice theory may tell more about the environment of some observed behaviour than the direct psychological realities behind that behaviour.

There is a tendency for rational choice theory practitioners to read off truths about human psychology from the successful uses of the theory. But the behaviour of rats, pigeons, hunter-gatherers, and stockbrokers may not tell us much about the nature of our rationality or our moral psychology. It tells us more about the structure of certain environments, environments we hopefully shall not attempt to duplicate in other areas of our social life.²²²

Satz and Ferejohn show how the theory can be applicable and relevant as an explanation by reference to external circumstances alone which force rational behaviour on the individual or the group by the pressures of automatic, if blind, selection. In situations of scarcity and where the stakes are high, those who do not behave rationally are soon eliminated and, thus, rational conduct tends to get automatically entrenched. Direct rational choice deliberations need not form part of this process at all. The direct and immediate reasons behind indirectly rational behaviour might be genetically programmed instincts, acquired habits and dispositions, customs, rules, or codes of conduct which became entrenched because they proved successful over time in delivering the goods, so to say, and not necessarily because they were consciously chosen.

²²¹Satz and Ferejohn 1994, p. 72.

²²²Satz and Ferejohn 1994, p. 87.

I shall argue in the course of later discussions that, although there are situations in which revenge behaviour may not meet strict criteria of direct rationality, it may nevertheless meet criteria of indirect rationality. I shall presently argue that either type of rationality is good enough to satisfy a prudential imperative to be rational. The absence of a direct rational choice motivational structure by itself constitutes no objection to an action which is otherwise rational. This view finds direct and indirect support in familiar criticisms of rational choice theory - which, by all accounts, is the dominant, if not the only developed, theory of direct rational choice. I intend to show that, although rational choice theory may be able to weather many of these criticisms, the costs incurred are considerable. Moreover, there is at least one problem, namely, the problem of cooperation in more-than-zero-sum games, where rational choice theory is found seriously wanting both in the explanatory and normative domains. My final objective with these discussions is to show that the same limitations are not present in an account of indirect rationality of the kind I presented above, and that, therefore, indirect rationality is as good, and sometimes better than direct rationality, at satisfying a prudentially grounded normative injunction to be rational.

In his book *Having Reasons*, Schick puts up a strong case against a dominant form of rational choice theory according to which all human behaviour is explainable by reference to self-regarding interests of the individual. He convincingly argues that human behaviour is often motivated, not by selfish rationality, but by genuinely other-regarding sociality.²²³ In his paper "Rational Fools," Amartya Sen similarly challenges rational choice theory based on traditional economic models which view individual human behaviour to be the maximisation of self-serving individual interests in consequential terms. He shows that commitment - be it to individuals (friends or relations), or groups (one's clan, club, class or political party), or to moral principles (honesty, impartiality, equity or fair distribution of resources) - "drives a wedge between

²²³Schick, 1984.

personal choice and personal welfare.”²²⁴ He also points out that, “Commitment sometimes relates to a sense of obligation going beyond the consequences ... [and that] Sometimes the lack of personal gain in particular *acts* is accepted by considering the value of *rules* of behaviour.”²²⁵

Challenged by such criticisms, rational choice theory can be salvaged by a thinly disguised variation on definitional egoism, whereby a person’s interests are no longer defined in terms of his individual welfare, but by his preferences as revealed by actual choices. Now the agent is seen to be maximising (or satisficing) *her own* utility regardless of whether she is a recalcitrant egoist, a saintly altruist, or a one-track ideologue. This move comes at a price, however, for now it appears that an agent’s behaviour is explained in terms of his preferences which in turn are postulated or derived only from his behaviour - effectively leading to what Sen calls “a remarkably mute theory” surrounding which, “not surprisingly, excursions into circularities have been frequent.”²²⁶ Nevertheless, as Samuelson observes, the theory “is not in a technical sense *meaningless*.”²²⁷ It is meaningful in the sense that it remains falsifiable, at least in principle, because the formal requirement of consistency would be violated by agents with inconsistent preference orderings. But, as Sen is quick to point out, there are serious difficulties in devising satisfactory observation conditions for testing the consistency requirement, since “on the one hand, our love of variety makes it illegitimate to consider individual acts of choice as the proper units (rather than *sequences* of choices) while, on the other hand, lapse of time makes it difficult to distinguish between inconsistencies and changing tastes.”²²⁸

²²⁴Sen, 1977, p. 329.

²²⁵Sen, 1977, p. 342.

²²⁶Sen, 1977, p. 325.

²²⁷Samuelson 1948, p. 92.

²²⁸Sen, 1977, p. 325.

Most importantly, however, Sen argues that the theory is unable to make sense of the complex nature of human decisions:

traditional theory has *too little* structure. A person is given one preference ordering, and as and when the need arises this is supposed to reflect his interests, represent his welfare, summarize his idea of what should be done, and describe his actual choices and behavior. Can one preference ordering do all these things? A person thus described may be “rational” in the limited sense of revealing no inconsistencies in his choice behavior, but if he has no use for these distinctions between quite different concepts, he must be a bit of a fool. The *purely* economic man is indeed close to being a social moron. ... If he shines at all, he shines in comparison - in contrast - with the dominant image of the rational fool.²²⁹

Satz and Ferejohn also agree.

Human psychology, studied in the laboratory, the psychoanalyst’s office, or through introspection, appears to be far more complicated and has much more structure than rational-choice theory allows. The conception of human rational agency in terms of maximizing over a complete and consistent set of preference orderings is not psychologically realistic.²³⁰

While there may be ways to overcome this particular difficulty - perhaps, by resorting to “*rankings of preference rankings*” as suggested by Sen²³¹ - it is virtually certain that rational choice theory has no solution to the problem of cooperation in more-than-zero-sum games. The prisoner’s dilemma is the most important of these games and will serve well to illustrate the point. Prisoner’s dilemma situations are very common and they are an inescapable feature of inter-personal relationships. Yet, in these situations the prescriptions of rational choice theory lead to failure in securing the optimum feasible outcome for those concerned. The theory advises against mutual cooperation which, otherwise, is the obviously most desirable, if not rational, thing to do because it is the only realistic strategy by which those involved can adequately further their interests. Advising defection instead, the theory *ipso facto* defeats its own

²²⁹Sen, 1977, pp. 335 - 36 & 344.

²³⁰Satz and Ferejohn 1994, p. 74.

²³¹Sen 1977, p. 337. For full details, see Sen 1974 (and other related articles) in Korner 1974.

prudentially grounded normative purpose.²³² Consequently, rational choice theory is vulnerable to the criticism that it is too narrow in the earlier discussed basic, pre-theoretical sense of the word - the sense which is normally captured by reportive dictionary definitions. In prisoner's dilemma situations players often decide to cooperate because they make a reasoned assessment of the situation in this pre-theoretical but compelling sense of the word 'rational.' Rational choice theory, however, fails to account for that rationality. To see this in detail, consider the dilemma.

Suppose that Fred and Betty are faced with the question whether to cooperate with each other for purposes of furthering their own individual interests. This can be anything from clearing a common driveway of a fallen tree, to escaping a lengthy prison sentence for an offence of which they are suspected, but for which the authorities need their confessions to convict them.²³³ Whatever the context, they each can cooperate and each can defect. For the four possibilities, their payoffs are as follows: If both cooperate, the interests of each are advanced by the value of, say, **10**; if both defect, they each get **2** (the punishment payoff); if only one of them cooperates, the cooperator gets **nothing** (the sucker payoff) while the defector gets **12** (the temptation payoff). Represented in a matrix, this looks as follows.

²³²This was discussed earlier in the chapter in connection with the significance of rationality.

²³³See Poundstone's *Prisoner's Dilemma* (1992) for a fascinating reading on the dilemma, its history and practical relevance.

		Betty	
		Cooperate	Defect
Fred	Cooperate	10(F) / 10(B)	0(F) / 12(B)
	Defect	12(F) / 0(B)	2(F) / 2(B)

Figure 5

If neither of them knows for sure what the other one is going to decide, what is the rational thing for each of them to do? According to rational choice theory, they should both defect. Fred will do better for himself if he defects regardless of what Betty decides to do. If Betty cooperates, his payoff is **12** as opposed to **10** and if she also decides to defect, then his payoff is **2** as opposed to **nothing**. Exactly the same rational choice reasoning is supposed to apply to Betty. But if they both defect, they can only secure the punishment payoff of **2-2**, which is far less desirable than a payoff of **10-10** which they can easily secure for themselves by cooperating instead.

Given that mutual cooperation is often the option which pays best, it is not surprising that voluntary cooperation is all around us in everyday life. Significantly, however, even in carefully controlled laboratory conditions, people playing the prisoner's dilemma recognise the self-defeating nature of senseless defections and, making a reasoned assessment of the situation, decide to cooperate instead in a very high percentage of cases.²³⁴ Confronted with such powerful evidence, some game theorists seem to be more willing to question people's intelligence than the adequacy of rational choice theory.

²³⁴See, for example, Lave 1962; Luce and Raiffa 1957; Rapoport and Chammah 1965.

Evidently, the run-of-the-mill players are not strategically sophisticated enough to have figured out that strategy DD [Defect-Defect] is the only rationally defensible strategy, and this intellectual short-coming saves them from losing.²³⁵

I propose a different conclusion: *Cooperation in these situations is fully rational, but rational choice theory is unable to account for the rationality behind it.* In other words, the narrow canons of rational choice theory make for an incomplete theory of rationality. I am not suggesting that some fatal flaw inherent in the theory renders it completely useless, but only that the theory has significant limitations. While the theory successfully accounts for some of what rationality as an attribute is, it fails to explain a rather important part of rationality - a case in point being its failure to explain the rationality of cooperation in prisoner's dilemma situations.²³⁶

Cooperation is an inescapable part of social life and, unless backed by coercion, many ordinary exchanges between people, as well as groups of people and organisations, assume the form of a prisoner's dilemma. The failure by rational choice theory to explain the rationality of cooperation is rendered particularly acute by the backward induction argument. This argument shows that *if*, in a prisoner's dilemma situation cooperation is irrational in a one-shot game, *then* it is similarly irrational in iterated games with a finite sequence. Consider first the implications of this argument. Its conclusion is clearly false. For example, if the above game is repeated between Fred and Betty, say, one hundred times, they would have to be eminently unwise and irrational to keep defecting on each other throughout the game. The strategy of systematic defection will further their individual interests only to the value of 200 units each, whereas mutual cooperation can see them five times better off. Yet, the backward induction argument shows that on the assumption that the only rational strategy for them in a one-shot game is to defect, they should rationally defect throughout an iterated game.

²³⁵Rapoport and Chammah 1965, p. 29.

²³⁶To be sure, this failure is not mere appearance. The payoff table is not confined to immediate payoffs, but shows the value of *all* of the relevant outcomes.

Therefore, unless there is something wrong with the backward induction argument, the assumption that the only rational thing to do in a one-shot game is to defect must be discarded. I want to show now that there is nothing wrong with the backward induction argument and that, therefore, this assumption must indeed be discarded. The argument is as follows.

If (1) the only rational thing to do in a one-shot prisoner's dilemma game is to defect,

then (2) similarly, defection is the only rational thing to do in the last shot of a finitely iterated game. This is because in all relevant respects, the last game of any finite sequence is the same as a one-shot game. From the point of view of rationality, the crucial characteristic they have in common is that, regardless of how they are played, it will not have a bearing on any further games. Therefore, they stand or fall together. More specifically, there being no further considerations to take into account, both games will have to be played on their own merits. If defection is the only rational strategy in one, it must be the only rationally justifiable strategy in the other.

(3) But if the rational strategy in the last game of a finite sequence is to defect, then for all rational intents and purposes the penultimate game becomes a *de facto* last game in that sequence, and hence, as a one-shot game in which one should defect. This is because, if in the last game all players must rationally defect on independent grounds, there can be nothing that could rationally warrant cooperation in the penultimate game. Defection in the last game is already a foregone conclusion and cooperation in the penultimate game is not going to make a rational difference to change that.

(4) But if the rational strategy in the penultimate game is to defect, then, by the same reasoning as before, the game preceding it must be treated as if it were the last game of the sequence, and hence, as a one-shot game in which one should defect.

Therefore, by the same, if somewhat tedious, reasoning all the way backwards,

(5) all games of a finite series must rationally be treated as the final game, and hence, as a one-shot game in which all players must rationally defect.

There is nothing discernibly wrong with this argument, even if it has, what Poundstone rightly describes as, "a distressing conclusion."²³⁷ Accordingly, the

²³⁷Poundstone 1992, p. 229.

argument enjoys widespread acceptance. "The prevailing opinion has long been that it is 'valid' in some abstract sense but not practical advice."²³⁸

Notwithstanding, as with virtually every philosophical argument and position, this argument also has its critics. Pettit and Sugden, for example, claim to have found "the solution" to the backward induction paradox.

The solution consists in recognizing that neither player is in a position to run the backward induction required. ... Taken together, these arguments yield the conclusion that the partner will defect in every round; thus, the only rational response is to defect in every round too. But it is mistaken to think that a player is in a position to run these arguments before making his first move. At this point, he believes that his opponent is rational, he believes that his opponent believes he is rational, and so on; this we have already assumed. But that does not entitle him to believe that in subsequent rounds his partner will still believe he is rational (etc.), irrespective of how he, the first player, has acted in the interim.²³⁹

This argument, and indeed the entire article in which it appears, misses the point of the backward induction argument in more than one way. To start with, if any assumptions are to be made at all about the players' beliefs about each other's rationality, it can simply be the assumption that both players correctly believe each other to be fully rational throughout the game, from start to finish. The question is posed then as to what they should rationally do in each of the finitely iterated sequence of prisoner's dilemma interactions. The only possible, and inescapable, game theoretic answer is via backward

It has been suggested to me by Michael Tooley that the above argument holds only in cases where the exact number of games of a finite sequence is *known* to the players. Even if this were the case, as Tooley also agrees, the conclusion would still be "distressing." I do think, however, that the argument holds for all finitely iterated sequences. I do not wish to deny that psychologically, or in practice, ignorance of when the sequence is going to end can make it easier for the players to overlook the fact that the sequence is finite. It seems to me, however, that on a theoretical level, in principle, the mere knowledge of the sequence being finite is enough to kick off the backward reasoning in question. Given an assumption of finitude, we are perfectly entitled to make an arbitrary assumption (for the sake of argument) as to how large the sequence is and complete the requisite reasoning backwards all the way to the first game of the sequence. It helps to keep in mind in this regard that the backward induction argument is not so much a problem for practice as it is for theory - a point which is going to be stressed further below as well.

²³⁸Poundstone 1992, p. 229.

²³⁹Pettit and Sugden 1989, pp. 171 - 72.

induction which prescribes defection in each round of the game. Pettit and Sugden agree that under conditions of common knowledge of each other's rationality, both players can run the requisite backward induction and "the solution fails" because "each will therefore defect."²⁴⁰ They argue, however, that under the same conditions the backward induction paradox itself fails.

[I]f the solution to our paradox fails under common knowledge of rationality, so does the paradox itself. The paradox arises from the conflict between the backward induction argument for permanent defection and the more intuitive thought that a player might do better to play a conditional strategy involving an initial act of cooperation. But this intuitive thought depends on a particular way of thinking about games which, although entirely natural, is not allowed by the assumption of common knowledge.²⁴¹

Throughout their paper, Pettit and Sugden argue as if the backward induction argument posed a practical problem for people in real life prisoner's dilemma situations. The problem posed by the argument however, is not so much practical as it is theoretical. Actual players are hardly at a loss as to what they should do in an iterated game. They usually settle on mutual cooperation fairly soon into the game because of the recognition that no alternative will serve their interests as well as that. To quote Poundstone,

Morton Davis said that the average person's reaction to the prisoner's dilemma is not so much to ask what you should do but rather to ask how you justify cooperation. Going by the literature, that's been the reaction of a lot of game theorists, too.²⁴²

Thus, the problem presented by the backward induction argument is one for theory, not so much for practice. If this is correct, it is hard to see how the problem of backward induction should cease to be a problem for game theory. The argument shows that rational choice (game) theory is inadequate and self-defeating in the situation because the logic of that theory would require rational players to act contrary to what otherwise is so obviously in their interests. Laying the blame instead, as Pettit and Sugden do, on the

²⁴⁰Pettit and Sugden 1989, pp. 180 & 181.

²⁴¹Pettit and Sugden 1989, p. 181.

²⁴²Poundstone 1992, pp. 222 - 23.

players' knowledge of each other's game theoretic rationality is misguided. Such knowledge need not be assumed at all by the backward induction argument.

This, in fact, is another, and more serious, way in which these authors miss the point of the argument under discussion. Pettit and Sugden mistakenly assume that the backward induction argument relies on the parties' knowledge of each other's game theoretic rationality at every point of the game. The formulation I have given above shows that this is not the case. The backward induction argument merely extends the logic of game theory's prescription for defection in a one-shot game. In doing this, beliefs about the players' game theoretic rationality or irrationality simply do not figure. In a one-shot game each player is recommended to defect regardless of what the other player believes or does because that will see each of them better off regardless of what the other player decides to do - and regardless of whether they are rational or irrational in doing what they do.

On the assumption that the game theoretic logic of defection proffered in a one-shot game is tenable, the backward induction argument merely shows that the same logic holds good for each single shot of a finitely iterated sequence. Nowhere in the argument need there be a reliance on the players' beliefs about their opponent's rationality or irrationality. One's opponent's beliefs, or actions, simply do not count. The recommendation is that one should defect, regardless of what the other person does. If the other person decides "irrationally" to cooperate at certain points of the game, then that is all the better for the defector. This reasoning is part and parcel of rational choice theory's logic in prescribing defection in a one-shot game - a logic extended by the backward induction argument to finitely iterated sequences. By the lights of game theory, an expectation that one's opponent is going to be irrational, and hence is going to cooperate, the temptation payoff provides all the more reason for one to defect.

Moreover, since the backward induction argument does not assume any beliefs about opponent players' beliefs or actions, the relevant reasoning which would require players to defect throughout a game *is* available to them. For players without the relevant

theoretical background, the day may simply be saved by not being aware of the reasoning in question. For those aware of the logic of the backward induction argument, however, the way out of the dilemma is to acknowledge that rational choice (game) theory cannot account for, or prescribe, the intuitively and obviously sensible, and indeed rational, course of action, which is mutual cooperation, not mutual defection. More strongly, but by no means without justification, that theory must bear the cross of being self-defeating in prisoner's dilemma situations because it prescribes mutual defection which is clearly against both players' interests.

I do, of course, agree with Pettit and Sugden that a conditional strategy such as TIT FOR TAT is a highly rational way of playing iterated prisoner's dilemmas.²⁴³ I also agree with them that

it will be in the partner's interest to cooperate - in the case of tit-for-tat, in every round except the last ... [because] in every round, except the last, each player has to take account of the possibility that his choice of move may influence the later moves of his partner.²⁴⁴

Where I disagree with them is the idea that the rationality of such cooperation can be explained compatibly with the dictates of rational choice (game) theory. What we have here is *strategic rationality* of a kind which not only goes beyond game theory's narrow prescriptions but is in fact in outright conflict with those prescriptions. Hence the so-called backward induction paradox, which is the most obvious symptom that there is something wrong with the logic employed by game theory in prisoner's dilemma situations.

This is what the backward induction argument shows, but the message is lost not only on Pettit and Sugden but also on game theorists in general who accept the argument as valid. It is also lost on another prominent critic of the backward induction argument, Russell Hardin.

²⁴³I consider the rationality of TIT FOR TAT patterns of behaviour in more detail in the following chapter.

²⁴⁴Pettit and Sugden 1989, p. 170.

The conclusion of the backward induction argument has become a virtual dogma despite the fact that many, perhaps most, discussants think it perverse. Its appeal as a dogma may simply be that it is cute and perversely contrary to common sense.²⁴⁵ [But] Whatever the apparent force of the backward induction argument for rarefied game theorists, it appears that actual people in going societies regularly take the risk of initially cooperating to upset that argument. Only for that reason do we have going societies.²⁴⁶

As can be seen from these remarks, Hardin is no fan of the backward induction argument which he thinks "is flawed."²⁴⁷ Significantly, however, he fails to point to any specific flaws in the argument. Although he rejects the argument, it is not clear whether the flaw is supposed to lie with the validity of the argument or the truth of the premises. This leaves him open to the charge that he simply does not like the argument's conclusion.

Hardin claims to be able to wreck his opponent's backward induction "by simply cooperating at our first encounter."²⁴⁸ He seems to expect that as a result his opponent will reconsider the logic of the backward induction argument, or at the very least will start cooperating. However, if we accept the assumption in (1), that the only rational thing to do in a one-shot game is to defect, then Hardin's expectations are unrealistic. Being a believer in (1) and also in backward induction, the opponent is likely to conclude that Hardin is irrational to have started off cooperating. Moreover, the opponent would have been rewarded for the defection in the first round with the handsome temptation payoff, while Hardin would have had to content himself with the sucker payoff for his cooperation. Sitting pretty after the first game, it is hard to see why the opponent should be willing to cooperate in the second round. By his lights, that would be irrational and his early success with defection bears witness to the effectiveness of his chosen strategy. Hardin could, of course, start putting pressure on his opponent by starting to defect as

²⁴⁵Hardin 1991, p. 206, fn. 2.

²⁴⁶Hardin 1991, p. 188.

²⁴⁷Hardin 1991, p. 187.

²⁴⁸Hardin 1991, p. 188.

well, but even then it is not clear that his opponent would rationally have to change tack. If he accepts (1), and also the backward induction argument, he will simply conclude now that, although Hardin is a slow learner, he has finally smartened up to the game and started doing what he should have been rationally doing from the very beginning, which is to defect. So this line of attack on the argument is unconvincing.

Hardin does succeed, however, in presenting a convincing case for rejecting the conclusion of the backward induction argument. He invites us to ponder why we have a tendency initially to trust people whom we do not yet know well. He also points out that all our relationships with people are of necessarily finite duration. He convincingly reasons that, given this state of affairs, since “[t]he backward induction argument recommends initial distrust and, further, continued distrust”, “[t]his is surely a recommendation for slow death by abnegation.”²⁴⁹ It is hard to disagree with this argument which does lend a strong support to Hardin’s position that trust and cooperation are eminently rational. The rub, however, is that this still does not tell us where the backward induction argument goes wrong.

If the backward induction argument was construed as a categorical argument, I could not disagree with Hardin that it was flawed, because I fully agree with him that, at least under certain conditions, trust and cooperation are fully rational. There is nothing wrong with the argument, however, if it is construed as a conditional argument, the way I have done above. Construing the argument conditionally is more profitable, I believe, because the crucial premise of the argument, namely (1), is not only a thesis which is held independently of the argument, but also a thesis that stands as the cornerstone on which the rest of that argument is built. Premise (1) contains the dubious prescription of rational choice theory for prisoner’s dilemma situations. This is that in a one-shot game the only rational thing is to defect. If this most critical part of the backward induction argument is granted, all is lost because the rest of the argument cannot be prevented from

²⁴⁹Hardin 1991, p. 188.

going through. The argument is deductively valid and it is hard to see how the other premises could be faulted, if (1) is accepted as true.²⁵⁰ Therefore, the earlier conclusion remains: Since there is nothing wrong with the validity of the backward induction argument, and since its conclusion is untenable, we must discard the dubious prescription of rational choice theory assumed in premise (1), namely, that the only rational thing to do in a one-shot game is to defect.

This failure of rational choice theory in prisoner's dilemma situations challenges the popular view which regards its recommended rational choice calculations as somehow being the pinnacle or epitome of rationality. I suggest that in cases where rational choice theory delivers the obviously wrong and self-defeating advice as to what the rational choice is, it is irrational to adhere to it. If there is an air of paradox in this claim, then that is only testimony to the inappropriateness of defining rationality by reference to the narrow canons of that theory. By contrast, indirect rationality operating through codes and norms of conduct, through entrenched attitudes of trust, honesty, non-exploitation and fairness, has the capability to ensure the optimum outcome for all. In other words, the rationality of cooperation in prisoner's dilemma situations can be explained by reference to the indirect rationality of being nice and cooperative unless provoked by defections.

This is not to say that the rationality of cooperation can only be accounted for by reference to indirect rationality alone. The failure of rational choice theory in its dominant form may not preclude the development of different accounts of direct rationality. The prevalent view among game theorists, however, is far from optimistic. The depth of this pessimism can be seen in the comment by Luce and Raiffa that "The hopelessness that one feels in such a game as this cannot be overcome by a play on the words 'rational' and 'irrational'; it is inherent in the situation. 'There should be a law against such

²⁵⁰It is uncertain from Hardin's arguments if he would agree to challenge directly the truth of (1). However, many of his comments, including his insistence that trust and cooperation are eminently rational, point in that direction.

games!”²⁵¹ In spite of enormous efforts over the last half a century to find a solution to the dilemma by showing, somehow, that cooperation is rational after all, Poundstone recently had to sum up the state of the game as follows.

In a true, one-time-only prisoner's dilemma, it is as hard to justify cooperation as it is to accept mutual defection as the logical outcome. Therein lies the paradox. Both Flood and Dresher say they initially hoped that someone at RAND would “resolve” the prisoner's dilemma. They expected Nash, von Neumann, or someone to mull over the problem and come up with a new and better theory of non-zero-sum games. ... The solution never came. Flood and Dresher now believe that the prisoner's dilemma will never be “solved,” and nearly all game theorists agree with them. The prisoner's dilemma remains a negative result - a demonstration of what's wrong with theory, and indeed, the world.²⁵²

I suggested earlier that the rationality of cooperation in a prisoner's dilemma situation can be adequately accounted for by a theory of indirect rationality. If the views expressed above are correct, then they only prove that the adequacy and importance of direct rationality in general, and especially those of rational choice theory in particular, have been overestimated and overemphasised. The fact that a theory of indirect rationality can provide a solution to the prisoner's dilemma is a strong point in its favour and is the best demonstration of its normative force.

But, before leaving this topic, it is worth taking another look at the pessimism regarding the chances of finding a solution to the prisoner's dilemma by way of an alternative theory of direct rationality. What should be noted here is that, since the underlying grounds for the rationality of being cooperative is hardly a mystery, the outlines of an alternative theory should already be visible. It is all too plain that in prisoner's dilemma encounters only mutual cooperation can be a winning strategy for those concerned, while defections are the sure way to, what Hardin has referred to as, “slow death by abnegation.” Resorting to somewhat more technical language, this means that the game has a *dominant point* which can act as the starting point for a different

²⁵¹Luce and Raiffa 1957, pp. 96 - 97.

²⁵²Poundstone 1992, pp. 122 - 23.

account of direct rational choice.²⁵³ Upon closer examination of the prisoner's dilemma it becomes evident that rational choice theory focuses far too narrowly on dominant individual strategies. In this particular game, however, dominant strategies can only ensure that players avoid the sucker payoff regardless of what their opponents decide to do. This focus is obviously too narrow and is definitely not sophisticated enough to outwit the trap posed by the dilemma. This is a tricky game in which the dominant strategy draws the players away from the dominant point of the game and this should be reason enough to reconsider the wisdom of staying fixated on the dominant strategy alone.

A well-rounded, fully rational person with judgement, who makes reasoned assessment the basis of action, will recognise this trap quickly and will try to shift the focus away from the dominant strategy to the dominant point of the game. If a solution based on direct reasoning is at all possible, such a shift in focus is what must hold the key to it. There is plenty of evidence that even under strict laboratory conditions where there can be no coercion to force cooperation onto the players, the need for this shift in focus seems to be quickly recognised, and is accordingly exploited for mutual advantage, by the majority of players. This is true even of those who are unfamiliar with the theoretical aspects of the game²⁵⁴ and there can be no doubt that this change in focus is crucial to achieving and explaining cooperation in prisoner's dilemma situations. The suggestion by Rapoport and Chammah that this shift in focus is due to an "intellectual short-coming" in the players is but a testimony to the short-comings of their own conclusion. Laboratory results clearly show that, initially, most players start off with defections and tend to change to mutual cooperation as the game progresses. Given this

²⁵³By 'dominant point' I mean a situation of compromise in which both players still try to maximise their individual interests as much as is permitted by the situation but without either of them expecting or exacting a cost from the other which they themselves would not be prepared to pay. In a prisoner's dilemma this is the situation afforded by mutual cooperation. Understood in terms of a strategy, the idea is equivalent to the dominant strategy for action by a group. The concept of collective rationality is explored in more detail further below.

²⁵⁴Lave 1962.

pattern, what Rapoport and Chammah are in effect saying is that the more dumbly the players behave, the more it profits them. To say the least, this is an extraordinary conclusion, a classic case where interpretation of the observed phenomena is severely distorted by an undue and unquestioning faith in the adequacy of rational choice theory. Surely, the correct conclusion is that players do initially fall into the trap of playing their dominant strategies, but soon realise that they are thereby defeating their own purposes. They look around for alternatives and quickly settle on the dominant point of the game by mutual cooperation.

What is crucial to bear in mind here is that this shift in focus pays, and players choose it for that very reason - an unmistakable sign of intelligence, one should think, rather than the lack of it. Moreover, since their shift in focus is a result of a direct decision to pursue the obviously much more profitable strategy, it can be properly regarded as a *direct rational choice*. The challenge to game theorists is to find an alternative explanation for the rationality of that choice. By no means do I want to suggest that this is a simple and effective solution to all aspects of the dilemma²⁵⁵ but dismissing such direct, interest-motivated and obviously successful choices as the result of an "intellectual short-coming" is both implausible and unhelpful in motivating alternative solutions to the problem of how we should explain cooperation. Against the above suggestion of Rapoport and Chammah, we must keep in mind the earlier observation that the general reaction to the prisoner's dilemma is not so much to ask what you should do but rather to ask how you justify cooperation.

²⁵⁵It is not clear, for example, how the free rider problem would be solved along the lines of the suggestions I have made, since in a prisoner's dilemma with many players, it would be very hard to motivate individually the required shift in focus away from the dominant strategy of the individual players involved. By comparison, it seems that this shift is much easier to motivate in games with only two players. It is also worth noting here that not every N -party prisoner's dilemma is a free rider problem. As Pettit has convincingly argued, some of them constitute a foul dealer problem which is significantly different. (Pettit 1986.) While the defection of the lone free rider is a mere irritation to the others for whom cooperation still remains profitable and desirable, the foul dealer's defection makes their cooperation harmful to themselves because they end up worse than they would be under mutual defection. "The free rider seeks to benefit by the efforts of others, the foul dealer to benefit at their expense." (Pettit 1986, p. 374.)

But, whether the reasons and explanations behind particular cases of mutual cooperation amount to direct rationality or not, there is no denial that mutual cooperation in a prisoner's dilemma situation is indirectly rational in the sense that it best serves the individual interests of the players. This reinforces the claim that rational action or conduct is all that we need to satisfy a prudential injunction to be rational, and that it is the most that can be reasonably required of anyone. As long as one's conduct is rational in the sense that it best serves the individual interests of the players concerned, the question whether it is the direct result of rational choice deliberations or not, though interesting in itself, remains largely academic.

There is no reason why revenge should be an exception in this regard. As long as some instance of revenge can be shown to be rational, in the sense that the action in question is in line with the agent's rationally held beliefs and desires, it is difficult to see how the fact that it is not the direct result of rational choice deliberations could be an objection against it. By implication, it is no objection to an act of vengeance if the best explanation for it is not via a direct, psychological, rational choice theory, but by reference to other sorts of explanations that appeal, instead, to such notions as fairness, justice, duty, loyalty, social custom, or norm-following. In other words, where reasons for and against revenge are concerned, the important question is not whether direct rational choice explanations provide the best explanation of actual cases of revenge.²⁵⁶ It is, rather, whether those acts of revenge are rational or not in the broader sense indicated above. I shall argue later in this chapter, and the next, that, with possible exceptions, the answer to this question is affirmative.

To sum up the discussions of this chapter so far, I have explained what rationality is and why it is important that revenge be rational. Following this I discussed the power of rational choice explanations on a normative level in terms of formal and substantive criteria. I argued that Elster's account of substantive rationality is flawed and proposed

²⁵⁶This question might still be important, however, for psychological, anthropological and predictive purposes.

that the substantive rationality of beliefs and desires is better defined in terms of relevance and agent interests. Finally, I made a distinction between direct and indirect rationality and argued that there is no general normative difference between them. I argued that in prisoner dilemma situations, where the interests of all those involved can be best served by mutual cooperation, direct rational choice deliberations are prone to be self-defeating. I also argued that, since the same problems are not present for indirect rationality, the latter is as good, and sometimes better, at meeting a prudentially grounded normative injunction to be rational.

4.3 Collective rationality and revenge

In discussing a subject, such as the rationality of revenge, a distinction can be drawn between individual and collective rationality. The individual rationality of revenge will be considered at length in the next chapter. The remainder of this chapter defends the collective rationality of revenge in terms of the common interests of the concerned collective.

Prisoner's dilemma situations are good examples where the individually rational choices of two or more individuals in the direct sense can lead to sub-optimal results for the group and its members - thus showing that direct individual rationality need not coincide with group or collective rationality, let alone the former automatically leading to the latter. Along the same lines Jackson argues that there is a morally significant difference between individual and group actions and that there are cases where morally wrong group actions are made up entirely of morally right individual actions.²⁵⁷ Finally, Arrow's controversial Impossibility Theorem (according to which there is no formula for aggregating three or more consistent individual preferences into a consistent ranking which will satisfy certain reasonable conditions) provides further reason for not

²⁵⁷Jackson 1987.

assuming that collectively rational social choices are automatically achieved through choices which are individually rational in the standard, direct sense.²⁵⁸

The notion of collective or group rationality, however, allows more than one interpretation. As already indicated, in the case of revenge it may refer to the collective practice or institution of revenge in a society where revenge is the expected means of ensuring social order and retributive justice, but it may also be taken to mean the rationality of a particular act of revenge taken by a group of individuals, such as members of a family, tribe or country for an offence committed against them as a group. The next section is a case study of group revenge of this latter kind and my aim is to demonstrate that the practice of *utu* (repayment) in the form of group revenge among the traditional New Zealand Maori can be plausibly explained by reference to the theoretical frameworks of rationality considered and discussed so far in this chapter. I argue that traditional Maori revenge meets the requirements of both formal and substantive rationality and that it is an example of collectively rational group action, in both the direct and indirect senses discussed above. Following this, I shall consider the collective rationality of revenge as an institutionalised practice.

4.3.1 The collective rationality of *utu* as group revenge

It is well known that the Maori of New Zealand gave overwhelming importance to the avenging of injuries. To avenge a wrong was held as a sacred duty. If immediate revenge was not possible, for example, some injuries were kept alive, sometimes for generations, until *utu* was obtained.²⁵⁹ However, if a tribe was confident in its ability to

²⁵⁸See, for example, Arrow 1977; Plott 1976; and other papers on the topic in Barry and Hardin 1982.

An interesting solution to the aggregation problem has been suggested to me by Michael Tooley. According to him, if we worked with quantitative measures of strengths of preferences, then the aggregation problem would look very different, and that Arrow's Impossibility Theorem would, so to speak, no longer arise. The idea here would be to add up the strengths of individual preferences to get a measure of the strength of the resulting social preference.

²⁵⁹Best 1952, p. 167.

take revenge, there was no room for hesitation. The determination and ferocity with which the Maori exacted their vengeance is clearly illustrated in the following story:

In bygone days, there was a settlement of some of the descendants of the Titahi folk near the present Whare-roa. These Titahi were a warlike people, and were constantly involved in disputes and conflicts with their neighbours. ... The Titahi people in some way insulted the inhabitants of Okahu-titi pa. ... For this reason, the Waingongoro people determined to seek revenge or payment, utu. A great vengeance was determined on, with the object of completely exterminating the offenders. ... In the dead of night, the war party arrived at the sleeping settlement of the Titahi folk, who had no night sentinels posted to warn them of impending danger. Silently, the Okahu warriors deposited their bundles of dry fern round the great sleeping-house. When all was ready, fire was applied to the fern, which immediately sprang into a blaze. Then the work of extermination commenced ... until not one of the Titahi remained alive. Nothing was left, save the blackened, smoking ruins. ... and the story of the vengeance exacted by Ngati-Okahu as utu for an insult.²⁶⁰

It is also worth, perhaps, briefly pointing out here the interesting link between revenge and cannibalism. Being eaten was considered to be the greatest possible disgrace that could befall a Maori and his tribe:

To kill a man and to partake of his flesh was to cast disgrace upon him and his people. It was done to satisfy a hatred, or to be revenged for a tribal injury or insult - the usual cause of war. ... The body would be eaten in order to bestow upon it the greatest possible indignity.²⁶¹

Given this picture of Maori revenge practices, the question is whether such extreme forms of group revenge can be shown to be rational. This question can, and must, be separated from the question whether Maori revenge practices are morally defensible. In this instance there is more to morality than sheer self-interested rationality. Indeed, it could be argued that, generally, an important function of morality is to place limits on the self-centred rationality of groups and individuals who in their pursuit of narrow self-interest may brush aside and place in serious jeopardy the interests of others. Therefore, even if Maori revenge practices can be objected to on such moral grounds as cruelty and lack of proportionality, it by no means follows that those practices cannot be rational in

²⁶⁰Houston 1965, pp. 172 - 73.

²⁶¹Del Mar 1924, pp. 55 & 56.

terms of the selfish interests of the tribe or the kinship group and their individual members. In order to see the most important way in which Maori revenge practices satisfy the requirements of direct group rationality, we need to clarify two important concepts of the Maori: *utu* and *mana*. Take *mana* first:

Mana is a (supernatural) power that can be present in a person, place, object, or spirit. It is commonly understood as prestige, power, or authority; but really such status is derived from possessing *mana*. ... Ultimately *mana* derives from the doings of agents: the transmission is originally from the High Gods to the people. ... In Maori terms it is a power of an agent in the absence of which no flourishing can take place.²⁶²

This last point is very important. *Mana* is valuable intrinsically but it is equally, if not more, valuable instrumentally. To have *mana* is to have the prerequisites for a meaningful, worthwhile life. Life without *mana* was the life of a dog, or more appropriately, the life of a slave. Understanding the concept of *mana* takes us a long way toward understanding the concept of *utu*. For although there is a rather unfortunate contemporary tendency to use the word 'utu' as if it was the Maori term for revenge, this is a mistake. The two concepts do not coincide. *Utu* is a wider concept than revenge. All cases of revenge may be *utu* but certainly not all cases of *utu* are revenge. Therefore, the key to understanding *utu* is not revenge. It is, rather, the already explicated notion of *mana*. Understanding this concept is essential for understanding the Maori and their actions in general. And, as the passage below also makes it clear, *mana* is the key for understanding the concept of *utu* in particular:

Insofar as *utu* means more or less the same as any English word, it means more or less the same as 'repayment' or 'compensation', restoring some sort of balance, exacting what is due, what is demanded by the situation. In some situations, what is traditionally demanded will be revenge. ... But other situations make other demands: sometimes a reward, sometimes a transfer of goods or cash, sometimes even an insulting song. When *utu* is taken it is always in response to some action, and aims at effecting at least a return to the prior position. ... Normally the members of a tribe live in a state of approximate perceived balance with outsiders. From time to time some action - friendly or unfriendly - upsets this balance. *Utu* is taken to restore the balance. ... [But]

²⁶²Perrett and Patterson 1991, pp. 192 - 93.

to say that *utu* consists in restoring balance is to tell only part of the story. There are several dualities in Maori life which need to be kept in balance, but not all of these involve *utu*. ... [In the case of *utu*] it is *mana* that has to be brought back into balance. When outsiders injure a member of a tribe, or present a valuable gift, this changes the *mana* of each party involved. The injured party loses *mana*, the other party gains; the party which makes the gift gains *mana*, the recipient loses. So from the point of view of the party whose *mana* has been reduced, something has to be done. Whatever is done to restore that *mana* is the *utu*.²⁶³

Insulting someone in return for a gift is not *utu*, of course, but a serious provocation. The crucial principle of exchange here is reciprocity in all things. In *utu*, normally, the repayment has to be in kind: a gift has to be repaid with one of *at least* equal value, preferably of greater value, and harms are, with some exceptions, repaid with harms in a similar fashion. Recognisably, this is a TIT FOR TAT pattern of behaviour which, in light of Axelrod's work on the evolution of cooperation, is a matter of considerable significance.²⁶⁴ The potential and actual advantages of not letting others get away with one-sided, exploitative behaviour are obvious enough. The only concern which remains from the point of view of rationality, therefore, is that the costs involved should not exceed the advantages gained. Axelrod's findings, to which I shall return in more detail in the next chapter, clearly show that in iterated prisoner's dilemma situations where the future is sufficiently important relative to the immediate payoffs, the strategy of repaying like with like (kindness with kindness, hostility with hostility) is the best way to elicit the cooperation of others and thereby, on the whole, the most efficient way to do well for oneself. In the light of Axelrod's findings, the recognition that *utu* is a TIT FOR TAT type pattern of behaviour provides good grounds for regarding the Maori practice of *utu* as having been rational in the indirect sense described above.

But *utu*, of which revenge was an essential part, was rational in the direct sense as well. The primary aim of revenge had always been the restoration of the tribe's *mana*. This was of paramount importance because survival itself depended on it. Recall that in

²⁶³Patterson 1989, p. 53. See also Patterson 1992a, Ch. 5; Patterson 1992b.

²⁶⁴Axelrod 1984.

Maori understanding no human flourishing can take place in the absence of *mana*. Injuries and insults that seriously harmed a tribe's *mana*, threatened the tribe's prosperity and existence. If an injury or insult was serious enough, although other options were often open, in a great many cases revenge was the best way of obtaining *utu*. The key idea here is *mana*-maximisation. The more severe the response to an offence, the more *mana* was gained. If circumstances allowed, the ultimate retaliation which was meted out to serious offenders was to eat them and make use of their bones and tattooed skins in the form of such mundane items as fish hooks, needles, coat hangers and toys.²⁶⁵

A disturbing feature of Maori group revenge is its lack of concern for considerations of proportionality. Though this may be morally indefensible in the light of contemporary moral standards, this feature can be explained from the point of view of narrow, self-serving rationality. The individual tribes were not only politically autonomous from each other but they saw each other as competitors and rivals with whom they lived in perpetual tension. Raids on other tribes were frequent. Consequently, considerations of moral principles, such as the principles of justice were not extended to rival tribes, especially not at times of inter-tribal raids and warfare. As I suggested earlier, in the absence of limiting moral principles, self-serving rationality rules. In the traditional Maori context, the more devastating a tribe's retaliation was on a rival, the better it was for the tribe - and not only in terms of *mana*-maximisation but also in terms of the selfish material interests of the tribe concerned.

In virtue of the fact that pre-European New Zealand was a land of limited and often scarce but vital resources, overdoing the revenge benefited the tribe in very practical and tangible ways, and to an extent to which no other form of response to an injury could have ensured. Among the obvious resources are weapons, tools, clothes, food and land. *Muru* (plunder) had invariably been at least a secondary motive for inter-tribal warfare. Raids which were primarily motivated by the desire to obtain *utu* in the form of revenge

²⁶⁵Del Mar 1924, pp. 55 - 56.

were certainly not immune to this factor. In addition, the benefits to be gained by weakening, enslaving, or completely eliminating one's competitors for the scarce material resources are all too obvious. In this regard recall the great vengeance of the Okahu people on the Titahi. The Titahi were a constant source of unrest, irritation and insecurity to the Okahu and other neighbours. By totally exterminating them for an insult, the Okahu not only gained control of much wanted natural resources (so far controlled and used by the Titahi) but they also increased their own security by eliminating a dangerous neighbour. By overdoing their revenge as a group, they paved their own way to increased security, prosperity and growth in a way in which no practical alternative could have ensured. These considerations should leave no doubt that Maori group revenge was not only rational in the formal sense but in the proper, substantive sense of the word as well.

Before moving on, it is worth noting that the collective rationality in the direct sense of such instances of Maori group revenge is not challenged in any way by Arrow's earlier mentioned Impossibility Theorem. From what we know of Maori values and attitudes of the time, it is reasonable to assume that members of the avenging tribe were virtually unanimous in their preferences concerning the great revenge referred to above. If, and since, the individual members of the tribe acted with one mind, the aggregation of their individual preferences into a consistent ranking for collective rational decision making should be unproblematic.

4.3.2 The collective rationality of revenge as an institutionalised practice

While not everybody is convinced by functionalist explanations of revenge as a collective practice,²⁶⁶ it is fairly well recognised that revenge practices in pre-industrial societies which have no centralised law enforcement agencies, serve important functions equivalent to functions served by contemporary legal justice systems.

²⁶⁶See, for example, Elster 1990.

[T]he practice, alike of immediate revenge and of postponed revenge, establishes itself as in some measure a check upon aggression; since the motive to aggress is checked by the consciousness that a counter-aggression will come: if not at once then after a time. Among human beings in early stages, there hence arises not only the practice of revenge but a belief that revenge is imperative - that revenge is a duty.²⁶⁷

Similarly, speaking of the indignation or resentment at wrong which “arises naturally and spontaneously in the human mind,” Rashdall speaks of the

profound ethical conviction that for societies - though not always for individuals - it is morally good and healthy that this indignation should be encouraged and expressed.²⁶⁸

Again, Boehm makes the following observations about the (informal) institution of the feud among the Montenegrins:

In many respects the morality of several thousand people who are living in a permanently settled tribal territory is unlike our own morality. There is less concern for legalistic manoeuvring, and more for personal and clan reputation. There are no specialized police officers or judges, nor are there any prisons. But there are very powerful sanctions that shape behavior, and these operate both directly and indirectly, intentionally and automatically. Feuding was essentially a positively valued institution insofar as the moral system was concerned, in that it involved the upholding of honor. In addition, ... feuding served as a kind of sanction, because it suppressed certain immoral behaviors that people knew were likely to start feuds. They also knew that feuds were dangerous, stressful, economically costly, and generally inconvenient from a practical standpoint. ... [T]he effects of feuding as a social sanction were highly significant for this very aggressive people, who deliberately chose to live their lives with the near absence of any coercive authority in human form. ... [F]euding served as a substitute for such authority, in that the probability of lethal retaliation and then a costly feud sharply curtailed certain socially disruptive behaviors.²⁶⁹

As evidenced by these passages, considerations of the collective rationality of revenge as a socially enforced institution or practice tend to be mingled up with considerations of its morality. The difficulties of keeping the two spheres apart are due to the fact that the rationality of revenge as a practice is evaluated by reference to the common good which, recognisably, is a moral concept. Moreover, in the case of

²⁶⁷Spencer 1966, p. 362.

²⁶⁸Rashdall 1924, p. 304.

²⁶⁹Boehm 1984, pp. 88 & 183.

revenge as an institutionalised practice, the notion of the common good includes not only such considerations as public safety and security (to be promoted by means of deterrence), but also considerations of such highly esteemed, non-instrumental moral values as retributive justice and honour which are universally shared among members of the societies in question.²⁷⁰

While the collective rationality or morality of particular revenge practices such as the Montenegrin blood feud may be controversial, it is not at all clear that, in principle, such informal social institutions are inferior to, say, contemporary systems of criminal justice. Supposing that a collective of freedom loving individuals needs to find some kind of (formal or informal) mechanism for achieving the above mentioned array of values and goals, and given that they prefer to do this cost-effectively and without the creation of a centralised state or authority, it is hard to see what better options there could be than victim-based private systems of law enforcement for which we have creditable historical precedents.²⁷¹

As against the idea that legal justice systems such as those currently operating in contemporary Western societies are rationally preferable to victim-based revenge practices, it will be enough for the moment to keep in mind that the former are not only prohibitively expensive but also questionable in terms of efficiency and integrity, as pointed out in the preface. Corruption and abuse of power in police forces, for example, are ever present public concerns and it is hardly an exaggeration to say that the ongoing tolerance many contemporary legal systems show towards lawyers' excessive preoccupation with winning court cases, rather than finding out the truth of the matter, is a mockery of justice.²⁷² A more detailed case for the collective rationality of allowing certain institutionalised forms and practices of revenge will be provided in the last chapter where I defend the morality of victim-oriented legal institutionalised systems of justice.

²⁷⁰A defence of these moral values will be undertaken in subsequent chapters.

²⁷¹These will be discussed in more detail in the last chapter.

²⁷²See, for example, Dickie 1989; Kennedy 1989; Whitton 1994.

4.4 Conclusion

To sum up the discussions of this chapter, I have explained what rationality is and why it is important that revenge be rational. Firstly, I showed that rationality is important because, in general, rational behaviour has the best chance of enabling us to pursue and achieve our short-term and long-term goals. I then discussed rational choice theory and argued that if it is to have normative strength, the purely ends-means conception must be supplemented by appropriate evaluations of the agent's goals and beliefs. Following this, I distinguished between direct and indirect rationality. Behaviour is directly rational if, as a direct result of a rational choice calculation, it is the best means of achieving the agent's goal. Behaviour is indirectly rational if it forms part of a rational habit, disposition, attitude, custom or social norm. I further argued that it is not significant from a normative point of view whether one's behaviour is directly or indirectly rational. What matters is not that we constantly engage in cost-benefit analyses, but that we do not jeopardise our short-term and long-term interests. Having rational habits, dispositions, attitudes, customs and norms is enough to ensure that this does not happen. I also made the suggestion that problems encountered by rational choice theory in situations, such as the prisoner's dilemma, can be satisfactorily overcome only by a theory of indirect rationality.

These arguments and discussions have also prepared the ground for the next chapter where I argue in detail for the formal and substantive, as well as for the direct and indirect, rationality of revenge behaviour by reference to the self-regarding material and non-material interests of individual agents. They also provided a context against which I considered the two senses in which revenge could be collectively rational. With reference to the collective rationality of particular group actions, I argued that, given the importance of *mana* for the traditional New Zealand Maori, and given the general conditions prevailing in pre-European New Zealand, there is good evidence that in a great many situations Maori group revenge was perfectly rational in the required sense of the word. I argued that traditional Maori revenge behaviour satisfied the requirements of

both direct and indirect rationality, as well as rationality's formal and substantive requirements. I suggested that this was so regardless of whether the range of proper goals was restricted to strictly material interests in the way preferred by Elster or not. Finally, I considered the collective rationality of revenge practices as social institutions. Relying on commonly accepted functionalist explanations of revenge practices, I made the observation that social institutions such as the feud served the same kind of functions in societies in which people preferred to deal with conflict on a personal level without involvement of a centralised state as legal justice systems do today. Pointing to some of the more obvious drawbacks of modern justice systems, such as high cost, questionable integrity, and the lack of interest in the courtroom by professionals of the law to discover the truth of the matter, I suggested that victim-based institutions of justice amounting to practices of revenge were not obviously inferior choices, given the combined aim of promoting deterrence, honour and retributive justice at an economically competitive cost. The case for the collective rationality of certain forms of institutionalised practices of revenge will be further supported by the arguments of the last chapter where I argue for adopting a more victim oriented legal institutionalised system which in the interests of victim justice allows for certain limited forms of institutionalised revenge.

5

The individual rationality of revenge

5.1 Introduction

In this chapter I present Elster's attack on the rational status of revenge. In order to motivate the discussion, I consider the most important objections against the possibility, and reality, of individually rational revenge behaviour. My focus throughout will be to defend an affirmative answer to the question as to whether revenge can be rational in terms of the individual agent's self-regarding interests.

According to Elster, revenge behaviour cannot be rational from the individual agent's point of view. I isolate four crucial areas where Elster's arguments are mistaken and then proceed to discuss them in detail. In section 5.3 I consider the argument that revenge is irrational on account of its backward-looking nature. In section 5.4 I call into question Elster's claim that honour is the key to understanding all cases of revenge. I also argue for the defensibility of honour as a proper goal by reference to which honour motivated revenge can be shown to be directly rational in a substantive sense. In section 5.5 I deal with a suggested dilemma for the rationality of revenge presented by the material/non-material distinction. I defend the rationality of revenge behaviour with respect to both the material and the non-material interests of the individual agent.

5.2 The case against revenge

As I mentioned in the introduction to the previous chapter, only Elster seems to have made a point of arguing against the rationality of revenge in any serious detail. I start by presenting Elster's position, quoting key passages wherever possible, and then move on to discuss in detail the rational status of revenge.

Elster begins by defining revenge as

the attempt, at some cost or risk to oneself, to impose suffering upon those who have made one suffer, because they have made one suffer, ...²⁷³

He correctly points out that revenge is a "universal phenomenon" and proposes to "discuss proximate and ultimate causes of this behaviour" with a view to finding a satisfactory explanation for it.²⁷⁴ One possible explanation which he considers, and rejects, is that revenge behaviour is individually rational for people to engage in. This is the only explanation which I want to consider and defend. I want to show that, not only is it possible for revenge to be rational for the individual agent in terms of self-regarding interests, but that generally, real life individual revenge behaviour does meet the requirements of rational choice theory. I shall not concern myself too much with alternative functionalist explanations discussed by Elster, which appeal to such notions as community interests and reproductive fitness. For current purposes at least, their discussion is largely unnecessary. As Elster also acknowledges, "if revenge can be shown to be rational, no further explanation may be needed."²⁷⁵ This is certainly true in so far as the current project is concerned, which is to show that there can be no blanket

²⁷³Elster 1990, p. 862.

As Roy Perrett pointed out to me, Elster is mistaken in supposing that revenge necessarily requires the suffering of the revengee. It is possible, for example, to kill someone painlessly in revenge - in which case, even though harmed, the person would not strictly speaking suffer.

²⁷⁴Elster 1990, pp. 862 - 63.

²⁷⁵Elster 1990, p. 863.

normative injunction against revenge based on the standard of individual rationality.

Against the viability of such a project, Elster puts up the following case:

Rationality is conditional and future oriented. Its imperatives are hypothetical, that is, conditional on future outcomes one wants to realize. ... Rationality says: If you want Y, do X.²⁷⁶

People can act in a rational, [future] outcome-oriented manner, choosing the best means to achieve their ends. Prima facie, this motivation is incompatible with revenge behavior. The very definition of revenge ... shows that it involves only costs and risks, no benefits. Rational individuals follow the principles of letting bygones be bygones, cutting their losses and ignoring sunk costs, but the avenger typically refuses to forget an affront or harm to which he has been exposed.²⁷⁷

If broader motivations [than the selfish, material interests of the agent] - like the concern to uphold one's honor - are allowed, the task of proving the rationality of revenge becomes easier but also less interesting. Whenever a norm is invoked in explaining an action, one can turn the account into a rational-choice explanation by saying that the agent is concerned with upholding his self-image and reputation as a norm follower. While true enough, such statements are uninformative. In particular, they contain no guide to how the agent could rationally choose between alternative means to the stipulated end. Many of the means, in fact, will themselves be regulated by norms. It seems better, therefore, to make a clear break and stipulate that revenge, to be rational, must be so in terms of material ends.²⁷⁸

Following this last passage, Elster goes on to consider three arguments for the rationality of revenge from the individual agent's perspective. Finding these arguments wanting, his conclusion is that "revenge behaviour cannot be shown to be individually rational."²⁷⁹ As to the possibility that there might be other arguments to establish the rationality of revenge, Elster's paper seems to contain only the following as an answer:

More generally, I do not believe rational-choice arguments can capture the phenomenon of honor, which figures so prominently in all accounts of revenge.²⁸⁰

²⁷⁶Elster 1990, p. 863.

²⁷⁷Elster 1990, p. 862.

²⁷⁸Elster 1990, p. 872.

²⁷⁹Elster 1990, p. 876.

²⁸⁰Elster 1990, p. 876.

There are several problems with Elster's argument and I shall discuss them in detail. The following is a list of points on which, I argue, Elster is mistaken:

- 1) The claim that, by nature, revenge is backward-looking and that the backward-looking nature of revenge is evidence against its rationality.
- 2) The claim that honour is universally present in all accounts of revenge.
- 3) The claim that in material terms revenge cannot be in one's interests, that revenge can only involve costs and risks, no benefits.
- 4) The claim that with respect to non-material interests and goals rational choice explanations of revenge are unsatisfactory because they are uninformative.

I discuss and argue against each of these claims. The last two of them are suggestive of a dilemma and will be considered as such.

5.3 The backward-looking element in revenge and rationality

The first crucial claim by Elster is that, by nature, revenge is backward-looking and that this is evidence against its rationality. One interpretation of this claim is that the backward-looking feature of revenge, consisting of the fact of a past offence, counts against its rationality. This way interpreted, the claim is mistaken. In this regard at least, revenge is not more backward looking than, say, judicial punishment. This kind of backward-lookingness is a necessary precondition for not only revenge, but also quite generally for any other kind of punishment. Even punishment of the purely consequentialist sort would be a logical impossibility if there was not some past offence or wrong to punish in the first place. Even jailing could be no more than harassment of the citizen or, perhaps, preventive and deterrent incarceration at best, if there was not some past offence or wrong (real or imagined) which enabled punishment on a purely logical level to be a possible response. What is more, it would be a mistake to think that acts of punishment are alone in possessing such so-called backward-looking features. Acts of gratitude, loyalty, promise keeping and forgiveness also display the same or

similar backward-looking features and they are far from being diminished in their rational, moral and social standing on that account.

However, the phrase “because they have made one suffer” in Elster’s definition seems to indicate not only a necessary, but also a sufficient condition for revenge. This suggests a different interpretation, that revenge is claimed to be backward-looking by reference to its own internal logic. Before discussing this particular interpretation I want to consider Hamlin’s response to the idea of taking a past offence as a sufficient condition for revenge. Hamlin’s first reaction is to argue that “there is no need to accept this reading.”

The fact that someone has inflicted suffering (of a particular type) on you may be a necessary condition for your taking vengeance on that person, but some further and forward-looking condition may need to be satisfied before vengeance will actually be taken. This possibility is sufficient to open the door to possible rational choice explanations of revenge.

There is a clear analogy here with judicial or state punishment. For punishment to be punishment rather than simple injury, there must be an element of the backward looking, but this does not imply any logical difficulty in constructing an account of punishment that mixes backward- and forward-looking elements. Any plausible deterrence or rehabilitation theory of judicial punishment will provide just such a mixed account. We may say that we punish a thief because he stole, indicating the necessary precondition for punishment, while simultaneously saying that we punish that thief in order to deter theft, indicating the forward-looking and contingently rational justification for the punishment.²⁸¹

While I agree with Hamlin’s argument, I want to point out something about the way he made his point. He seems to be under the impression that his argument relies on rejecting Elster’s reading of the phrase ‘because they have made one suffer’ as a sufficient condition for revenge in the definition. The fact however is that, even if we do accept Elster’s reading, Hamlin’s argument still goes through. For, although his argument may not show the possibility of rational revenge when the only reason for revenge is purely backward-looking (‘because they made one suffer’), Hamlin’s argument still shows the possibility of rational revenge in cases where there is in addition

²⁸¹Hamlin 1991, p. 375.

some forward-looking reason which forms part of the motivation for taking revenge. In fact, because of this, regardless of the reading one gives to the definition, much of the revenge behaviour discussed by Elster would come easily within the range of Hamlin's argument. For, since Elster mostly discusses the blood feud, in this special type of revenge the need to save face, honour and the family's good name are very powerful forward-looking considerations which are never absent.

It might be tempting to counter the above argument by interpreting Elster's claim that revenge is irrational to be restricted to purely backward-looking acts of revenge. But if we assume, as Elster does, that rationality is purely future oriented, the resultant claim becomes, not only self-evident, but also trifling. For since the rationality we are dealing with here is necessarily, and only, forward looking, the truth of the claim that *purely backward-looking revenge cannot be rational* is so obvious and trivial that it would be pointless even to mention it. So interpreted, both the thesis and the arguments for it are otiose.

An alternative way of getting around the difficulties raised by Hamlin's argument could be tried by interpreting Elster's claim to be that revenge, by definition, *cannot* be forward-looking, but only backward-looking. This interpretation, however, is indefensible for at least three reasons. First, the irrationality thesis would become, once again, trivial. It would be straightaway established in virtue of two definitions: Rationality can only be forward-looking and revenge can only be backward-looking, so the two simply cannot engage. If this were correct, no further arguments would be needed. The second reason why this interpretation is indefensible is that it would exclude uncontroversial cases of revenge from being identified as revenge. It would be unacceptable, for example, to exclude definitionally the blood feud from being revenge just because it has a forward-looking motivational element aimed at saving or maintaining good reputation and honour. Finally, this reading is clearly not what Elster intended for his definition and arguments. His claim about forward-looking, outcome-oriented considerations is that "prima facie, this motivation is incompatible with revenge

behaviour,” and not at all that the two are obviously or definitely incompatible with each other.

The above considerations expose the fine line which Elster is treading between talking substance and embellishing triviality. For if, by definition, rationality is only forward-looking, and similarly, by definition, revenge is only backward-looking, it would seem to be logically impossible for revenge to be rational. There seems to be only one way to avoid this conclusion. This is to interpret Elster’s claim that revenge is only, or essentially, backward-looking to mean that revenge is backward-looking by reference to its own internal logic. This fits in with the reason ‘because they have made one suffer’ being not only a necessary but also a sufficient condition for revenge, while leaving open the possibility that revenge might still be rational if it could be shown that, in spite of its own internal backward-looking logic, it can be the best means to some independently stipulated end. But, while this reading of Elster’s definition saves his case against revenge from being trivial, it also renders it vulnerable to Hamlin’s argument that in cases where revenge is taken with some forward-looking goals in mind, such as honour and good reputation, the door is wide open to possible rational choice explanations. I shall argue for the defensibility of such explanations in the next chapter.

But even though Elster’s definition of revenge allows for the possibility that there could be forward-looking reasons for taking revenge, is it true that revenge is only backward-looking by reference to its own internal logic? I dispute that claim. The internal logic of revenge is more complex than Elster seems to think. There is more to the logic of vengeance than what is expressed by the simplistic phrase ‘because they have made one suffer.’ Principally, Elster fails to make explicit the essentially retributive nature of revenge. As a result of this failure, an important forward-looking element in revenge, which is part of its internal logic, remains hidden from view. As can be seen from the account of revenge I presented in the first part of this thesis, there is a significant sense in which the aim of revenge is retribution, variously expressed as

desert, payback, getting even, etc. Its aim is the retributive punishment of those offences and wrongdoings which have been committed against oneself and those close to oneself.

It is true, of course, that revenge *is* retribution. What is also true, however, is that revenge is not the *only* form of retribution. This means that in a straightforward sense of the word, the aim of revenge (retribution) can be achieved through other ways than revenge. Vigilantism and purely institutionalised judicial retribution are possible alternatives. Viewed from this perspective, questions about the forward-lookingness of revenge simply shouldn't arise. Furthermore, along the lines suggested by Hamlin, additional elements of forward-lookingness can be present in many cases of revenge, especially in societies with honour-based systems of morality where it is very important to maintain honour and a good reputation. As I shall demonstrate in the course of the following chapter, quite often, revenge was the best way to secure these all-important ends. Presently, I want to move on to consider the normative status and importance of honour for revenge.

5.4 Honour and revenge

In this section I consider Elster's claim that the phenomenon of honour figures prominently in all accounts of revenge. I shall argue that this is mistaken and so is, therefore, Elster's related claim that honour is the key to understanding revenge. Following these discussions I also challenge Elster's low, ethnocentric evaluation of honour as a worthless, if not nasty, pursuit.

The claim that honour is universally present in all accounts of revenge is mistaken. The reason why Elster came to form this impression might well be that, although he proposed to discuss revenge behaviour in general, in an attempt to illustrate and lend empirical support to his arguments, he concentrates almost exclusively on the special case of blood feuds. But, as Hamlin also points out, "blood feuds are a special, and extreme, case for two distinct reasons: the punishment itself - death - is extreme, and punishment is itself seen as punishable, leading to the possibility of a spiral of counter-

vengeance.”²⁸² A third, perhaps more pertinent reason, however, is that the phenomenon of blood-feud is present only in societies with strong, honour-based systems of morality. Given Elster’s singular attention to blood feuds, it is easy to see how he might have formed the impression that honour figures in all accounts of revenge.

But, however that may be, the truth of the matter is that in many modern societies the importance of the concept of honour is virtually obsolete, its place having been taken by a concept of human dignity. The fading of the importance of honour and its replacement by human dignity seems often to coincide, roughly, with urbanisation and industrialisation and in certain parts of the world has been a historically documented process.²⁸³ Honour and dignity will be discussed further below where I shall also defend honour against Elster’s disparaging remarks. What I want to draw attention to here is that, in contrast to the obsolescence of honour, revenge is far from being a thing of the past. In spite of heavy penalties, people still decide to see to it personally that their wrongdoers get their retribution, especially in instances where they feel the law has let them down in this regard. Good examples are the case of Stephen Owen, a father who shot the unrepentant and spiteful killer of his son,²⁸⁴ and the more recent case of Malcolm X’s daughter who tried to hire a hit man in order to see her father’s alleged killer dead.²⁸⁵ In many such cases honour is not even a consideration for individual people seeking revenge.

²⁸²Hamlin 1991, p. 375.

²⁸³In his book *Vengeance and Justice*, Ayers (1984) gives a fascinating account of this process of change which took place at a much faster rate in the 19th century American North than in the American South “where nothing seems to go down without a fight.”

²⁸⁴Daley, J. Blinded by emotion. *The Times*, Tuesday May 26, 1992, p. 12.

This case is also remarkable because Owen walked away a free man after “the jury declared Mr Owen innocent of all charges, even those that were simple statements of fact to which he had admitted (such as wounding Mr Taylor).” (p. 12.)

²⁸⁵Wyndham, S. Malcolm X’s daughter in revenge case. *The Weekend Australian*, January 14 - 15, 1995, p. 1. See also, Terry, D. Malcolm X: The Payback. *The Sydney Morning Herald*, Saturday, January 14, 1995, p. 1.

This also has vital implications for Elster's second claim that the phenomenon of honour is the key to understanding revenge. If this were right, then such cases of revenge in the modern world would have to go unexplained. This is not the case, however, for these cases of revenge can be readily understood by reference to a retributive principle of value, reinforced by feelings of resentment, which, as Butler points out, get stirred up by perceptions of unfairness, injustice and moral wickedness in general, but especially in response to those directed at us and those dear to us.²⁸⁶ Perhaps, not even Elster would deny these arguments. On one level he certainly seems to be aware that

... many societies have no codes of honor or norms of revenge. ... spontaneous revenge behavior is universal: norms of revenge are not. In many societies there would, if anything, seem to be a norm against seeking revenge and, correlatively, injunctions to turn the other cheek.²⁸⁷

Yet, notwithstanding his awareness that in many instances revenge has nothing to do with honour, it is almost in the same breath that Elster expresses the conviction that "the phenomenon of honor ...[is] the key to understanding revenge."²⁸⁸ This is clearly impossible and, although the contradiction generated by these two claims could be read away by taking the latter claim to be about blood feuds or norms of revenge, enough has been said, I believe, in support of the view that honour is neither universally present in all accounts of revenge nor is it the key to understanding revenge. This is not to deny, of course, that in cases where considerations of honour form part of the motive for revenge, it has to be given its rightful place in the explanation of those cases of revenge. Ironically, in spite of his emphasis on honour as the key to understanding revenge, Elster fails to grant honour its rightful place where rational choice explanations of revenge are concerned. I shall take issue with this aspect of his argument in more detail in the next chapter where, among other things, I argue that rational choice explanations of honour

²⁸⁶The relevant passage from Butler was quoted in the first chapter in Section 1.4.

²⁸⁷Elster 1990, p. 883.

²⁸⁸Elster 1990, p. 883.

motivated revenge are not only informative and satisfactory but are best suited to explain the relevant behaviour.

Now I want to move on to consider the normative defensibility of pursuing honour as a proper value and goal. Once again, Elster's low esteem of honour provides a useful starting point.

Asserting one's honor, like enjoying other people's envy of one's assets, is an aspect of a deep-rooted urge to be superior to other people. ... Its aim is sheer self-assertion and self-esteem. ... I believe the urge for honor, like the enjoyment of other people's envy, are universal phenomena. They can be controlled but not fully suppressed. They arise in the mind spontaneously but need not have any further effect if we can recognize them and avoid acting on them.²⁸⁹

Although in putting up his case against the rationality of revenge, Elster chooses not to rely on the disparaging, normative evaluation he gives to honour, if such views were sustainable, they could be used as a basis for normatively excluding honour from being a proper goal of revenge. Even if his actions met the formal requirement of purely ends-means rationality, honour just couldn't be the sort of value or goal for which a rational agent would or should want to take revenge from a normative point of view. Elster's evaluation, however, is indefensible. It displays a serious lack of understanding of the nature and vital significance of honour for the individual in the relevant cultures. In cultures with an honour-based system of morality, honour is constitutive of a person's very identity in the epistemic sense (but not necessarily in the metaphysical sense, of course) in the same way that dignity is constitutive of our identity in contemporary Western societies. Identity in the epistemic sense is a narrative, a story one constructs and tells about oneself regarding who and what sort of person one is. In a culture of honour the individual defines him/herself in terms of his/her position, status and the roles s/he has in his/her community. Failing to fulfil one's roles means losing one's honour and together with that one is losing one's self-respect, and one's identity as a worthy individual. It will be worth quoting Peter Berger in this regard:

²⁸⁹Elster 1990, pp. 883 - 84.

Dignity, as against honour, always relates to the intrinsic humanity divested of all socially imposed roles or norms. It pertains to the self as such, to the individual regardless of his position in society. ... [By contrast] in a world of honour, the individual discovers his true identity in his roles, and to turn away from the roles is to turn away from himself . . .

Both honour and dignity are concepts that bridge self and society. While either pertains to the individual in a very intimate way, it is in relations with others that both honour and dignity are attained, exchanged, preserved or threatened. Both require a deliberate effort of the will for their maintenance - one must strive for them, often against the malevolent opposition of others - thus honour and dignity become goals of moral enterprise. Their loss, always a possibility, has far-reaching consequences for the self.²⁹⁰

Since in honour-based cultures honour is an essential ingredient of an individual's identity, self-worth and self-respect, it is something that in valuing oneself the individual necessarily nurtures and, if need be, fiercely protects. But, a culture of honour is also likely to make different demands on the individual from a culture of dignity.

In a culture of dignity, people were expected to remain deaf to the same insults that Southern men were expected to resent. "Call a man a liar in Mississippi," an old saying went, "and he will knock you down; in Kentucky, he will shoot you; in Indiana, he will say 'You are another.'" Dignity might be likened to an internal skeleton, to a hard structure at the center of the self; honor, on the other hand, resembles a cumbersome and vulnerable suit of armor that, once pierced, leaves the self no protection and no alternative except to strike back in desperation.²⁹¹

It is clear from this passage that in a culture of honour an injury to one's honour can be as damaging to the individual as injuries to our dignity today would be and, probably, much worse on account of the intense social disapproval which accompanies one's loss of honour. Such threats to one's self-respect and identity as a valuable, worthy individual cannot be ignored. Something must be done about them. In the relevant cultures in particular - where in many cases there is no efficient external authority to take possession of the conflict - it often happens that the best, even if not the only, available defence against such threats is revenge. In such cases taking revenge is perfectly rational in a proper and defensible normative sense. By taking revenge one

²⁹⁰Berger 1984, pp. 153 & 154.

²⁹¹Ayers 1984, p. 20.

employs the best available means to save one's honour which is an essential and integral part of one's identity as a worthy individual.

If all this is correct, honour cannot be excluded from being a proper aim of revenge and, therefore, it cannot be excluded from being a goal capable of providing a basis for substantive, rational choice explanations of honour-based revenge. I consider in proper detail the direct rationality of honour motivated revenge behaviour in the last section of this chapter. I move on now to give a broader context in which these and other arguments against revenge by Elster can be profitably discussed.

5.5 Revenge and the material/non-material distinction

Regarding the rationality of revenge, Elster distinguishes between material and non-material interests of the agent.²⁹² Thus, although Elster does not explicitly put it this way, his key assertions and arguments against revenge are suggestive of a dilemma: We either try to show that revenge is rational by reference to the material interests of the individual agent, or by reference to the non-material ones. If the former, rational choice explanations of revenge fail because it can never be in one's (material) interests to take revenge. If the latter, rational choice explanations fail because they are uninformative, unsatisfying and uninteresting. Either way we take it, revenge behaviour cannot be shown to be rational from the individual agent's point of view.

I shall argue in the remainder of this chapter that neither horn of this suggested dilemma presents a problem for revenge. In the first main section I consider the rationality of revenge with respect to material goals. I shall argue that revenge behaviour can be shown to be rational in this regard, both in the direct and the indirect senses of the term. In the other remaining main section I consider the rationality of revenge with respect to non-material goals. I shall argue that with respect to the non-material interests

²⁹²Elster 1990, p. 872.

of the individual agent, rational choice explanations of revenge are appropriately informative and satisfactory. Indeed, very often they are the best explanations on offer.

5.5.1 Revenge and material interests

In this section I consider four things: First, I show that it is mistaken to suppose that in material terms revenge can involve only costs and risks, no benefits.²⁹³ Second, I show that the desirability and rationality of being revengeful by disposition provides a defensible basis for the indirect rationality of individual acts of revenge forming part of that disposition. Third, I show that in communities with an honour-based system of morality, social expectations and pressures to carry out the duty of revenge and the real threat of severe sanctions make compliance individually rational. Fourth, I argue that even if an action is not strictly speaking rational from a purely economic perspective on account of its costs and risks, provided that these are not ignored but are taken into account in a responsible manner, and can be afforded in the interests of some other, greater good, then the action can still be rational from an all-things-considered perspective.

5.5.1.1 Immediate costs and benefits in revenge

According to Elster, by definition, revenge “only involves costs and risks, no benefits.”²⁹⁴ This is a mistaken supposition and I want to show in this section that there can be circumstances in which revengefulness is accompanied by substantial material rewards, even though the aim of revenge is not these rewards, but retribution. To this end, consider the following example.

Transylvanian Life. Suppose that Jakob and Julia live as good neighbours in a Transylvanian township sometime during the *Ceausescu* era. Just like most other people around them, they work hard and are paid little. And, just like most other people around

²⁹³Elster 1990, p. 862.

²⁹⁴Elster 1990, p. 862.

them, they both engage in unlawful activities of various sorts in order to improve the quality of life for their respective families. Jakob, for example, regularly steals agricultural goods, such as corn, wheat and potatoes, from the nearby cooperative and whatever he does not use he sells on the black market for a profit. Julia is similarly involved in improving her life by illegally importing and selling western luxury items such as jeans, colour television sets and push button phones. A third neighbour is involved in diverting building materials from state construction sites to trusted customers of his own, and so on. They all know of each other's sources of flourishing and also that there is a substantial reward for reporting on other people's illegal activities, say \$5 000. Despite the incentives, they all abide by the implicit agreement that they will keep each other's activities in confidence because, after all, they are all in the same boat, so to speak.

Now suppose that on one fine day the *securitate* raids Julia's house, seizes her luxury items and she ends up in a forced labour camp. She is also questioned regarding other illegal activities that she might know of and, as an inducement she is promised not only the usual reward for reporting on illegal activities but also a substantial reduction in her sentence. Obviously, in the circumstances it would be perfectly rational for her to report any of her neighbours. However, having a sense of solidarity, she does not report any of the others despite the fact that in the circumstances it would pay her to do so even more than it would have before she were caught.

For present purposes we can put aside the question whether her silence over her neighbours' illegal activities is rational or not. The only question that matters now is whether it would be irrational of her to take revenge on any one of her neighbours if she found out that a particular one of them was the informer. Suppose, for example, that after a period of bitterness and hard labour in a detention camp she finds out that it was Jakob who reported her. She is outraged by the betrayal and vows to get even with Jakob, to pay him back for everything she has endured. But will it be rational of her to seek retribution and avenge Jakob's treachery by, say, similarly reporting on his

booming illegal business? Of course it will be. She will be rewarded for her reporting with \$5 000 and her sentence will also be reduced.

This example shows clearly, I believe, that the material costs and benefits of revenge are circumstantial. It is a mistake to suppose that on account of its retributive character, revenge can involve only costs and no benefits. A residual worry that some people might express at this point is whether Julia's reporting on Jakob is appropriately described as revenge. Is it not otiose, the objection might go, to call it revenge when it would have been rational of her to report on Jakob even if she was totally ignorant of his treachery? The answer is straightforward enough: Julia's primary motive was fair retribution; she wanted to get even with Jakob, she wanted to pay him back for what he did to her. Her getting back at him is both retributive and personal in the most appropriate senses for her action to qualify as a genuine, unambiguous case of revenge.

A further feature of this case is the fact that she did not consider the rationality of reporting on someone whom she believed was innocent (of harming her) as a sufficient reason for getting others into trouble. Her discovery of Jakob's treachery and her desire to hit back was what made the relevant difference. It was her desire to hit back for the treachery which made her reporting a genuine case of revenge, and the fact that her action was also going to benefit her considerably in the particular situation was what made it appropriately rational. This is a clear, unambiguous example of rational revenge where one is rational to engage in revengeful behaviour even in a single, one-shot game.

This example does not show, of course, that revenge can be rational on account of its own nature. It was never intended to show that and no one needs to show that. My claim is not that revenge is *intrinsically* rational, that it can be rational independently of any further considerations besides retribution, regardless of circumstances. In any case, as I already pointed out, there would be few, if any, human actions which could meet such stringent criteria of rationality. Rather, I am arguing that in certain circumstances revenge can be an individually rational option to take. There are circumstances in which, far from being counterproductive, detrimental or self-defeating, revengefulness can be

quite compatible with important material interests even in one-shot game situations. In the next section I want to show that, even if revenge behaviour is costly in the short term, it can have substantial benefits if a long term perspective is adopted. Absorbing short-term losses in the interests of greater long-term returns is eminently rational.

5.5.1.2 Long term interests and indirect rationality in revenge

I want to show in this section that a revengeful attitude or disposition can have important material benefits for the individual concerned and that, therefore, individual cases of revenge which form part of such a desirable disposition are rational in the indirect sense I explained towards the beginning of the previous chapter. Axelrod's famous work *The Evolution of Cooperation* clearly demonstrates that in iterated prisoner's dilemma situations where the future is sufficiently important relative to the immediate payoffs, the strategy of repaying like with like is the best way to elicit the cooperation of others and thereby, on the whole, the most efficient way to do well for oneself.²⁹⁵ TIT FOR TAT - the robust, winning strategy of several prisoner's dilemma computer tournaments - cooperates on the first round and thereafter always does what the other player did on the previous occasion. This simple decision rule has several virtues which jointly explain its robust performance, the relevant one here being that it is not exploitable. If, for example, an opportunist rule tries to exploit it by "unprovoked" defection(s), a TIT FOR TAT player swiftly retaliates with a like defection in the next round and will continue with the "punishment" by defecting until the other starts to cooperate again. This unfailing retaliatory feature - which is structurally identical to a strategy of retribution and revenge - ensures that in iterated prisoner's dilemma situations TIT FOR TAT cannot be taken advantage of to any significant extent.

That a keen disposition to retaliate for defections is a desirable attribute to have can be seen by considering a more forgiving strategy than TIT FOR TAT which does not punish isolated defections. A strategy of ONE TIT FOR TWO TATS, for instance,

²⁹⁵ Axelrod 1984, pp. 27 - 54.

could be exploited quite easily by a cunning, opportunistic strategy. One could, for instance, easily defect every alternate round and a ONE TIT FOR TWO TATS player would keep on foolishly cooperating to her own disadvantage. And if this is what can happen to someone who was somewhat slow to resent exploitation actively, how much the worse it would be for someone who *never* retaliated for exploitative behaviour.

In the light of these considerations, I take the potential and actual advantages of not letting others get away with just anything to be obvious, and there can be no doubt that revengefulness is capable of serving such a function. What is more, it is probably the best means of serving that function. An important concept which can be profitably appealed to at this point is resilience. The basic idea here is that an attitude of genuine revengefulness based on retributive considerations and resentment is more resilient than a consequentialist approach to calculating negative “retaliation.” Thinking consequentially on each individual occasion to decide whether to retaliate or not exposes one constantly to the temptation of giving in to immediate, short term considerations and, thus, to the increased danger of becoming a ONE TIT FOR TWO OR MORE TAT-er - in less flattering terms, the sucker. By contrast, retaliation based on a genuine retributive attitude and resentfulness is not in such danger. If anything, retributive dispositions and feelings would need to be restrained in cases where the costs of retaliation would be inordinately large or too hard to bear. Consequently, with the exception of cases where revenge is against one’s short and long term interests overall, quite generally revenge behaviour is indirectly rational in that it is part of a pattern of behaviour springing from a retributive disposition and attitude towards those who harm one’s personal interests - a retributive disposition and attitude which itself is desirable and rational to have.

Indeed, even Elster is prepared to admit that along these lines there is a mechanism through which one could engage in rational revenge. He calls it “the rationality of appearing to be irrational.”²⁹⁶ To see what exactly is being claimed here, consider the

²⁹⁶Elster 1990, p. 873.

following game-theoretical model where there are two players, I and II. (See Figure below.) I has the chance first to choose between going left and going right. If she goes left, the payoffs for the players are 2;2. If she chooses to go right then player II will have the option of choosing between going left and going right. If II goes left, she gets 1 for herself and 3 for I. If she goes right, neither of them gets anything.

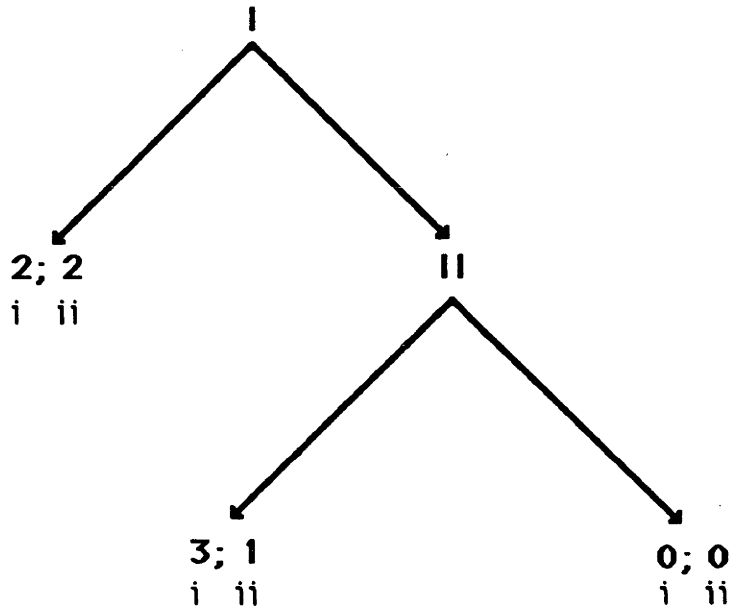


Figure 6

Regarding this model Elster claims that if both players are fully rational and not moved by backward-looking considerations then “player I knows that if he goes right, II’s self-interest will induce her to go left, thus ensuring the best outcome for I.”²⁹⁷ He does acknowledge, however, that if II has a reputation for being revengeful - because in the past she has often gone right in similar situations - then I knows that if he goes right, there is a high probability that II will go right as well. I being rational, and preferring 2 to 0, will do best if he goes left. The consequence, of course is that II will get 2 as well. This means that in the long run it pays for II to be revengeful and therefore revengefulness can be a perfectly rational attitude and disposition for her to adopt. In fact, this model represents the mechanism of indirectly rational revenge I have been

²⁹⁷Elster 1990, p. 874.

discussing above. But, curiously enough, Elster wants to dismiss the model as useless in arguing for the rationality of revenge:

By this mechanism, it could indeed be rational to engage in acts of revenge. Note, however, that the mechanism is parasitic on the existence of some genuinely irrational persons in the population. What drives the argument is the common knowledge that society has some rational and some irrational members, but that they do not bear their rationality or lack of [it] on their face. In a population of individuals known by each other to be fully rational, nobody would ever exact revenge.²⁹⁸

The central claim of this passage is mistaken. The argument need not presuppose the existence of irrationally revengeful persons at all. If anything, its presupposition does not have to be more than the ability of people to recognise revengeful behaviour for what it is: a TIT FOR TAT pattern of behaviour which is based on a genuine attitude of retribution and resentment and which is unfailing in visiting punishment on all (and only) those who choose defection over cooperation. Returning to the model above, it should make no difference to I to know that II is a TIT FOR TATTER, not out of blind resentfulness, but because to some degree she consciously has chosen to maintain and cultivate in herself a retributive attitude and disposition towards all defections.

Many people have a fairly strong disposition to feel resentment and moral indignation and have a retributive attitude to injustice and wrongful conduct, especially when these are directed towards them. Such dispositions and attitudes, however, are not immutable as they are subject to rational and moral evaluation and it is within the ability of most of us to make a decision as to whether we continue to maintain and nurture retributive dispositions, or whether we try to change them and replace them with what are regarded by some to be “higher,” “spiritual,” or more “humane” responses, such as unconditional pacifism or forgiveness. On the value and moral appropriateness of these latter attitudes I need not pronounce judgement straightaway, although I shall discuss these issues in the final chapter. What is important to keep in mind here is that, given an

²⁹⁸Elster 1990, p. 875.

intensely competitive world where, generally, people are looking out for their own interests first and foremost - and all too often unscrupulously, to the detriment of others - choosing to maintain retributive dispositions, attitudes and emotions is fully rational. Upon reflection, such a choice can spring from the recognition that, besides it being fair, retribution is an essential part of that winning strategy of *reciprocity* which is the best way overall to elicit fair cooperation from others.

It will help, perhaps, to illustrate these points with another example. Suppose that Paul and his cousin, Mary, have an eccentric, and somewhat bored aunt who likes to amuse herself by devising little games for her nephew and niece whom she knows do not particularly like each other, or her, but try to pretend otherwise with their eyes on her money. She presents them with the following game: Every time Paul rings her first on her birthday she rewards them with \$20 000 each but gives \$25 000 to Paul and only \$5 000 to Mary if it is Mary who first rings her. If neither of them rings, they receive nothing. This game situation can be represented diagrammatically as follows:

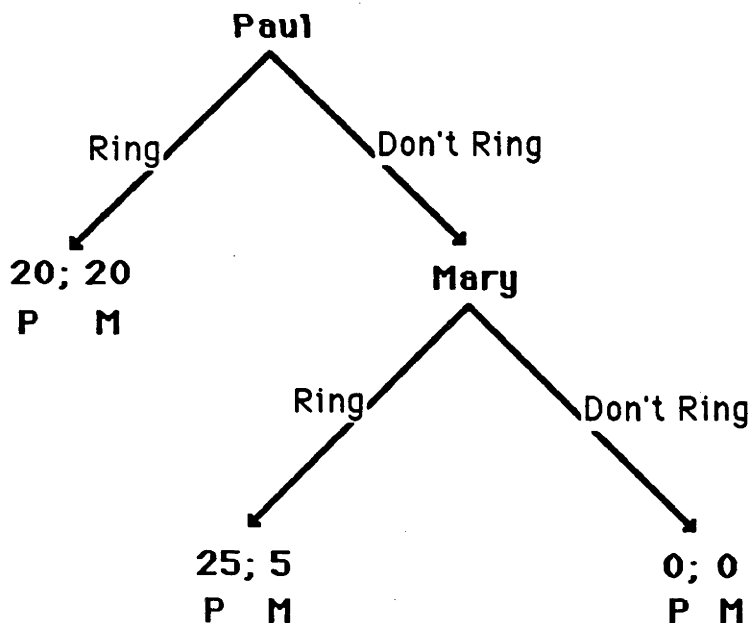


Figure 7

If it is only money that matters to them both, Paul will do best if Mary rings first because Paul will receive \$5 000 more than what he would receive if he rang. But by the same token Mary will do best if Paul rings first because Mary will receive

\$15 000 more than what she would receive if she rang. Knowing each other to be fully rational, and given this conflict of interests, what are the rational strategies for each of them to pursue? Suppose that Paul decides on the following: He makes it clear to Mary that he is not going to ring their aunt on her upcoming birthday. To leave no doubt in Mary's mind that he means what he says, Paul starts out on a day-long bush walk. Aware of Mary's knowledge that he will have no access to a telephone on the day, and knowing that in money matters Mary is fully rational, and therefore, that Mary prefers \$5 000 to nothing, can Paul feel confident of a \$25 000 reward by the end of the day as a consequence of Mary's call to their aunt?

I want to show that, if Paul is fully rational, he cannot have that expectation. To start with, Mary cannot help but resent Paul's decision. She considers it to be very unfair, especially given that she has gone out of her way to help Paul on previous occasions. Feeling let down and angry at the treatment, she wants to get back at Paul, to get even with him by giving him his due deserts. It occurs straightaway to her that she can get her revenge for having been disadvantaged by refusing to ring their aunt. This will come at a price to her, of course, because she will miss out on a \$5 000 reward every year, but she can deprive Paul of a yearly \$20 000 - \$25 000 and that will serve him just right. But then she wonders whether this is going to be in her best interests in the long term. Isn't she cutting off her own nose to spite her own face by allowing herself to be retributively punitive in this matter? Being fully rational in the sense stipulated, shouldn't she drop her determination for revenge, let bygones be bygones and be contented with her yearly \$5 000 and just not worry about Paul's attitude, or about the yearly \$15 000 she is missing out on?

Upon reflection, the answer to these questions is negative. Taking revenge cannot pay off for Mary if, say, their aunt is close to the grave and the game is not likely to be repeated in the future. In such a situation Mary may well decide to ring and receive \$5 000 rather than nothing. But if both of them expect that the same game-situation is going to be repeated on numerous future occasions, Mary will be better off in the long

term if she does not put up with Paul's greed. The short-term result, of course, might well be that neither of them receives anything. The long-term result, however, is bound to be different. For if Paul sees this sort of response to his strategy a couple of times, he would have to be a fool not to change - and change he will, if he is rational. By ringing he can ensure that he receives \$20 000 every time and it is no skin off his nose to know that Mary benefits the same. It would be imprudent of Paul to provoke Mary afresh and, in the hope of receiving an extra \$5 000, repeatedly risk losing \$20 000.

But the most interesting and crucial side to this particular game situation is that only Mary would be rational to engage in a struggle to outlast Paul. This is because both of them are well aware that the costs of an impasse, and the benefits gained in the event of winning it, would accumulate rapidly on the two sides at a very uneven rate. If they both adopt the strategy of tiring out the other, Paul has much more to lose than Mary, and in the event of success, also significantly less to gain than she does. Let us count up the losses first. A loss can be properly defined in the present context as that amount which a person in fact misses out on but could have ensured for him/herself independently of what the other person decided to do. In the case of an impasse, Paul's refusal to ring would cost him \$20 000 in each round whereas Mary's refusal to ring will cost her only \$5 000. This asymmetry has to make a difference to their respective rational determination to hold out in a struggle to outlast the other. Mary has far stronger reasons to outlast Paul than the other way around.

A similar picture presents itself if we compare the gains they respectively can look forward to on the event of outlasting the other. If Paul outlasted Mary, for example, his extra gain would be only \$5 000 per round, whereas if Mary outlasted Paul, her extra gain would be \$15 000 per round. Again, if they are rational, this asymmetry also makes a difference to their respective rational determination to hold out in a struggle to outlast each other. Once again, Mary has much stronger incentives, and thereby stronger reasons, to outlast Paul than Paul has to outlast Mary.

Their respective reasons are not different in kind, of course. They are different only in strength. Nevertheless, they are crucial at every point in deciding whether to continue holding out in the course of an impasse. Paul's knowledge of Mary's degree of rational determination relative to his own makes it irrational for him to expect that Mary is going to cooperate. By the same token, Mary's knowledge of Paul's degree of rational determination relative to her own makes it rational for her to expect that before long Paul will start to cooperate. This remains true at least as long as Paul remains free to give in and ring their aunt.²⁹⁹ Recognising these facts about the situation, Mary can reassure herself that her resentfulness and intended retribution are in the right place from a rational choice perspective. Rather than working against her best interests, her act(s) of revenge, which has a retributive motivational structure, will lead to the very same pattern of behaviour towards Paul which she should rationally adopt anyway as a direct result of purely non-retributive, totally forward-looking rational choice considerations. Knowing this can only strengthen her resolve to be, and remain, punitive towards Paul - at least as long as he refuses to adopt a non-exploitative attitude towards her.

Elster is mistaken, then, in claiming that in a population of individuals known by each other to be fully rational, nobody would ever exact revenge. It simply isn't the case that one has to be believed by others to be genuinely irrational in order to adopt successfully an attitude of revengefulness towards them. Revengefulness does not become any less efficacious in eliciting fair cooperation from others even if it is known by these relevant others that to some degree the attitude has been adopted and keeps being cultivated for a purpose. Even if one was known to be fully rational and was completely "transparent" to others regarding one's motivations and goals, there still would be situations in which one could maintain revengefulness to personal advantage in the most concrete and material of terms. In the example above, Mary's revengefulness towards

²⁹⁹ If Paul managed to make it irreversibly impossible for himself to ring, the picture might well be different, of course.

Paul would work regardless of whether Paul believed Mary to be mad or whether Paul believed Mary always to remain calculating even when strongly resentful.

5.5.1.3 Social sanctions and direct rationality in revenge

In this section I argue for the direct rationality of individual revenge behaviour. I want to show how a mechanism of sanctioning makes revenge individually rational in the direct sense for an otherwise reluctant individual. Especially in communities with no centralised system of punishment, social expectations and pressures to carry out the duty of revenge and the real threat of severe sanctions can easily make compliance with the norms of revenge individually rational. As I suggested towards the end of Chapter 3, in communities without a centralised system of punishment, socially shared conceptions of honour, codes of conduct, norms of revenge, as well as the real threat of sanctions for not abiding by the established norms, all form part of a highly evolved mechanism of social control by means of which society puts pressure on its individual members to carry their burden in the punishment of wrongdoers. They make a duty out of revenge and make it individually rational for individual people to take revenge in accordance with the norms even in circumstances in which they would be tempted to do otherwise because of the potential costs in terms of time, stress and other resources. The mechanism in question seeks to outweigh the costs associated with revenge by increasing the costs of the alternative by means of sanctions and social ostracism which are imposed on persons who fail to carry out the duty of revenge, and which, as we have seen, can be nothing short of crippling.

It is worth noting here that such mechanisms of social control assume and rely on people's individual rationality in general, and on their individual rationality in matters of revenge in particular. Increasing the costs of non-compliance with the norms by means of sanctions would make no sense, and it certainly would not work, if people weren't moved by rational choice considerations in matters of norm following in general and revenge in particular. Thus, returning to the argument under consideration, to the extent that a person considers the tangible costs of *not* taking revenge to be greater than the

tangible costs of *taking* revenge, to that extent it is individually rational for that person to go ahead and fulfil the expectations of the would-be sanctioners by taking revenge. As far as I can see, this argument for the individual rationality of revenge behaviour in the contexts discussed, is flawless. But Elster disagrees and claims that the case of revenge is one in which this account is particularly implausible for two reasons.

If it were correct, revenge behavior would be singularly overdetermined, because the internal, emotional forces also seem to provide a sufficient explanation. The argument also fails more generally, for reasons that apply to all norm-guided behavior. We have seen that violators of a norm shared with other members of their society are indeed exposed to sanctions by these others. But then we have to ask what reason others could have for sanctioning the violators. What's in it for them?³⁰⁰

Neither of these objections cuts any ice as long as we adhere to the limited objective of answering the question: why would a rational person ever carry out an act of revenge? The answer proposed here is as follows: In contexts in which people go out of their way to sanction severely one who fails to carry out the duty of revenge when required to do so by the accepted norms, one may be better off taking revenge and avoiding these sanctions than not taking revenge and enduring these sanctions. The objection from overdetermination is simply irrelevant here. It is hard to see how overdetermination could pose a problem and, therefore, there is no reason to pose the question: *What is the one, single reason for which (rational) individuals take revenge?* There simply need not be only one reason. If a decision to take revenge based on the desire to avoid sanctions is a rational one, the rationality of this decision can in no way be affected by the fact that the individual had, in addition, other good reasons for taking revenge, such as the value he attaches to responding retributively to wrongs and resentment. In any case, there is no mystery here about how overdetermination of revenge in communities with no centralised system of punishment comes about. I explained this in the latter part of Chapter 3.

³⁰⁰Elster 1990, p. 873.

The other objection, that the sanctioners do not gain anything by sanctioning the violators of a norm, is equally irrelevant.³⁰¹ For even if their sanctioning is an irrational activity, this in no way can detract from the fact that one will be worse off as a violator of the norm than one is as a norm follower. And, if we are to keep our eyes on the ball, the point here is simply that this and this alone can come into consideration for determining the *individual rationality* of a person's action in following the norm. Contra Elster, for our limited objective of considering the individual rationality of revenge behaviour, we need not ask what reasons others could have for sanctioning violators of the norms of revenge. Whether there is anything in it for them or not is only relevant in determining the rationality of *their* decision to sanction the norm violators. It is not relevant to determining the *individual rationality* of someone's action in following or violating them.

To drive home the point, suppose that a mob of hooligans stops me in a deserted street and asks me to slap myself on the face or else they will severely beat me up. To determine whether I am doing the individually rational thing by giving myself a big slap, we need not ask what's in it for them in beating me up if I don't. Whether they are rational or not in their "game", and whether the overall situation is ultimately explicable in rational terms or not, the facts remain that I am better off hurting myself with my own hands than being hurt by their fists and boots. This, and nothing but this, makes it individually rational for me to choose the former over the latter choice. And the norm follower's choice to take revenge out of fear of being severely sanctioned by others is entirely analogous.

In conclusion, the argument for the *individual rationality* of acts of revenge in highly specific cultural contexts, such as those outlined earlier, does work. But, of course, this leaves unresolved the interesting and important question of whether or not in these cultural contexts people have a defensible reason for sanctioning violators of the

³⁰¹But the underlying claim is also contestable because, as John Braithwaite remarked, those who sanction violators may stand to gain/maintain "honour and a sense of asserted superiority to the man without honour, for example."

accepted norms. Consequently, as it stands, the argument does not give a complete explanation of norm-guided revenge behaviour. In this area Elster's queries also become relevant and his observation that "some sanctions must be performed for other motives than the fear of being sanctioned" becomes fully applicable.³⁰²

Elster's mistake in this area lies in supposing that there aren't any such defensible motives. He explores only one possibility, that "for any ordinary norm there is a meta-norm that enjoins people to punish people who fail to punish violators of the first-order norm," and observes with some justification that "this argument soon runs out of steam."³⁰³ However, a satisfactory answer to the problem is provided by the value moral communities accord to a retributive response to injustice and wrongdoing in general. The same value also explains the norms of revenge and the relevant parts of the codes of honour, which form part of the mechanism of social control through which the more narrow individual rationality of revenge is ensured even in cases where revenge would otherwise be irrational from the individual agent's limited perspective. In other words, sanctioners are trying to uphold a very important value, the value of retribution which, I have argued, is based on individual moral responsibility and accountability.

5.5.1.4 Material versus overall rationality in revenge

In this section, I want to consider those cases of revenge which are detrimental to the immediate, and sometimes long term, economic interests of the individual concerned. It is important to consider the degree of significance that can be attached to such economically irrational actions. The key question to answer here seems to be whether the economic irrationality of such acts of vengeance tells decisively against them. I argue that the correct answer, once again, is negative. We are complex creatures and live multi-faceted lives and, even if it is a very important one, economics is only one facet among

³⁰²Elster 1990, p. 873.

³⁰³Elster 1990, p. 873.

many. This is well recognised even by (philosophically astute) economists.³⁰⁴ In making decisions of any kind we have a whole host of other considerations besides economic ones with which we must contend. Even on the individual level there are psychological and spiritual considerations that we all must consider and it is unrealistic to expect either that these latter ones will invariably coincide with one's economic interests or that it will always be in the best overall interests of the person to give economic considerations a precedence over the other two when they conflict. Hence, the fact that some action, including an act of revenge, is economically irrational from the individual's point of view cannot by itself constitute a decisive objection to it even on the individual level because the individual can have, and often does have, a number of other interests against which the economic ones will have to be traded when there is a conflict between them.

This is even more the case when we move from the individual to the social level. We live in communities with other persons and this brings in additional considerations or interests for each of us: political, social, humanitarian and moral. Once again, it is unrealistic to expect either that these will invariably coincide with one's economic interests or that it will always be in the best overall interests of the person always to give economic considerations a precedence over them when they conflict. This point reinforces the earlier conclusion that the fact that an act of vengeance is economically irrational from the individual's point of view cannot constitute a decisive objection to it and, similarly, the economic irrationality of any act in and by itself is only part of the equation and it is a mistake to regard it as decisive.

There is no need to deny in any of this that the overall irrationality of an act, be it an act of revenge or not, is of serious concern. For, even if an action is not directly rational from a purely economic perspective because it involves certain costs and risks, provided these are not ignored but are taken into account in a responsible manner, and can be

³⁰⁴See, for instance, Sen 1977.

afforded in the interests of some greater good, then the action is still rational from an all-things-considered perspective. I discuss this point in more detail in the following section.

5.5.2 Revenge and non-material interests

In this section I consider the second horn of the suggested dilemma: Elster's claim that rational choice explanations of revenge with respect to non-material goals in general are uninformative. I want to show that, particularly in cultural contexts where part of the motivation for revenge is honour, the behaviour can be given informative and interesting rational choice explanations. With possible exceptions, in such cases revenge behaviour is rational in the direct sense. Given the defensibility of honour as a reasonable and proper goal, such explanations are also defensible from a full-blown, normative approach to rationality. As I have already indicated, while Elster holds the view that the phenomenon of honour is the key to understanding revenge, he is of the opinion that "rational-choice arguments can[not] capture the phenomenon of honor, which figures so prominently in all accounts of revenge."³⁰⁵ I already argued against the claim that honour is universally present in revenge. I presently argue that, when appropriate, rational choice explanations can properly appeal to honour as a goal in revenge. For, since the conviction that rational choice explanations of revenge with respect to non-material goals in general are uninformative seems to be the only reason for Elster's claim that rational choice arguments are unable to capture the phenomenon of honour, it will be useful and appropriate to start with a more detailed discussion of this latter claim.

There are at least three ways in which Elster's claim that rational choice arguments cannot capture the phenomenon of honour could be argued. One way would be to dispute the claim that revenge can in fact be a means of acquiring and maintaining honour. However, this is empirically disconfirmed by revenge practices in communities

³⁰⁵Elster 1990, p. 876.

with an honour-based system of morality, such as the blood feud and, accordingly, Elster doesn't consider it. Another way to defend that claim would be to challenge the rationality of honour itself on the ground that it is a worthless, misguided, indefensible goal. Although it is clear that Elster holds honour in very low esteem, he in fact chooses not to rely on this in putting up his case against the rationality of revenge. A likely reason for this might be his conviction that revenge cannot measure up to even the minimal requirements of rationality in the thin sense. But however that may be, I have already argued in detail against such an ethnocentric view of honour, and for the defensibility of honour as a worthwhile value and goal.

We are left with a third possibility. Elster's claim that rational choice arguments cannot capture the phenomenon of honour seems to be based on his belief that explanations of revenge behaviour with respect to non-material interests, such as honour and a good reputation as a norm follower, are uninteresting and uninformative. This is a view which I now want to consider in detail and a useful starting point for discussing it is provided by Hamlin who argues for a particular interpretation of the word 'capture' in Elster's claim.

Elster argues that "I do not believe rational-choice arguments can capture the phenomenon of honor" (p. 876); what does 'capture' mean in this context? If it means that rational choice theory does not, of itself, argue for the rationality of honor as an end, then we may all agree, but surely this is not a relevant interpretation (not least because it is difficult to see how rational choice arguments "capture" selfish, material interests in this sense). The relevant interpretation of 'capture' must surely involve the question of whether a rational choice framework of analysis, together with the specification of honor as a primitive motivation, can offer insights into the pattern of behavior observed. I would suggest that the answer to this question is yes.³⁰⁶

Hamlin is correct in pointing out that the thin conception of rationality used by Elster is both unable and inappropriate to evaluate the rationality of interests and goals of any sort. Interests, desires and goals need a different kind of evaluative approach from that afforded by ends-means rationality in the thin sense, where the only normative

³⁰⁶Hamlin 1991, p. 375.

injunction concerns the choice of the best means to some stipulated end, whatever that end may be. Hamlin is also correct in arguing that the relevant sense of 'capture' must be one where honour figures as the interest, the end, or the goal, to which revenge may be the best means. These two points are evident and it is unlikely that anybody would want to find fault with either of them. What is more of an issue is his suggestion above that "a rational choice framework of analysis, together with the specification of honor as a primitive motivation, can offer insights into the pattern of behavior observed."

This is a point on which Elster disagrees. His reasoning is brief, but it seems to be that when revenge is a means of saving one's honour, the reasoning between the means and the end is just too direct to admit of real and meaningful rational choice deliberation for the individual. It will be worth quoting again the relevant passage from Elster:

If broader motivations [than the selfish, material interests of the agent] - like the concern to uphold one's honor - are allowed, the task of proving the rationality of revenge becomes easier but also less interesting. Whenever a norm is invoked in explaining an action, one can turn the account into a rational-choice explanation by saying that the agent is concerned with upholding his self-image and reputation as a norm follower. While true enough, such statements are uninformative. In particular, they contain no guide to how the agent could rationally choose between alternative means to the stipulated end. Many of the means, in fact, will themselves be regulated by norms.³⁰⁷

The crucial claim in Elster's argument is that for the individual faced with the task of saving self-esteem, reputation and honour, there is no room left for (rational choice) deliberation. For, if an individual is to save honour, as a norm follower he must take revenge. The accepted norms will not allow any other way around it. And there can be no room for deliberating between revenge and other alternatives when there are no acceptable alternatives for achieving the proposed end. However, this crucial line of reasoning does not seem to register with Hamlin who tries to show that rational choice explanations can help us understand how norms of revenge could be a cost effective

³⁰⁷Elster 1990, p. 872.

mechanism for producing and maintaining honour in societies in which honour is important. He argues that

In a society in which honor is important and individuals are rational, an honor-producing or honor-supporting mechanism has the aspect of a public good. ... These societies will realise honor in a variety of ways. A norm of revenge is one such way. Norms of revenge are, therefore, consistent with the rational choice model provided that it has a rich enough conception of basic human motivation, and a rich enough appreciation of history and social circumstance.³⁰⁸

There are two reasons why Hamlin's arguments are open to criticism. One is that he does not confront the question of whether honour could or should be regarded as an acceptable value worthy of pursuit. This lack of concern with evaluating the goals by reference to which the rationality of an action is to be determined is ultimately unsatisfactory from a normative point of view, even if he does show that rational choice explanatory models can offer insights into the question of why some societies have norms of revenge (while others don't). The other reason why Hamlin's reply to Elster is not quite satisfactory is that he fails to give a direct answer to Elster's charge that, given that there are norms of revenge in a society, there is no room left for rational deliberation for the individual concerned. If he is to save his reputation and honour, as a norm-follower he must take revenge because the norms in question leave no viable alternative. In other words, the argument here is that even if Hamlin has shown that rational choice explanations can offer insights into the question of why certain societies have norms of revenge, he does not show thereby that revenge can be *individually* rational for particular agents in those societies, because of the possibility that the norms regulating revenge behaviour simply leave no scope for rational choice considerations and deliberations.

But, even though Hamlin does not give a decisive response to the argument that with respect to non-material interests, rational choice explanations of *individual* revenge behaviour are unsatisfactory because they are uninformative, this line of reasoning by Elster can be shown to be defective at two points. Firstly, it is invalid. Secondly, the

³⁰⁸Hamlin 1991, pp. 380 - 81.

premise that revenge is the only way to save honour is false. I start by discussing the invalidity of the argument.

Even if it were the case that one could save honour only through revenge, it simply would not follow that there was no room left for rational choice deliberation. One can still deliberate and find out whether taking revenge (and saving honour) or not taking revenge (and losing honour) is the rational thing to do. The only way to avoid this invalidity charge would be to make the mistaken assumption that honour is the only thing that matters to an agent. There is no denying that in societies where the moral life revolves around a socially shared conception of honour, honour may be very highly valued by an agent. It may seem to us that in such cultures honour is accorded a disproportionately high value relative to the other values such as material prosperity and security. However, an awareness that in such societies honour is to people the equivalent of what dignity and self-esteem are to us should make us wary of making rash judgements in this regard. In any case, as Boehm observes, even among the early nineteenth century Montenegrins of the Balkan peninsula where blood feuds were so prominent, the “demands of honor were weighed against personal and collective considerations of energy, stress, and bodily risk.”³⁰⁹

But if honour is not the only thing that matters, then whether to take revenge or not will be best decided by rational choice considerations. Perhaps no one will disagree that deciding to take revenge is a serious business and it would be unkind of us to suppose that anyone takes lightly to it. One thinks hard about the risks involved and the probable consequences. In making the final choice at least the following are taken into account: The safety of one's own life and that of one's family, economic prosperity and the approval or acceptance of others. If so, we no longer have the simple-minded model where not taking revenge cannot be a realistic option for the individual. Rather, we have the more complex picture where these other values are also taken into account in deciding

³⁰⁹Boehm 1984, p. 152.

whether or not to take revenge. The way such decisions are made can be demonstrated on the formal level by calculating the expected value of the relevant options.

First we fix the value (V) of each of the possible outcomes ($O_1, O_2, O_3, \dots O_n$) consequent to each action that is available to the agent.³¹⁰ This will be the sum of the degrees to which the various values (such as safety, honour, economic prosperity, social approval ... etc.) are realised in each of these possible outcomes. For example, the value of one possible outcome, O_1 , is given by honour being realised in it to the degree of 10, safety to the degree of 5, prosperity to the degree of 6 and social approval to the degree of 9. The value of a different outcome, O_2 , is given by honour being realised in it to the degree of 2, safety to the degree of 7, prosperity to the degree of 8 and social approval to the degree of 3. Similarly, the value of a third outcome, O_3 , can be given by honour being realised in it to the degree of 3, safety to the degree of 7, prosperity to the degree of 5 and social approval to the degree of 2. Thus we have: $V(O_1) = 30$, $V(O_2) = 20$ and $V(O_3) = 17$. Clearly, the list could go on.³¹¹

The next step is to fix the probability function of each of these outcomes with respect to the possible courses of action which are open to the agent. These will be the probabilities that any one of these outcomes will in fact obtain if some particular action is performed.³¹² For example the probability that O_1 will in fact obtain if revenge is

³¹⁰It is worth noting here that we are often in the dark about the *real value* of things and states of affairs. An example is the real value of the unborn human foetus. While some anti-abortionists place almost infinite value on it, many pro-abortionists radically disagree. It is questionable, however, that any of us is in a position to claim certain knowledge in this area. Hence, it is important from a normative point of view that agents do not exhibit culpable ignorance about real values. (I owe this point to Graham Oddie who first drew it to my attention in 1988 in a context not related to this thesis.)

³¹¹It is assumed here that values are commensurable in the required sense, and that it makes sense to compare them and trade them off against each other. For a defence of value commensurability, see Griffin 1986 (especially Ch. 5.).

³¹²On the question of whether the probabilities in question should be subjective or objective, I concur with Oddie and Menzies that while maximising objective value is "the correct regulative ideal," this is not in itself an "epistemically sensitive selection rule." In view of the fact that people are often in the dark, not only about the real value of things and states of affairs, but also about the objective chances that certain states of affairs will obtain, in making a choice, "an agent must use the partial information in her possession to provide best estimates, in the light of her information, of the objective value of her choices." (Oddie & Menzies 1992, p. 532.) See also Oddie 1996.

attempted may be, say, .5 and it may be as low as .1 if no revenge is attempted. [Symbolically, $P(O1/R) = .5$ and $P(O1/\sim R) = .1$] Similarly, the probability that O2 will in fact obtain if revenge is attempted may be as low as .2 and it may be somewhat higher, say .4, if no revenge is attempted. [$P(O2/R) = .2$ and $P(O2/\sim R) = .4$] And the probability that O3 will in fact obtain if revenge is attempted may be .3, perhaps, as opposed to .5 if no revenge is attempted. [$P(O3/R) = .4$ and $P(O3/\sim R) = .5$]

Once we have all the relevant value and probability assignments, the expected value (EV) of any particular action can be calculated by summing up the values (V) of each possible outcome (Oi) after they have been weighted by the appropriate probabilities that they will in fact obtain if the particular action (A) is performed ($P(Oi/A)$). In other words, the expected value of some action is the probability-weighted average of the values of all the possible outcomes that can obtain if the action is performed. Symbolically: $EV(A) = \sum_i V(Oi) \times P(Oi/A) = [V(O1) \times P(O1/A) + V(O2) \times P(O2/A) + V(O3) \times P(O3/A) + \dots + V(On) \times P(On/A)]$

In fact, the details given above as examples can be usefully summed up in a table to give us a simple case with three possible outcomes as follows:

Values	01	02	03
Honour	10	2	3
Safety	5	7	7
Prosperity	6	8	5
Approval	9	3	2
Sum values	30	20	17
$P(O_i/R)$.5	.2	.3
$EV(R) =$	$15 + 4 + 5.1 = 24.1$		

Revenge

Values	01	02	03
Honour	10	2	3
Safety	5	7	7
Prosperity	6	8	5
Approval	9	3	2
Sum values	30	20	17
$P(O_i/\sim R)$.1	.4	.5
$EV(\sim R) =$	$3 + 8 + 8.5 = 19.5$		

No Revenge

Figure 8

In this simple example the expected value of attempting to save honour through revenge happens to be somewhat higher than the expected value of not attempting to save honour at all. But, of course, this need not be the case all the time. Both the value and the probability assignments following an offence are bound to differ from case to case and, therefore, there is always an interesting and useful decision-theoretical calculation to be made to find out whether it is worth trying to save one's honour through the only way possible or not.

I now turn to the second reason for the failure of the claim by Elster that rational choice explanations of individual revenge behaviour with respect to non-material goals are uninformative. I have assumed so far that taking revenge is the only way to save honour. But this assumption, which is the crucial premise in Elster's argument, is false. Even in societies where a mostly revenge-based conception of honour is accepted, there are a number of ways allowed by the codes of honour beside revenge in which the honour of the offended party can be saved. Some of these are: obtaining an apology for the insult, obtaining material compensation - like blood money, convincing the offending party to commit suicide, and possibly others. But even so, in some cases at least, the option with the highest expected value may well turn out to be revenge. This certainly seems to be the case among the Tausug:

It is important to note . . . that for the Tausug, the use of violence in retaliation for an injury to a person's body or honor is not the only means afforded by his society and culture. There are alternative means of conflict resolution such as mediation, amicable settlement usually with the interference of a third neutral party, the use of money settlement, and resort to supernatural or ritualistic methods. However, for the Tausug, especially in cases of murder or homicide, a violent confrontation would seem to be the only effective and realistic solution to the larger problem of justice and social order. As some people say: "It may not be legal to retaliate, but it takes so much time, effort, and money to take cases to court. Besides, we are not assured of justice."³¹³

Furthermore, even if the probability that honour will be successfully saved was the same in each of the different options, there still is a rational choice calculation to be made along the lines suggested above in order to select the option with the highest expected value. This is because although it is true enough that in certain cases the *only* reason behind wanting to take at least one of these courses is to save honour, in making the final choice there is much more to take into account than that. As we have already seen, the security of one's own life and that of one's family, one's economic prosperity and the approval of others also have to be taken into account in order to calculate the expected values of all these options.

³¹³Tan 1981, p. 55.

Moreover, even if the only thing to take account of was saving honour, selecting an option randomly without rational deliberation would be silly all the same. For the circumstances following an insult differ from situation to situation and each of the options has a different probability of being successful at achieving the proposed end. If you know, for example, that your aggressor is extremely strong and well-prepared for a retaliation but you also know him to have been at fault and to be very reasonable and that he considers himself to be a true man of honour, attempting to take revenge may not have the highest expected value because of the low probability that it will succeed. In such a situation asking a third party to mediate between the two of you with the aim of eventually receiving an apology may well turn out to be the way to go because of the high probability that it will succeed in preserving your honour.

On the other hand, if you know that your aggressor is extremely arrogant and vain, you would be silly to demand an apology because the probability that you will get it is very low. Further, if your aggressor is also known to be a distinguished miser, it is unlikely that he will pay blood-money. In such a case, depending on the other circumstances - like the ability and the preparedness of the aggressor to defend himself, for instance - revenge may well turn out to have the highest probability in saving your honour and, hence, be the option with the highest expected value.

It is well to recall in this regard the story of the 47 Ronin.³¹⁴ The accepted codes of honour prevailing in 18th century Japan allowed for the saving of their insulted master's honour (and their own) by successfully requiring Lord Kira to acknowledge himself to have been at fault and commit ritual suicide. However, they knew well that Lord Kira was not only very powerful but also extremely vain and arrogant - so that the probability that he would apologise if asked to do so was very small. So instead of foolishly demanding an apology and thereby putting Lord Kira on the alert, the 47 went to great lengths to convince everyone that they cared nothing about the loss of honour

³¹⁴Benedict 1977, pp. 140 - 44.

which resulted from the insult. As they expected, after some time Lord Kira did in fact drop his guard and the 47 were swift with their much recited and venerated vengeance on him for an insult.

5.6 Conclusion

In this chapter and the previous chapter I considered the rational status of revenge. In this chapter I focused on the arguments put forward by Elster against the individual rationality of revenge behaviour. I gave a detailed presentation of Elster's overall position and identified four problem areas for his arguments. First I argued against his claim that revenge is backward-looking and that, therefore, it cannot be rational. Following this, I disputed his claims that honour is universally present in revenge and that it is the key to understanding revenge. I then argued that when honour is an essential part of people's sense of identity in the epistemic sense, it is a legitimate and defensible value and goal for them to have from the point of view of a substantive, normative rationality.

I further pointed out that, when combined, the remaining two claims are suggestive of a dilemma: Either revenge is rational because it is in the material interests of the agent or it is rational because it is in the non-material interests of the agent. According to Elster, the former case is false since revenge can never be in the material interests of the agent, and the latter case, if true, is trivial and uninformative. I argued against both these claims and showed that revenge can be in both the material and non-material interests of the agent. With respect to material interests, in certain cultural contexts there is a real threat of severe social sanctioning when the individual fails to carry out the duty of revenge. In such instances revenge is individually rational in the direct sense. I also argued that revenge can be rational in the indirect sense because a revengeful attitude or disposition is in the long-term material interests of the agent. This is so because as part of a resilient TIT FOR TAT pattern of behaviour, it is the best way of eliciting cooperation from others, and thus, a better way of doing well for oneself than otherwise.

I also showed that revenge can maximise agents' non-material interests, such as the preservation of their honour, and that this is not a trivial or uninformative claim. Having shown that revenge can be rational and that it often is, I now move on to consider its defensibility and contemporary relevance to institutions of justice from a moral point of view.

6

Victim justice and the morality of institutionalised revenge

6.1 Introduction

In Chapter 2 I clarified the notion of retribution and argued that it is a mistake to confuse it either with revenge, or with what, in all plausibility, should be regarded as principles of justice. Furthermore, in Chapter 3 I argued that revenge is personal retribution and that it can take institutionalised as well as non-institutionalised forms. In this chapter I defend the moral justifiability of institutionalised revenge, by which I mean personal retribution within frameworks provided for by a legal justice system rather than informally institutionalised practices in revenge cultures.

As indicated in the main introduction, I am not advocating the unconditional, inherent moral desirability of revenge in legal institutionalised contexts. Rather, I am arguing for the moral *permissibility* of institutionalised revenge with a view to promoting victim justice, which in many cases requires acknowledgment and satisfaction of the legitimate personal needs of victims for retributive justice. This requires legal institutionalised forums where victims can express their legitimate feelings of resentment and anger, forums where they can demand satisfaction in terms of adequate restoration and in terms of apologies or retributive justice in a controlled and civilised manner for the wrongs committed against them, without their being trivialised, moralised or patronised on that account. Subject to possible vetoes on grounds of unreasonable leniency or harshness, such forums must also invest victims with the prerogative and the power to

forgive and show mercy to their offenders. These victim rights are important, I shall argue, for purposes of victim restoration in social-psychological terms and are an integral part of providing justice to victims.

To be satisfactory, however, a defence of institutionalised revenge must show three important things. First, revenge being retributive in character, the defence must justify retribution as a legitimate and morally acceptable motive which can be present in punishment. Second, revenge being personal in nature, the defence must provide good reasons for bringing the personal element back into the justice system in the form of greater victim participation and empowerment. Third, the defence must address concerns over possible violations of the principles of justice with respect to offenders, such as the requirements of proportionality and the principle of not punishing the innocent, as well as other fears and reservations about revenge, the danger of escalation being a good example. I argue in these regards that, under appropriate institutionalised conditions, the personal and retributive components in revenge neither endanger justice to offenders, nor do they lead to morally repugnant, or otherwise unwanted, results. The three main sections of the current chapter address these three broad areas of concern with respect to institutionalised revenge.

6.2 A defence of moral retributivism

It was noted above that a satisfactory defence of institutionalised revenge must show that retribution is a legitimate and morally acceptable reason for punishment. This, in turn, requires clarification of the key notion of moral retributivism. Discussion of the nature of retributivist justifications will also be useful because of a tendency by retributivists to slip into consequentialist modes of thinking. I start by covering these two points.

6.2.1 The thesis of moral retributivism

Moral retributivism, sometimes referred to simply as retributivism, is a philosophical position which offers an answer to the question: What justifies punishment? Consequentialist and other ways of connecting wrongdoing and punishment make the moral status (value/disvalue) of punishment completely dependent on its (good/bad) consequences³¹⁵ or functions.³¹⁶ By contrast, while being compatible with other connections such as the above, retributivism makes a positive, approving moral connection between punishment and moral wrongdoing, independently of the consequences or functions punishment might serve. In other words, retributivism is characterised by the thesis that there is a moral connection between wrongdoing and punishment which is somehow *intrinsic* and is not there in virtue of consequentialist or functional benefits which punishment might have. We need to be more precise, however, because moral retributivism does not concern itself with the whole domain of retributive punishment. Its claim is not that all retributive punishment has positive moral value, but only that retributive punishment within the limits set by the principles of justice, such as proportionality and not punishing the innocent, has positive moral value *qua* being punishment of wrongdoing.

This core retributivist thesis provides the basis for two familiar versions of moral retributivism which in the literature are variously referred to as ‘pure’, ‘hard’, and ‘positive’ retributivism, and ‘soft’ and ‘permissive’ retributivism, respectively. Pure, hard, or positive retributivism is the view that those found guilty of wrongdoing *should*, or *must*, be punished in proportion to the seriousness of their offence. Soft or

³¹⁵For instance, deterrence, rehabilitation of the wrongdoer, etc. A contemporary, undesirable consequence of imprisonment as a form of punishment seems to be that, far from being rehabilitated, the wrongdoer learns how to be more efficient at criminal activity. This suggests, among other things, that there is more to take into account in the justification of the modality of punishment in any given case than whether or not the mode of punishment is a humane one or not.

³¹⁶For example, expression of disapproval, symbolic and ritualised affirmation of correct values, etc. Arguably, a bad, undesirable function of (at least some) forms of punishment might be the implicit message that violence is an acceptable response to conflict situations.

permissive retributivism is the view that while those found guilty of wrongdoing *may* be punished in proportion to their crime, they do not have to be so punished.

There is, however, another kind of retributivism which could be added to these two. According to this view, although the retributive just punishment of wrongdoing is right and good in itself, mortals and their institutions possess no moral warrant to impose it on the wrongdoer.

It is only just and right, Bradley and Kant would argue, that the guilty should suffer. Yes, I agree; their punishment is morally right and just. But that proposition doesn't imply whether *we* ought to punish them or how. ... What is just when the Fates or the natural world brings it about is unwarranted - morally unwarranted - for human agents. No one and no institution is morally justified in doing such things.³¹⁷

This is a somewhat curious retributivist position. The primary purpose or function of moral judgements is to guide action, including *human* action. Indeed, moral judgements are supposed to have prescriptive force which requires agents to act in specific ways: what is judged to be good or right, morally speaking, must be pursued and whatever is judged to be bad or wrong must be avoided. But, this version of retributivism turns these fundamentals on their head. It morally forbids us doing what it judges to be morally right and good, namely the imposition of due punishment on wrongdoers. But since this form of retributivism allows the *gods* and the forces of nature to pursue or bring about what in this instance is judged to be the good and the right, it raises an important issue with regard to punishment. This is the question of who has the right to get involved in impositions of punishment. I discuss this issue in the next main section with respect to the rights of victims to be substantially involved, should they so choose, in the legal resolution of their cases.

For now, what is important to note about moral retributivism is that, although it is concerned with *just retribution*, it does not beg the question as to the need for an independent justification or rationale for the moral defensibility of just retributive

³¹⁷Wolgast 1987, pp. 160 - 62.

punishment. For the justness in this context consists solely in the non-violation of certain basic principles of justice, such as those prohibiting the overpunishment of offenders and the punishment of the innocent. Such principles do not provide reasons for imposing punishment. They only have limiting roles. They demarcate the moral limits of punishment by an appeal to independent standards of fairness and justice, standards which are, or can be, autonomous and independent of the reasons for which punishment is imposed.

This notion of just punishment is different from the notion that retributive punishment is imposed as a matter of justice, that justice demands the punishment of wrongdoing. Whatever else is going on behind such claims, it is clear that they appeal to the notion of justice more as a justificatory device for imposing punishment, rather than as a device aimed at placing limits on it. I shall discuss in more detail the notion of justice as a justificatory device in the section which follows.

I also want to suggest that, generally, no moral view or theory should be labelled 'retributive' unless it subscribes to, or implies, the thesis that punishment within the limits set by the principles of justice has positive moral value, that such punishment is right and good in and by itself, simply in virtue of being punishment of wrongdoing. As Cottingham points out, some construals and justifications of 'retribution' are blatantly consequentialist. An example is the claim that retribution is justified by the fact that punishment of wrongdoing brings satisfaction to the aggrieved, which in turn helps to prevent vendettas in the interests of social stability.³¹⁸ There is nothing retributive about such explanations or justifications. Braithwaite and Pettit have raised similar complaints and point out that, in spite of retributivists generally being critical of consequentialism, "sometimes they themselves move into a consequentialist mode of thinking" and that, even if the phrases used in framing retributivist rationales bear the mark of retributivist

³¹⁸Cottingham 1979, p. 242.

concerns, “what we are being offered may still be a consequentialist theory.”³¹⁹ Honderich’s account of retribution is a good example of this. The critical feature of Honderich’s account is the claim that “the penalty will give satisfactions equivalent to the grievance caused by his (the offender’s) action.”³²⁰ According to him, “the truth of the retributivist tradition, more precisely, is that it seeks to justify punishment partly or wholly by the clear reason that it satisfies the grievances created by offences, through causing distress to offenders.”³²¹ But, as Atkinson points out, even “garden variety” retributivists would disown this characterisation of retribution.

Although Honderich contrives to take the sting out of the retributivist position, he does so by a fundamental misrepresentation. He in fact seems to have followed just those ‘Serpentine wanderings of the Happiness theory’ that Kant so emphatically warned us against.”³²²

It would be difficult not to agree with Atkinson that through those “serpentine wanderings” Honderich lost retributivism out of his account. Honderich’s justification is recognisably consequentialist in nature. That his account is not retributive is also shown by the fact that if grievance satisfaction could be maximised through some less painful method, then punishment would no longer be justified by it. For example, a system of rewards, combined with counselling and hypnotherapy for those aggrieved, could be an alternative to punishment as far as grievance satisfaction is concerned. Moreover, there may be cases where the aggrieved party has been murdered and there is no one around who cares about him to derive the satisfactions necessary for justifying punishment along lines suggested by Honderich’s account.³²³ It is a necessary feature of a retributivist account or justification of punishment that it is able to render an offender liable to moral censure, blame and punishment simply by reference to the offence committed. The challenge facing moral retributivists, therefore, is nothing less than giving a coherent and

³¹⁹Braithwaite and Pettit 1990, p. 48.

³²⁰Honderich 1984, p. 34.

³²¹Honderich 1984, pp. 233 - 34.

³²²Atkinson 1974, p. 81.

³²³I owe this latter point of criticism to Michael Tooley.

plausible explanation of the intrinsic moral connection which it claims to exist between wrongdoing and punishment.

I do not wish to deny that good consequences, such as deterrence or grievance satisfaction, and desirable functions, such as expressing disapproval or emphatically denouncing the crime, are relevant to justifying punishment in instrumental and functional ways. My claim is that such instrumental and functionalist justifications and explanations are of no help where a justification of retributive punishment is called for. The core retributivist idea that there is an intrinsic moral connection between wrongdoing and a liability for punishment is not supported by justifications which make only a consequentialist or functionalist link between the two. In the next section I formulate a soft retributivist justification for the imposition of punishment which is not vulnerable to criticisms such as the above.

6.2.2 A moral retributivist justification of punishment

According to Hume, and argued more recently by Morris, imposing punishment on offenders is justified because criminal acts alter the moral status of wrongdoers in that such acts affect their moral rights that stand in the way of their being punished and lead to their forfeiture.³²⁴ Even if this line of argument is correct on a more general level, the task of giving a satisfactory justification of *retributive* punishment in particular requires further argumentation. This can be an intricate, if not daunting, task. In spite of there being a plethora of justifications and rationales on offer, retributivism seems to remain a largely unpopular and unconvincing theoretical position. Yet, it is likely to remain an issue on both practical and theoretical levels. For, as J. L. Mackie points out, while many people are inclined to say that

only some possible future benefit, such as deterrence or reformation, can justify punishment, ...
when we consider actual or possible cases of crimes or wrongdoing of kinds that we really regard

³²⁴Hume 1975, *Inquiry Concerning the Principles of Morals*, Section III, 'Of Justice,' Part I, § 10; Morris 1991.

as unmitigated and inexcusable, then we do after all tend to see them as in themselves calling for the infliction of some adequate penalty.³²⁵

Mackie goes further than this, however, and claims that what we have on our hands is nothing short of a paradox, the paradox of retribution.

The paradox is that, on the one hand, a retributive principle of punishment cannot be explained or developed within a reasonable system of moral thought, while, on the other hand, such a principle cannot be eliminated from our moral thinking."³²⁶

Mackie identifies the crucial, problematic retributive principle to be the view that *one who is guilty ought to be punished*, but it is clear from the context that he would regard as equally problematic the permissive variant that *one who is guilty may be punished independently of the good consequences such punishment may effect*. Mackie then mentions the major attempts to make sense of that principle as one with immediate, underived moral authority. Some of these he dismisses because of their respective shortcomings, while others he simply dismisses out of hand, a case in point being the possibility that the principle in question may be justified with reference to the idea of negative desert.³²⁷ Notwithstanding, with some unexplained leap, Mackie concludes that the principle in question cannot be explained or developed within a reasonable system of moral thought.

Having deemed moral justification of retribution an impossibility, Mackie proceeds to give a socio-biological explanation of it in a broad, emotivist framework of the nature of morality and moral judgements. According to this, ultimately, our moral thinking itself has both its origins and its substance in our retributive (both kindly and hostile) emotions. The explanation is in two broad stages. First, he gives a purely biological explanation for the tendency to feel non-moral resentment of injuries and gratitude for benefits. Second, he gives a sociological explanation for the development of moral

³²⁵Mackie 1985, p. 208.

³²⁶Mackie 1985, p. 207.

³²⁷Mackie 1985, pp. 209 - 12.

resentment from its non-moral counterpart.³²⁸ Mackie's claim is that unless we accept an emotivist framework in which the retributive *emotions* form the basis of our moral beliefs and reasoning, we cannot solve or explain the paradox of retribution. However, such a scientifically minded, socio-biological explanation of retribution is far from satisfactory because it leaves retribution in almost as much need of moral justification as before. As Mill points out, "that a feeling is bestowed on us by nature, does not necessarily legitimate all its promptings."³²⁹

Moreover, the main point here is that even if there is merit in such an emotivist socio-biological explanation, Mackie has failed to show that the principle in question can only be explained, but not properly justified.³³⁰ Apart from the fact that the list of possible justifications covered by Mackie is far from being exhaustive, many of the attempts covered by him are given very short shrift, while others simply aren't taken seriously. Given this state of affairs, he is hardly entitled to conclude that the principle in question cannot be explained or developed within a reasonable system of moral thought. Not having established this crucial thesis also raises the question as to whether there really is a paradox of retribution. Finally, even though Mackie is right in saying that a retributive principle of punishment is ineliminable from our moral thinking, absence of a paradox would render his emotivist, socio-biological explanation redundant.

I, for one, do not think that there is a paradox here. Contra Mackie, a retributive principle of punishment can be explained and developed within a reasonable system of moral thought. I suggested in Chapter 2 that while the notion of negative desert is somewhat problematic in a *definitional* role because its range does not extend beyond the limits set on punishment by the principles of justice, its proper role is a *justifying* one for

³²⁸For the details of these explanations see Mackie, 1985, pp. 215 - 18.

³²⁹Mill 1962, p. 296.

³³⁰Mackie 1985, pp. 209 - 12.

the very same reason. But, as with all other accounts of retributivism, scepticism surrounds this idea. According to Cottingham, for example,

viewed as an exercise in justification, this account shares the curiously jejune quality of theory (1) [the repayment theory]; and its detractors have made the complaint - not without force - that it reduces to the bald assertion that it is simply just that the offender should be punished.³³¹

Mackie also dismisses the idea in barely more than one line in a footnote by saying that “Desert Theory seems to be the bald, unexplained assertion that crime simply does deserve punishment.”³³² This need not be the case, however. According to Honderich, for example, there are at least nine (9) different ways of interpreting desert claims.³³³ More importantly, Robert Nozick has provided a detailed framework for retribution in an explicit attempt to answer the question: ‘In what sense is punishment deserved?’ Nozick formulates the moral retributivist thesis to be the view that “people deserve punishment for their wrongful acts in accordance with $r \times H$ [the person’s degree of responsibility weighted by the magnitude of the wrongness of the act], independently of the deterrent effect of such punishment.”³³⁴ Far from denying the need for a justification or rationale, Nozick then goes on to give an interesting and fairly complex rationale for the view by reference to the idea of connecting up the wrongdoer with correct values:

The wrongdoer has become disconnected from correct values, and the purpose of punishment is to (re)connect him. It is not that this connection is a desired further effect of punishment: the act of retributive punishment itself effects this connection. ... When he undergoes punishment these correct values are not totally without effect in his life (even though he does not follow them), because we hit him over the head with them. Through punishment, we give the correct values, qua correct values, some significant effect in his life, willy-nilly linking him up to them.³³⁵

Whether this is a plausible way of explicating desert claims is a question that need not be decided here and it is not my intention to defend Nozick’s rationale against

³³¹Cottingham 1979, p. 239.

³³²Mackie 1985, p. 209.

³³³Honderich 1984, pp. 26 - 33.

³³⁴Nozick 1981, p. 366.

³³⁵Nozick 1981, pp. 374 - 75.

criticisms.³³⁶ I mention it to make the point that desert claims do not always come as bald, unexplained assertions in the way alleged by Cottingham and Mackie.

What I propose now is to give a moral explanation of retribution along the lines suggested and developed by Oldenquist. This explanation appeals to notions of individual flourishing, life within a moral community and personal accountability, and the link that exists between them. I propose this explanation to be an explication of such shorthand, retributive justificatory claims as *'Retributive punishment is properly due as repayment for wrongs,'* *'Retributive punishment is justified because it is deserved,'* and *'It is only fair, just and right that an offender should be punished.'* Such claims are often intended to justify a retributive approach to punishment. As they stand, however, they are not very satisfactory and need further clarification and explanation. The following is an explication of such claims.³³⁷

- (1) Humans are innately social who can flourish and achieve their full humanity and potential in terms of moral maturity, only in society.
- (2) A human society is a moral community.
- (3) A moral community is such that members hold one another accountable for wrongs to fellow members and to the common good.

³³⁶See, for example, Ten 1987, p. 45; Oldenquist 1988, pp. 470 - 71. In addition to these criticisms we could ask of Nozick, tongue in cheek, whether there is a point or further rationale to connecting someone up with correct values by hitting him over the head with them.

There is an interesting parallel here between Nozick's rationale and Kafka's story "In the Penal Settlement" in which punishment consists in literally inscribing messages about values into the living flesh of the condemned person - such as, for example, HONOUR THY SUPERIORS and BE JUST. (Kafka 1961.) But, while the understanding or "enlightenment" which is generated in the person by the slow and painful carving of the message deep into his body seems to be the central point of the punishment in Kafka, Nozick's rationale for retribution does not incorporate an analogous point about effecting in the wrongdoer through punishment similar realisations about correct values.

³³⁷For the following, I am indebted to Oldenquist's first argument in the article "An Explanation of Retribution." 1988, pp. 464 - 65. However, there are at least two important difference between us: (1) While Oldenquist uses the notion of 'harm' in his argument, I appeal to the stronger notion of a moral wrong. (2) Oldenquist refers to desert as part of his explanation of retribution, while I take desert to be the object of the explanation. This ought to put beyond doubt the fact that these accounts are not vulnerable to the criticism by Cottingham and Mackie that desert claims are "jejune" or mere "bald assertion[s]."

- (4) To hold persons accountable for wrongs to fellow members and to the common good is to consider them liable for blame and punishment for such wrongs, independently of functionalist and instrumental considerations, such as expressing disapproval or deterrence - though obviously such considerations are not irrelevant to impositions of punishment.
- (5) To consider persons liable for blame and punishment for wrongs independently of functionalist and instrumental considerations is morally to accept retribution.

Using this explanation as part of an argument, there are two conclusions which follow:

- (6) Human individuals can flourish and achieve their full humanity, including moral maturity, only if they morally accept retribution and retributive liability for their wrongful actions.
- (7) Since individual flourishing and the achievement of one's full humanity, including moral maturity, are good things worthy of being pursued, retributive punishment within the limits set by the principles of justice is also a morally good thing which may be pursued and, other things being equal, ought to be pursued.

It is important to note that this account does not fall into the trap of consequentialism. The central claim here is not a consequentialist one, that a good consequence of retributive punishment is individual human flourishing. Rather, it is the claim that life within a moral community with its attendant responsibilities and a liability for retribution within limits set by the principles of justice are ineliminable aspects of human flourishing and the achievement of full humanity, a very important part of which is moral maturity. The account is not jejune either. It is not a kind of "unexplained assertion" which Mackie and Cottingham complained of. Nevertheless, the crucial analysis of personal accountability in premise four would benefit from the following explanation.³³⁸

Recognisably, personal accountability is one aspect of moral maturity in terms of moral responsibility. It is the retrospective part of it. When someone is considered

³³⁸There are also other important aspects to moral responsibility, such as free will and choice, which are not immediately relevant to this discussion and which fall outside the scope of this thesis.

morally responsible for something, typically, there are two aspects to that responsibility. One is a prospective, forward-looking aspect which places that person under some kind of obligation or duty. But tied to such duties and obligations is a retrospective accountability which renders the person liable for moral censure in the form of condemnation, blame, and possibly punishment, in the event the duty or obligation is not carried out.

For example, owning and using such items as guns and motor cars place people under moral, as well as a legal, obligations to take all reasonable precautions in order to avoid injuring other people. Maintaining one's car in a roadworthy condition and not driving under the heavy influence of alcohol are specific instances of such obligations. But such duties and obligations constitute only half of the story of what it is to be morally responsible as a car driver. The other half of the story is told, and it has to be told, in the event one fails to fulfil one's duties and someone gets hurt as a consequence. If George keeps driving his car with faulty brakes, or if he decides to go for a "spin" in his new car under heavy influence of alcohol, and runs down a pedestrian on a zebra crossing, it will not suffice as a response simply to shrug our shoulders, or say something glib like 'it is just too bad for that pedestrian and for poor old George for ruining his car as a result of ignoring his forward-looking responsibilities as a driver.' His moral responsibility as a car driver exposes him to legitimate, *retrospective* moral censure, condemnation, blame and possibly punishment for his moral failure.³³⁹

Similarly, if George decides in a moment of frustration to use his neighbours for target practice, it will not suffice as a response to say something even more glib than before, like 'it is too bad for the neighbours that they happened to be home when George lost his temper.' His moral responsibility as a member of the community exposes George to legitimate, *retrospective* moral censure, condemnation, blame and punishment

³³⁹Husak (1994) has argued that drunk driving is not as serious an offence as people generally believe it to be. He does not dispute, however, that when disastrous consequences result because of drunk driving, the responsibility of the driver renders him liable to blame and punishment in the way suggested above.

for his wrongful, criminal behaviour - and to an even greater degree if it resulted in injury or death to others.³⁴⁰

What gives legitimacy to these responses is the responsibility George, as a whole-minded, mature individual, bears morally (and legally) for his actions, and quite independently of instrumental considerations. The possibility that, in addition, these responses might also have some instrumental value is neither here nor there. The expected consequences of some proposed punishment should, by all means, be taken into account. But, as the above examples illustrate, it would be mistaken to think that consequentialist considerations were the only ones to have relevance. Moral censure, blame and punishment can be mandated by the person's retrospective accountability for things done wrong and for which, as a mature member of a moral community, he or she bears a degree of responsibility.³⁴¹

This line of argument is not new. Many philosophers of distinction have argued that retributive accountability must be taken as the basis of punishment, if the dignity and inherent value in a human being is to be honoured and treated with due respect. Thus, Kant argues against consequentialist justifications of punishment because such punishment treats the offender as a mere means to an end.

Juridical Punishment can [should] never be administered merely as a means for promoting another Good either with regard to the Criminal himself or to Civil Society, but must in all cases be imposed only because the individual on whom it is inflicted *has committed a Crime*. For one man ought never to be dealt with merely as a means subservient to the purpose of another, nor be mixed up with the subjects of Real Right. Against such treatment his Inborn Personality has a Right to protect him, even although he may be condemned to lose his Civil Personality. He must first be found guilty and *punishable*, before there can be any thought of drawing from his Punishment any benefit for himself or his fellow-citizens.³⁴²

³⁴⁰This last claim may raise the question of moral luck - an interesting question which falls beyond the scope of the current project.

³⁴¹Retrospective moral accountability is rarely strict, however. We do recognise certain excuses for failure to fulfil an obligation. Rather, the line is drawn at, or with, the notion of 'fault' - discussion of which is beyond the scope of this project.

³⁴²Kant 1974, p. 195.

Similarly, Hegel argues that

... right and justice must hold their seats in freedom and volition, and not in unfreedom which is addressed by threat. In this [latter] way punishment is established as if we raised a stick at a dog, and man will not be treated according to his honour and freedom but like a dog.³⁴³

Also, in his paper “The Humanitarian Theory of Punishment,” C. S. Lewis argues in defence of a retributive liability for punishment and against the deterrent and rehabilitative reasons proffered by theories of instrumental punishment for which, somewhat unfortunately, he uses the “humanitarian” label. Like Hegel, Lewis argues that punishing someone for the sake of good consequences, rather than for retributive considerations based on retrospective responsibility for his or her wrongful action, denies the inherently moral character of those actions. It denies the offender’s status as a dignified, fully responsible human being.

Their very kindness stings with intolerable insult. To be “cured” against one’s will and cured of states which we may not regard as disease is to be put on a level with those who have not yet reached the age of reason or those who never will; to be classed with infants, imbeciles, and domestic animals. But to be punished, however severely, because we have deserved it, because we “ought to have known better,” is to be treated as a human person made in God’s image.³⁴⁴

Though they argue in their own respective ways, Kant, Hegel and Lewis clearly converge on the central claim of the earlier outlined argument for retribution. This is the claim that retributive responsibility for one’s actions is an ineliminable part of achieving one’s full humanity. A refusal to hold someone retributively responsible for his or her wrongful actions is tantamount to refusing to grant him or her the status of being a mature, fully developed human individual. From this perspective it is understandable why Lewis finds the instrumentalist’s “humanitarian” kindness in matters of punishment so repulsive. Their kindness “stings with intolerable insult” because it is a kindness based on a paternalistic disrespect for the offender. By regarding the offender to be someone in need of rehabilitation, or as someone who can be used as an instrument for

³⁴³Hegel as quoted by Heller 1987, p. 160.

³⁴⁴Lewis 1949, p. 10.

purposes of deterrence, or both of these, the instrumentalist fails to respect and fails to take the offender seriously as a fully developed, mature, morally responsible person. While there may be cases where such an attitude may be warranted, it is indeed intolerably disrespectful to whole-minded, fully responsible adult individuals.

6.2.2 Mercy and retributive liability for punishment

None of the foregoing arguments for moral retributivism are incompatible with the importance of forgiveness and mercy. A recognition that offenders bear a retributive liability for their wrongful acts does not in any way entail that, therefore, come what may, they must be punished. Such retributive accountability does not undermine in any way the moral right, the power, and the prerogative of victims, or of the wider community, to forgive the offence and show mercy. As Garcia also argued,

since desert renders punishment *prima facie* permissible but not *prima facie* obligatory, it poses no real danger to the normal acceptability (indeed, the praiseworthiness) of mercifully foregoing punishment.³⁴⁵

However, we can, and perhaps we must, go further than the above passage from Garcia suggests. For, far from being inimical to mercy, a retributive *liability* for punishment is a precondition for it. This was recognised by Lewis who, while being a person of deep Christian convictions and aware of the value of forgiveness and mercy, was, at the same time, a fierce defender of moral retributivism - a view for which, incidentally, he could get no hearing in England.³⁴⁶ It will be appropriate, therefore, to give a more rounded view of his position on these matters with a quote.

... I think it essential to oppose the Humanitarian theory of Punishment, root and branch, wherever we encounter it. It carries on its front a semblance of Mercy which is wholly false. ... The essential act of Mercy was to pardon; and pardon in its very essence involves the recognition of guilt and ill-desert in the recipient. If crime is only a disease which needs cure, not sin which

³⁴⁵Garcia 1989, p. 264. See also Garcia 1986.

³⁴⁶Lewis concludes his article by saying: "You may ask why I send this to an Australian periodical. The reason is simple and perhaps worth recording; I can get no hearing for it in England." (p. 12.)

deserves punishment, it cannot be pardoned. How can you pardon a man for having a gum-boil or a club foot?³⁴⁷

Similarly, Armstrong argues that

The retributive theory is not ... incompatible with mercy. Quite the reverse is the case - it is only the retributive idea that makes mercy possible, because to be merciful is to let someone off all or part of a penalty which he is recognised as having deserved.³⁴⁸

Notwithstanding these arguments, it is still an open question whether on a practical level any particular individual chooses to respond retributively to wrongdoing, or with mercy and forgiveness. Perhaps there is even a legitimate sense in which mercy is somehow superior to, or more desirable, admirable and noble than, retributive justice - and I do not wish to dismiss out of hand a philosophy of unconditional, universal non-violence or pacifism, or even a Nietzschean ideal of overcoming our condition by becoming "*too pure* for the filth of the words: vengeance, punishment, recompense, retribution."³⁴⁹ Yet, on a practical level the moral and practical appropriateness of mercy do seem to depend very much on the circumstances of the case. It would be most inappropriate, it seems to me, if not repugnant, to go about unconditionally forgiving anything and everything everywhere at all times. Instrumental considerations alone seem to require that there must be some form of appropriate punishment for such crimes as, say, rape, child abuse and murder. To do otherwise is to breed moral monsters devoid of respect, let alone empathy, for the rights and needs of others and the suffering they cause through their actions - moral monsters who simply do not, and perhaps cannot, care. Confucius answered well, I believe, when he was asked:

'What do you say concerning the principle that injury should be recompensed with kindness?' The Master said, 'With what then will you recompense kindness? Recompense injury with justice, and recompense kindness with kindness.'³⁵⁰

³⁴⁷Lewis 1949, p. 11.

³⁴⁸Armstrong 1969, p. 155.

³⁴⁹Nietzsche 1993, p. 116. (*Thus Spake Zarathustra*, Chapter 27: "The Virtuous.")

³⁵⁰As quoted by Spencer 1966, p. 366.

Mercy, it seems, can be rendered appropriate by considerations such as remorse, as well as inappropriate by a lack of it and by misguided defiance, for instance. Its intrinsic and instrumental virtues also have to be weighed against other considerations which may have a legitimate bearing on how a particular case should be resolved. Thus, a need for retributive justice, not unlike a need for deterrence, may well conflict with granting mercy in particular cases and, it seems that each case is best decided individually on its own merits. As to how we should think on the general level of the tension between the competing values of justice and mercy and the way in which this tends to be handled in a climate of anti-retributivist justifications of punishment, Lewis may well have pointed us in the right direction.

[T]he Humanitarian theory wants simply to abolish Justice and substitute Mercy for it. This means that you start being “kind” to people before you have considered their rights, and then force upon them supposed kindnesses which they in fact had a right to refuse, and finally kindnesses which no one but you will recognise as kindnesses and which the recipient will feel as abominable cruelties. You have overshot the mark. Mercy, detached from Justice, grows unmerciful. That is the important paradox. As there are plants which will flourish only in mountain soil, so it appears that Mercy will flower only when it grows in the crannies of the rock of Justice: transplanted to the marshlands of mere Humanitarianism, it becomes a man-eating weed, all the more dangerous because it is still called by the same name as the mountain variety. But we ought long ago to have learned our lesson. We should be too old now to be deceived by those humane pretensions which have served to usher in every cruelty of the revolutionary period in which we live. These are the “precious balms” which will “break our heads.”³⁵¹

To summarise, on the basis of the conceptual analysis I provided in Chapter 3, I argued there that revenge is personal retribution. I also argued there that, to the degree that institutionalised personal punishment is done with retributive reasons, it qualifies, and can properly be referred to, as *institutionalised revenge*. I have argued in this section that retribution is a legitimate and morally justifiable consideration in punishment. I first clarified the notion of moral retributivism and then provided an argument in defence of it. I argued that liability for retributive punishment within limits set by the principles of

³⁵¹Lewis 1949, pp. 11 - 12.

justice is part of what it means to be a fully responsible member of a moral community, and that it is part of what it means for an individual to flourish as a moral person and achieve full humanity. The next thing which needs to be shown for a moral justification of institutionalised revenge, therefore, is that there are good reasons for putting the personal element back into punishment . I take up this task in the section immediately following.

6.3 Victim justice and empowerment

The following statement from Abel and Marsh is a fair description of the lack of status accorded to victims by contemporary criminal justice systems.

Our present criminal justice system has, in effect, no law relevant to victims; it leaves them as they are, ignoring their plight in its zeal to punish, deter, and rehabilitate.³⁵²

In their turn, Abel and Marsh argue for making restitution to victims the primary focus of criminal law. While theirs is a positive step forward from a victim justice point of view, it is unlikely to go far enough to be satisfactory. This is because victims' needs to be heard, respected, appreciated and recognised as an important participant in the judicial system often matters to them just as much as does the receiving of adequate compensation.³⁵³ Accordingly, in this section I argue for making legal institutionalised punishment more personal by way of greater victim participation in the punishment process. Some of the reasons I provide for greater victim participation are independent of retributive considerations. This has the desirable result that my arguments support greater victim involvement in the process regardless of whether retributive considerations are present in a particular case or not. Victim empowerment through constructive participation, I shall argue, is very important from the point of view of victim justice as victim restoration. I cover two main models for substantive victim involvement. My arguments for them spring from considerations of victim rights and justice to victims in

³⁵²Abel and Marsh 1984, p. 160.

³⁵³Erez 1991, p. 2.

the form of restitution, compensation and restoration through victim empowerment and victim satisfaction with the legal processing and final resolution of their cases.

Certain basic human rights, such as the right to a fair trial, are specific to punishment contexts. Moreover, some such rights are specific to people accused of wrongdoing, for example, their right to a fair trial and, in the event of being found guilty, not to be punished out of proportion to the moral gravity of the offence committed. But victims also have comparable rights. It is at least arguable that victims have an inherent, non-instrumental right to participate actively, and have a say, in the resolution of offences committed against them. After all, the incident is *their* business first and foremost, and it is the rest of society's only after that. In a great number of cases of criminal wrongdoing the real or primary injury is to the victim and society is harmed, if at all, only secondarily and derivatively by it.³⁵⁴ As one victim put it, "Why didn't anyone consult me? I was the one who was kidnapped, not the state of Virginia."³⁵⁵ Indeed, as Friedman noted,

The idea that law is primarily private, that most offenses are offenses against specific individuals or families, and that punishment of the crime is primarily the business of the injured party seems to be common to many early systems of law ...³⁵⁶

The idea of greater victim participation is further supported by considerations of victims' rights to restitution, compensation and restoration in social and psychological terms. Broadly construed, their right is to be restored as much as possible to their original, pre-victim position economically, physically, psychologically and socially. Such victim rights are clearly spelled out in the *United Nations Resolution on Victims of Crime and Abuse of Power*.³⁵⁷ Victims have the moral right to get all the help and opportunities they need to rebuild their shattered lives as quickly as possible. To deny them this right is to deny them justice - justice which can be regarded as the other side of

³⁵⁴Braithwaite and Pettit, 1990, p. 91.

³⁵⁵As quoted in Erez 1991, p. 2.

³⁵⁶Friedman 1979, p. 401.

³⁵⁷General Assembly 1985.

the coin of criminal justice: *criminal victim justice*. This is a concept which needs to be explored in detail.

Criminal injuries compensation schemes are an important step towards restitution and compensation to victims for harms and losses sustained by them. However, while such schemes are necessary and welcome, they are not normally part of the judicial process which deals with the offence in question. Thankfully, there are exceptions to this, such as the one in Hanover where the Association for Conflict Mediation and Redress

manages a fund for victims, from which offenders without means may receive an interest-free loan in order to pay compensation to the injured party. The offenders then either pay the money back through instalments or perform community service at a fixed hourly rate.³⁵⁸

There is a lot to be said in favour of such schemes. Unfortunately, in countries like Australia, for example, compensation is normally considered and granted by the State in complete isolation from, and well after, judicial proceedings have been finalised and the case closed. These schemes, therefore, do not affect in any way the lack of status accorded to victims of crime in the judicial process. But, as indicated above, research findings consistently show that victims' needs to be heard, respected, appreciated and recognised as an important and necessary participant in the judicial system is not less important to them than is the punishment of their offenders or the receiving of adequate compensation for losses and traumas endured.³⁵⁹ Consequently, victim dissatisfaction with judicial processes which do not accord them a role commensurate to their status as the persons who in fact have been wronged and whose lives have been the most affected by the crime is understandable and invites proper consideration.

It is not difficult to see why, and how, victim dissatisfaction with the judicial processing and substantive resolution of their cases can be an obstacle to recovering their damaged sense of security, self-esteem and peace of mind which are so necessary for the

³⁵⁸Severin 1995, p. 21.

³⁵⁹Erez 1991, p. 2.

quiet enjoyment of their lives. Satisfaction in the above terms is essential to engendering in victims a feeling that they can now put behind them whatever happened and get on with their lives, looking once again to the future, rather than being continuously preoccupied with the hurtful past because they feel that, in all fairness, justice still hasn't been done. In the absence of a feeling that their cases have been dealt with in a satisfactory manner, their recovery is likely to be protracted, if not impossible to achieve.

A crucial element of victim justice is victim restoration in the form of emotional and psychological recovery following an offence. In turn, victim restoration must start with resolving the conflict with as much satisfaction to the victim as possible. It is absolutely crucial in this regard that the legitimacy of the victims' feelings of resentment and, where appropriate, their need for retributive justice, receive proper acknowledgment, respect and due consideration, rather than dismissal or trivialisation. This is especially important, I suggest, in serious crime, such as assault, rape and murder.³⁶⁰ Otherwise victims are likely to feel victimised twice over: first by the original offender, and then by the legal system. It is the common experience of raped women, for example, that they are the ones on trial in the courtroom. As Holmstrom and Burgess concluded in their study *Rape: The Victim Goes on Trial*,

The rape victim is treated in the courtroom as if she had committed a crime. Technically, only the defendant is on trial. But socially and psychologically the victim is on the firing line too. It is her reputation against his. The defense lawyer's strategy is to blame the victim. ... This strategy is devastatingly effective, as the attrition rate in the present sample shows. Sixty-one rape victims were admitted to the hospital. Twenty cases went through lower court, and seven through superior court. There was one conviction for rape: The man pleaded guilty.³⁶¹

³⁶⁰In line with established usage, a victim is any person who suffers harm as a result of an offence. In cases, such as murder, "any person who was financially or psychologically dependent on the primary victim immediately before his or her death", is also a victim. (ACT *Victims of Crime Act 1994*, No. 83. § 3,1.) Accordingly, considerations of victim involvement and empowerment are as applicable to them as they are applicable to (living) primary victims. It would be implausible, and wrong, to say that in a case of murder, for example, there are no living victims to consider. This is acknowledged not only in legislation such as the above, but also by the *United Nations Resolution on Victims of Crime and Abuse of Power*. (General Assembly 1985.)

³⁶¹Holmstrom and Burgess 1975, p. 46.

But even more generally, including cases where convictions are in fact obtained, victim accounts of their experiences with the judicial system clearly indicate that unsatisfactory processes and resolutions generate severe feelings of injustice and, instead of soothing, tend to aggravate their feelings of disillusionment, anger, bitterness and resentment.

Where is the justice in the law that allows none of us to speak of the unimaginable horror of a father finding his dead daughter. ... At times it felt as though we were being punished twice, once by having our daughter murdered, and then having to go through the ordeal again in court.

I feel very angry and bitter about this and I would wish that the legal processes could be changed so that the families who are hurting so much could feel as though justice was indeed being done. To us, the victims, it appears as though the law does everything to protect the criminal and forgets about his victims.³⁶²

Contrast this with the satisfaction level of a murdered young man's mother who was at least allowed to confront the offender with the consequences of his act in court.

For the first time in Queensland legal history, a sobbing mother was yesterday allowed to confront her son's murderer in court and tell of the tragedy caused by the killing. ... Mrs Ward spoke for about 10 minutes from a typed statement, which had been approved by Justice Paul de Jersey who said he thought the courtroom was "not inappropriate for those convicted of murder and the public to be reminded of the consequences of these terrible crimes". ... Outside the court, Mrs Ward said she wanted to congratulate Justice de Jersey on taking Queensland a great step forward by allowing her to address her son's killer.³⁶³

Such accounts indicate that giving victims the opportunity to take a greater role in the resolution of their cases has strong potential to increase their satisfaction levels and, thus, aid their recovery and restoration. Giving them a proper say in the resolution of their cases would be an act of empowering them in the sense that it would place them in a position of once again being in control of their lives and their situation, which has been wrongfully and grievously disrupted by the wrongdoer. It would also empower them to

³⁶²Harvey, R. How Murder Hits the Victim's Kin. *VOCAL Newsletter (ACT)*. 1994 - 95. Vol. 5. No. 3 - 4. pp. 14 - 15.

³⁶³Court Allows Mum to Confront Son's Killer. *The Courier Mail*, Saturday, November 11, 1995, p. 3.

have, and feel, some control or influence over their wrongdoer's situation and, thus, aid them in re-establishing power equality vis-a-vis their offenders. As I argue the point in more detail below, this is likely to aid powerfully the return of their confidence and self-respect. As Braithwaite put it in the context of a need to empower similarly disempowered offenders, "coherent strategies of empowerment can succeed even with the most powerless."³⁶⁴ If this is right, denying victims the opportunity to participate in the resolution of their cases is inimical to according them justice.

What can be most profitably focused on here, perhaps, is that offences occur in social contexts. In cases where victim(s) and offender(s) are clearly identifiable an offence is a kind of interpersonal exchange, a social (or rather anti-social) exchange between offender and victim. A very important, if not the most important, element of such an exchange is that the victim suffers undue and unjustifiable harms and losses in material and psychological well-being, a loss of trust and sense of security. Their self-esteem is also likely to be seriously threatened and undermined. A brief look at the nature and psychology of interpersonal relationships and exchanges should help in appreciating the importance of substantial victim involvement and empowerment in the interests of victim restoration.

Although communal relationships such as between partners and parent-child, for example, do not depend so heavily on reciprocity between the partners,³⁶⁵ the idea that equality, reciprocity and fairness ought to characterise human interactions is deeply ingrained in the human psyche. As I go on to explain in more detail, research in social psychology shows that equality and reciprocity are among the healthiest and most desirable features which can, and ought to, characterise relationships and interpersonal exchanges. It is a widely held belief among exchange theorists that relationships are in general more satisfying and stable when reciprocity is perceived, and when the rewards

³⁶⁴Braithwaite 1995, p. 68.

³⁶⁵Clark and Mills 1979.

for each partner are perceived to be more or less equal.³⁶⁶ For example, in relationships between colleagues at work “concerns of reciprocity are of paramount importance.”³⁶⁷ In many different kinds of relationships, being over-benefited, as well as being under-benefited have been found to create negative feelings, distress, loneliness and a reduced sense of well-being. This is as much true of able bodied individuals as it is of people with disabilities who have also been found trying to maintain reciprocal supportive relationships.³⁶⁸ The evidence suggests that, generally, giving more support than one receives leads to feelings of unfairness and resentment, whereas receiving more support than one gives results in feelings of guilt and shame³⁶⁹ as well as to feelings of obligation and owing, uncertainty and fear of not being able to repay the debt.³⁷⁰

Two crucial notions which must be given key importance at this point of the explanation are respect from others and self-worth, or self-esteem. To be treated with respect by others is to be treated justly by them. Respectful, just treatment entails nothing short of what is regarded as equitable, reciprocal and fair treatment. Behaviour which contravenes (one or more of) these values is disrespectful and can easily arouse feelings of indignation, resentment, anger and, not uncommonly, an urge to retaliate in order to even the scores with the wrongdoer. As Kim and Smith observed,

Most people believe that they deserve respectful treatment from others. Cultural norms dictate that, fundamentally, everyone should be of equal worth as a human being. Thus, if the unjust act of a harmdoer serves to violate these beliefs, the threat and assault to the self can be severe. A large part of the message carried by the unjust act is that the victim is inferior and undeserving of respect. ... Particularly when the wrongdoer’s harm diminishes our self-worth, we are all the more angered and all the more compelled to avenge injustice, even at high personal cost.³⁷¹

³⁶⁶Buunk, Doosje, Jans, & Hopstaken 1993, p. 801. See also, LaGaipa 1977.

³⁶⁷Buunk at al. 1993, p. 802.

³⁶⁸Antonucci and Jackson 1990.

³⁶⁹Rook 1987.

³⁷⁰Greenberg & Westcott 1983 and Hatfield & Sprecher 1983; as reported in Buunk at al. 1993, p. 801.

³⁷¹Sung Hee Kim and Smith 1993, pp. 39 & 41.

Still, it might be asked: Why should our self-worth be so sensitive to the opinions and attitudes of those around us? And, given that it is, why don't we react to injustice and disrespectful treatment with acquiescence or unconditional and instant forgiveness, instead of anger, indignation and resentment? Once again, social-psychological research is suggestive of certain answers to these questions. Disrespectful, unjust treatment upsets the ideal state of equality and fairness between individuals (or groups) and the balance is tipped in favour of the offender. A recognisable function of the retributive emotions, such as indignation, anger and resentment, is to promote a rapid re-establishment of that balance which is very important from the victimised person's point of view. For, unless that balance is re-established soon, there is an increased probability that the victim of the original offence will be further victimised by the offender.

Laboratory experiments indicate that if the harm-doer doesn't compensate his victim, either because proper channels are not open or because he chooses to withhold compensation, he will then distort his perceptions in such a way as to justify his actions. Usually one justifies the harm he has done by derogating his victim, but one may also justify his behavior in other ways. He may minimise the harm he has done or he may deny responsibility for the harm. It appears that the harm-doer will attempt to eliminate, at least in his own mind, the inequity that he has created, either by compensating his victim or by justifying his act.

Removal of inequity through justification rather than compensation is potentially dangerous. Not only does the harm-doer end up with a distorted and unreal assessment of his actions, but he may commit further acts based on these distortions. When the harm-doer's response to his act is justification, the victim is likely to be left in sad straits. Not only has he been hurt, but as a result of justification of the harmful act the probability that the harm-doer will hurt him again has increased.

Obviously, from the victim's point of view, it is desirable to have equity restored before the perpetrator is forced to justify what he has done.³⁷²

These findings may explain why in the face of injustice we react with anger, indignation and resentment. We are more likely to do something about restoring equity following unfair treatment as a result of having retributive feelings than we would be in

³⁷²Berscheid, Boye and Walster 1968, p. 370.

the absence of such feelings. Additional findings also explain why it is so tempting for victims to retaliate, sometimes even in the face of high personal cost. If offenders are unable or unwilling to redress their wrongs, reciprocating through retaliation appears to be the next best option which may be open to their victims if equity is to be restored at all to the now inequitable relationship. The benefits of retaliation are not merely that the victim's self-worth is restored, however. Retaliation also has the highly desirable effect of engendering respect for the victim in the offender. This is illustrated and supported by both anecdotal and experimental evidence from social psychology.³⁷³ Research findings clearly show that in cases where other ways of redressing the offence are impossible, "a harm-doer will derogate a victim who is powerless to retaliate but will not derogate a victim from whom he anticipates retaliation."³⁷⁴

Such results confirm that retaliation can be an effective way to restore equity following an offence. To be derogated by one's wrongdoer following an offence is quite literally to add insult to injury and is inimical in every way to victim rights and the principles of victim justice. Victims' desire for retaliation even in the face of further costs is, therefore, entirely understandable and the victim-focused perspective afforded by the above findings provides an interesting context for interpreting and making sense of Solomon's following remarks on vengeance, retribution and punishment.

Vengeance is the emotion of "getting even", putting the world back in balance, and this simple phrase already embodies a whole philosophy of justice, even if (as yet) unarticulated and unjustified. ... "Getting even" is and has always been one of the most basic metaphors of our moral vocabulary, and the frightening emotion of righteous, wrathful anger is an essential part of the emotional basis for our sense of justice, ... Our resentment of injustice is a necessary precondition of our passion for justice, and the urge to retribution its essential consequence. ... The impulse to punish is primarily an impulse to even the score.³⁷⁵

³⁷³For anecdotal evidence see Sung Hee Kim and Smith 1993, p. 40.

³⁷⁴Berscheid et. al. 1968, p. 370.

³⁷⁵Solomon 1990, pp. 293 & 292.

Depending in part on the gravity of the wrong committed, victims usually have a deep personal need to see and feel that retributive justice is done, that criminals are made to pay for what they have done, that they get their just deserts. If such needs are not satisfied in victims, punishments imposed on their wrongdoers, no matter how harsh, are never going to be good enough to give them satisfaction. Such feelings of dissatisfaction with the criminal justice system are often exploited by politicians who try to increase their popularity ratings at election time by promising tougher and tougher punishments for criminals. Apart from perpetuating the frightful, but at the same time utterly unhelpful, image of the criminal as a monster, such policies are proven failures in terms of crime control, let alone the social reintegration of offenders, and it is doubtful that they will succeed in the long run in engendering more respect for the law and the criminal justice process.³⁷⁶

Victim accounts suggest that a great deal of their dissatisfaction is due to their marginalisation and disempowerment in the legal processing and resolution of their cases. It is likely that sentences which currently leave them dissatisfied would in fact satisfy them if they were involved and had proper input into the final outcome of their cases. As Braithwaite observes,

The surprising thing is that victims, who so often call for more blood in traditional Western justice systems, in New Zealand frequently plead with the police to waive punishment and “give the kid another chance”.³⁷⁷

Thus, it is very important from the point of view of victim satisfaction and social-psychological restoration that victims are appropriately empowered and involved. Such

³⁷⁶However, Block and Lind argue that “for any punishment harsh enough so that no individual would commit these crimes if he were certain to be punished there also exists a probability of punishment less than unity that will deter all such criminal acts.” (Block and Lind 1975, p. 247.) If this is right, then increasing the chances of apprehension above a certain level would bring about more positive results in terms of deterrence.

³⁷⁷Braithwaite 1993, p. 39. Braithwaite seems to attribute this phenomenon to people being “enmeshed in institutions that invite them to care about each other instead of hate each other.” (p. 39.) However, it would be more appropriate to say that the two systems *allow* and *engender* the development of different attitudes. Under restorative justice processes victims and offenders do not have to be (and most of the time are not) explicitly invited to care about each. They naturally start to care as a result of new perspectives afforded to them by their direct involvement and empowerment.

victim involvement and empowerment should also include according to victims their rightful status to be the bearers of the prerogative and the power to prosecute, settle with, or forgive their aggressors. As I pointed out earlier, in a great number of cases of criminal wrongdoing the real or primary injury is to the victims, and society is harmed, if at all, only secondarily and derivatively by it. Hence, the affair is *their* business first and foremost, and it is the rest of society's only after that.

By analogous reasoning, the business of forgiveness of wrongs is primarily the domain of those against whom the wrongs in question have been committed, the victims, and their decisions with regard to prosecution or mercy should be vetoed by judges, magistrates or community representatives only when they are clearly reckless from the community's point of view. An interesting historical example of a system where this right was honoured is provided by medieval Iceland.

The function of the courts was to deliver verdicts on cases brought to them. That done, the court was finished. If the verdict went against the defendant, it was up to him to pay the assigned punishment - almost always a fine [paid to the victim]. If he did not, the plaintiff could go to court again and have the defendant declared an outlaw. The killer of an outlaw could not himself be prosecuted for the act; in addition, anyone who gave shelter to an outlaw could be prosecuted for doing so. Prosecution was up to the victim (or his survivors). If they and the offender agreed on a settlement, the matter was settled. ... If the case went to a court, the judgement, in case of conviction, would be a fine to be paid by the defendant to the plaintiff.³⁷⁸

Apart from acting as a boost of confidence and empowerment, giving victims the power of decision with regard to matters of prosecution and mercy has the potential to aid in generating respect in offenders for victims by means of victim empowerment. From a deterrence point of view alone it would make sense to make the prospect of personal victim retaliation a reality to offenders who at the moment don't have to face the often disastrous impact their actions have on victims, let alone face their victims' anger, indignation and resentment. Through substantial victim involvement and empowerment a legal justice system sends an unambiguous message to potential offenders that by law,

³⁷⁸Friedman 1979, p. 405 - 406.

and in virtue of the power of the state, their victims (and potential victims) possess the powers of punishment and forgiveness and that they will be given the opportunity to respond to wrongs suffered. This will not only help to restore victim confidence and dignity but will also help to promote a more realistic and healthier outlook towards others in potential and actual offenders. In light of research findings such as those quoted above, it is hard to think of a better way to teach offenders and impress on them the value and respect which must be accorded to others whom they victimise, intentionally or otherwise, through unlawful behaviour.

Greater victim involvement and empowerment could be achieved, for example, in the context of a privatised legal justice system. This could be modelled, for instance, on the medieval Icelandic system of law enforcement, along lines suggested by Friedman.

The first step in applying the Icelandic system of private enforcement to a modern society would be to convert all criminal offenses into civil offenses, making the offender liable to pay an appropriate fine to the victim. In some cases, it might not be obvious who the victim was, but that could be specified by legislation. ... For some minor offenses anyone could sue; presumably, whoever submitted his case first would be entitled to the fine. It must be remembered that specifying the victim has the practical function of giving someone an incentive to pursue the case. The second step would be to make the victim's claim marketable, so that he could sell it to someone willing to catch and convict the offender. The amount of the claim would correspond approximately to the damage caused by the crime divided by the probability of catching the criminal. In many cases it would be substantial. Once these steps were taken, a body of professional "thief-takers" (as they were once called in England) would presumably develop and gradually replace our present governmental police forces.³⁷⁹

Current debate of, and interest in, such models is focused on questions of comparative efficiency and cost-effectiveness, as a result of privatisation,³⁸⁰ while the desirability of such privatisation of the justice system from a victim justice point of view

³⁷⁹Friedman 1979, p. 414 - 15. (Appendix B). For more details, and arguments in favour, of a completely private system, see Friedman 1973 and Becker and Stigler 1974. For a discussion of an historical example of a private system from the nineteenth century American West, see Anderson and Hill 1979.

³⁸⁰See, for example, Becker and Stigler 1974; Landes and Posner 1975; Block and Lind 1975; Schwartz and Tullock 1975; Polinsky 1980; Friedman 1984.

receives little or no attention. Nevertheless, because of the proposed victim involvement, it is reasonable to expect that victim justice and empowerment would be prominent under such private law enforcement conditions. Privatisation, however, represents not only a radical departure from the dominant criminological theories of modern times but also from the experience of contemporary, urbanised societies with crime and the way it is handled. This explains, perhaps, why the debate on privatisation is sluggish at gathering pace. Notwithstanding, it is reasonable to expect that, in the long run, a complete privatisation of justice system is going to become reality. The price paid by the public for crime and law enforcement through our current institutions is prohibitively high. Considering that the very same money supports corruption, organised crime, and abuses of power which are prevalent in some of those institutions of enforcement (such as the police, for example), privatisation which effects a shifting of the financial burden away from the tax payer to the criminal is bound to look increasingly attractive not only to the average citizen and voter, and hence the political parties eager to capture their votes, but also to governments struggling to balance their budgets.

But, however that may be, there are at least two models for facilitating adequate victim involvement and empowerment which do not require privatisation of the institutions involved. They are perfectly workable in the context of existing justice systems. One of them, to be discussed in more detail further below, is a victim-offender restorative justice model. The idea here is that victim and offender, accompanied by their respective supporters (members of their respective families, close friends, teachers, employers, etc.) meet each other face to face to discuss and find an all-round satisfactory resolution to their conflict in a controlled and supportive environment. It is reasonable to expect, however, that in many cases of serious crime, such as murder, for instance, victims and offenders would not be prepared to meet face to face with each other. Therefore, this model is not going to be appropriate in all circumstances. Consequently, there is a need to consider the possibility of greater involvement in the existing judicial

process where victims could, and should, be empowered and given the opportunity to take an active role in the trial and the sentencing of their offenders.

The easiest way to do this, it seems, would be in the sentencing stage of the process where victims may be allowed to contribute to the final sentencing of their offenders. Indeed, an increasingly common attempt to make judicial systems more victim oriented is by way of Victim Impact Statements (VIS). A VIS is a statement by the victim about the physical, financial, psychological and social impacts of the crime on their lives. In some countries it may also include an expression of victim's feelings about the crime and the criminal, as well as a statement of opinion by the victim as to what they would consider to be an appropriate sentence. Normally, VIS are submitted to the judge for consideration in sentencing, thus providing victims with an opportunity to make a contribution towards the final resolution of their cases. While VIS are by no means the most empowering way to involve victims in the judicial process, when they are properly administered and are given due consideration by the sentencing judge or magistrate, they definitely are a welcome first step towards proper victim empowerment. Increased victim satisfaction with judicial outcomes as a result of participation in the process by way of VIS has already been demonstrated.³⁸¹ However, the effectiveness of VIS tends to be undermined by a couple of crucial factors which need to be covered briefly.

One of these factors is that it is mostly left to the discretion of the sentencing magistrate or judge to read and take into account the contents of the VIS. The remedy to this is obvious. Sentencing judges and magistrates must be required to address each element of the VIS as part of the sentencing process. No element of a VIS should be left without mention or dismissed without adequate justification. Proposed sentences should then be examined by the victim and by an appropriately qualified victim advocate. Sentences which do not display adequate awareness and consideration of the statement's contents should be automatically returned to the sentencing judge or magistrate in

³⁸¹Erez 1991, p. 5.

question for proper consideration. To be truly effective, these requirements must be given appropriate legislative backing.

The other factor which tends to undermine the effectiveness of this important victim right is that, all too often, victims either aren't adequately informed about VIS, or they do not receive the necessary encouragement and support with their preparation. Thus, VIS can become "well-kept secrets that only few victims know about, or make use of, to their advantage."³⁸² The remedy to this problem is fairly straightforward. Victims must be provided with an appropriately qualified advocate who, besides keeping them informed at all stages of the judicial processing of their cases and representing them and their interests in court, would consult with the victim for purposes of preparing a proper VIS according to the guidelines set out for the purpose. Once again, legislative backing would be required for such provisions to be truly effective. Legislative backing is important because, as Erez points out,

... victim participation and input into sentencing decisions challenge traditions and established patterns within the criminal courts, [and as a result] legislative reforms often amount to lip service; they typically lack remedies for non-compliance.³⁸³

A key role in ensuring the safeguarding of victims' rights in the above proposals would fall on the victim's representative, or advocate. Furthermore, on a par with there being a qualified advocate for the offender, it ought to be the victim advocate's role and responsibility to provide the victim with all the relevant information, support, and representation with respect to all aspects of the judicial process. This would require some serious changes to current legal and judicial proceedings where victims are unrepresented, are severely marginalised and disempowered. From a victim justice point of view, victims must be empowered to speak up and take an active role in judicial proceedings and they must be given proper support and representation by a qualified advocate. This would be the best way to guard and promote their interests throughout

³⁸²Erez 1991, p. 6. See also Kelly 1990.

³⁸³Erez 1991, p. 6.

the process, as well as in the determination of the final resolution of their cases. Even if only by means of VIS, their views on what is fair and appropriate punishment for their offenders must be heard and given proper consideration. Fundamental law reform in these areas is likely to aid significantly victim restoration by raising victim satisfaction levels with the judicial process, as well as levels of satisfaction with substantive justice. This is supported by research.

Studies of victims who participated as subsidiary prosecutors or who acted as private prosecutors (as several continental justice systems allow) reveal that their satisfaction with justice is higher than those who did not actively participate in the proceedings.³⁸⁴

The current lack of provisions for personal victim participation or, for that matter, a lack of indirect participation by way of a qualified victim advocate, in the criminal trial process is perhaps the single most important contributing factor to the continued marginalisation and disempowerment of victims in judicial systems. The severity of this disempowerment is acutely felt, especially under the adversary system, by victims of sexual assault.

Many sexual assault victims who have given evidence at a criminal trial feel that they were the ones being judged, rather than the assailant. ... women find that they are disbelieved, their allegations are trivialised, and their honesty and credibility are severely questioned.³⁸⁵

Judicial disempowerment is not restricted to sexual assault victims, however. The phenomenon is virtually universal and the problem is structural. The following passage is indicative of the disempowerment other victims experience with the judicial system.

How devastating it was for me and for my family to sit in the court, day after day, feeling almost as though Leanne was on trial. To hear various witnesses speak on behalf of her murderer, while we are not permitted to say one word on Leanne's behalf.³⁸⁶

³⁸⁴Erez 1991, p. 5.

³⁸⁵McCarthy 1994, p. 180.

³⁸⁶Harvey, R. How murder hits the victim's kin. *VOCAL Newsletter (ACT)*. 1994 - 95. Vol. 5. No. 3 - 4. p. 14.

I mentioned earlier that there is another way to promote effective victim restoration through empowerment besides the highly formalised judicial processing of offences. This is the victim-offender restorative justice model. The earlier mentioned *United Nations Resolution on Victims of Crime and Abuse of Power* strongly recommends in the interests of victims the use of such alternative, less formal processes of conflict resolution.³⁸⁷ A recent survey in South Australia found that, although victims of violent offences viewed the concept of mediation with more caution, “nearly six out of ten property offence victims were in favour of a mediation approach ... [because] they felt it would be of benefit to both the offender and themselves.”³⁸⁸ Under this kind of model, victim and offender, accompanied by their respective supporters meet each other face to face to discuss and find a solution to their conflict which is acceptable to both parties involved. The process is facilitated and controlled by trained and skilled facilitators while the content is determined by the participants involved from both sides of the dispute. The aim is to facilitate and allow an in depth discussion of the offence in question, consider its consequences, promote mutual understanding and appreciation for each others’ situation, and resolve the conflict in a way which is agreeable to both sides.

Various versions of this kind of conflict resolution between victims and their offenders have recently been adopted in several countries, including the USA, Canada, Singapore, Austria, Belgium, Germany, New Zealand, and Australia. Braithwaite’s reintegrative shaming model, for example, is well suited to accommodate the need for proper victim participation.³⁸⁹ However, a theoretically neglected virtue of such models is that the public shaming of offenders is one of the most constructive ways to meet the demands of retributive justice. There is no need to shy away from, or be embarrassed about, this aspect of shaming, for it is often a crucial element in promoting satisfaction and restoration in victims. It is a mistake to suppose, as advocates of restorative justice

³⁸⁷General Assembly 1985.

³⁸⁸Gardner 1990, pp. 46 & 58.

³⁸⁹Braithwaite 1989.

tend to do, that retributive justice and restorative justice are incompatible with each other.³⁹⁰ From a social-psychological perspective, retributive justice is often a necessary precondition to the restoration of social peace and to the satisfaction and restoration of victims.

I do not wish to deny, of course, that shaming is also a powerful weapon of social control and, if it is combined with appropriate efforts to reintegrate the offender into the community, visible shaming holds the promise of being an effective method of crime control. Such a diversionary conferencing method is currently subject to an extensive study, the *Re-Integrative Shaming Experiment (RISE) Project*, by The Australian National University with the cooperation of the Australian Federal Police whose officers act as facilitators. Although the study is still in its early stages, early indications are that in terms of victim, as well as offender, satisfaction this method of conflict resolution is far superior to traditional court processing of offences.³⁹¹ A similar model in Germany, for example, has already been “judged favourably and considered to have been a success” and in December 1994 “explicitly was made an integral part of general criminal law.”³⁹²

From a victim justice perspective, an important reason for empowering victims and, whenever possible, giving them the opportunity to discuss the offence personally

³⁹⁰Thus, for example, Consedine argues that the criminal justice system is only concerned with retribution and vengeance and he seems to regard these to be obstacles to restorative justice. (Consedine 1995) It is doubtful that Consedine is right in his diagnosis. Even if we put aside the reality of victims' needs for satisfaction in terms of retributive justice as a precondition of their restoration, many victims of crime would disagree with Consedine's claim about the orientation of contemporary justice systems. The earlier quoted comment that “the law does everything to protect the criminal and forgets about his victims” is not an atypical impression among victims of crime. Consedine's claim is similarly at odds with Doolan's following observation about discriminatory and harmful practices in pursuit of *rehabilitation* which, recognisably, is one of the major aims of criminal justice systems: “In New Zealand, Maori and Pacific Island youth are more fundamentally at risk of the more coercive, intrusive welfare dispositions ... in pursuit of rehabilitation, than are their Caucasian counterparts.” (Doolan 1993, p. 18.) Finally, deterrence is one of the most explicit aims of modern day justice systems and it is not unusual to hear arguments by lawyers that “the sentence should reflect a general deterrent.” (Leedham, N. Jail ‘unduly harsh’ for guilty officer. *The Canberra Times*, Thursday, June 22, 1995, p. 4.)

³⁹¹The author has been involved in the study as an observer and has interviewed both offenders and victims after court processing and conferencing of their cases.

³⁹²Severin 1995, p. 18.

with their aggressors is that irrational and unrealistic fears, apprehensions, anxieties and misconceptions about the aggressor's person and power are likely to be quickly dissipated in the victim's mind when their offender is shamed and humbled right there in front of them. Furthermore, such discussions also give an opportunity for an apology, and from a social-psychological point of view, apology is very important for speedy and effective victim restoration. It is the most desirable way to redress psychological trauma and soothe hard feelings in victims. The power of apology as a mechanism of aggression control has been demonstrated through laboratory experiments in social psychology.³⁹³ A sincere apology soothes hard feelings in victims, such as resentment and anger, and is very powerful in controlling and diffusing retaliatory aggression. Sincere apologies take the wind out of the sails of resentment and obviate the need for retaliation as a means of re-establishing equity into inequitable relationships.

The key messages conveyed in apology explain its almost magical ability to soothe hard feelings following an act of aggression. For example, by publicly admitting fault and wrongdoing through apology, the aggressor reduces the victim's sense of responsibility for the negative consequences of the act and, therefore, helps to restore the victim's self-respect and social identity. It is a common tendency among sexual assault victims, for example, to blame themselves for what happened. Such misplaced self-deprecation is bound to go hand in hand with a loss of self-esteem. Apology also indicates self-disapproval on account of one's wrongful act, and conveys the message that the wrongdoer is not such a bad person, after all. Again, apology expresses respect for the victim from the offender and publicly acknowledges the victim's moral and social status.

Acceptance of disgrace in public and expressions of remorse through apology may also be taken by victims and third parties as self-punishment for the transgression. This may reduce, if not altogether obviate, the need for imposing further punishment from

³⁹³See, for example, Sung Hee Kim & Smith 1993; Ohbuchi, Kameda & Agarie 1989.

outside. Especially where remorse and sincere apologies are accompanied by adequate restitution and compensation, the need for retaliation or further punishment may well be redundant, since adequate compensation also addresses the need for social justice and equity.³⁹⁴ In such circumstances retaliation or further punishment would rightly be perceived to be completely uncalled for, misguided and wrong.³⁹⁵

The benefits of apology, and the benefits of direct victim-offender discussions generally, do not stop with the victim, however. Resolving conflicts in these ways holds the promise of being far less costly to offenders than victim or community retaliation would be and it is also likely to be far less disruptive for everyone concerned. As indicated above, the potential of apology as a constructive mechanism for reducing and controlling retaliatory aggression is a demonstrated fact. Moreover, meeting their victims and offering them an apology is good for offenders because it affords them a constructive approach to the wrong they have done. Also, signs of forgiveness from their victims would make it easier for them to start being positive about themselves and, where applicable, to forgive themselves.³⁹⁶ Such factors are important for the social-psychological reintegration and restoration of offenders. With rare exceptions, offenders are not bereft of moral sensibilities and they realise that they have done wrong and that they need to make amends through adequate reparation and apology in order to put things right. In familiar jargon, offenders need to be given genuine opportunities to wipe the slate clean for a fresh start and impersonally imposed punishment by a judicial system which completely disempowers them is unlikely to engender the requisite feelings in them in this regard.

It is not difficult to see that denying offenders the opportunity to make amends in appropriate ways is bound to make it harder for them to feel remorse and to forgive

³⁹⁴Ohbuchi, et. el. 1989, pp 219 - 20.

³⁹⁵Possible concerns that greater victim involvement will lead to vindictiveness and to the endangerment of offender rights and offender justice will be covered further below. (Section 6.4.)

³⁹⁶Karen van den Broek pointed this out to me.

themselves. Worse, imposing on them penalties which may be wholly inappropriate to meeting their needs in the above indicated ways is as likely to hinder their reintegration and rehabilitation as it is likely to hinder victim restoration and healing. The need for engendering constructive attitudes in the offender and the need for appropriate sentences both highlight the importance of involving close family and friends from both sides of the conflict, since it is the opinions of close ones which matter most to them and it is these people who know best what is likely to be constructive rather than destructive in the given case. As Moana Jackson put it in connection with the Maori,

In a Maori setting, offenders were never alienated from the victim of their actions or the authority which decided their fate. Their actions were the shared responsibility of a whanau or iwi, and the consequences and judgement of them was similarly shared. ... Whanau involvement would both reinforce the idea of group responsibility and ensure the remorse and shame an offender needs to feel before effective rehabilitation and redress is possible.³⁹⁷

It is also an important point that the needs of offenders and victims do complement each other on this general level of principle in that their respective giving and acceptance of appropriate apologies and reparation comprise the single most important and constructive first step towards resolving their conflict to the point where they can both leave the matter behind and get on with their lives. It stands to reason, therefore, that the first mission of any justice system ought to be to promote the resolution of conflicts in ways which most appropriately meet the needs of both offenders and victims. If this is right, it hardly needs further argument that we should promote the empowerment of victims and offenders to the point where they can enter into civilised, constructive forms of dialogue with each other.

To summarise, I argued in this section for putting the personal element back into punishment by way of greater victim involvement in the justice system. I described two main models of constructive victim involvement and empowerment and argued that the principal reasons for such victim participation spring from considerations of victim rights

³⁹⁷Jackson 1988, pp. 234 & 242.

and justice to victims in the form of restitution, compensation and restoration in the physical, mental, financial and social domains. In the course of my arguments I drew support from findings in social-psychological research on the nature, psychology and dynamics of interpersonal relationships. While the arguments of this section are sufficient on their own to justify enhanced victim participation in the legal resolution of their cases, the defence of moral retributivism I have given in the preceding section also provides a morally admissible basis for substantial victim participation. When retributive considerations such as payback and just deserts are among a victim's reasons for getting involved in the punishment of their aggressors through opportunities provided for by institutionalised systems of justice, the result may well be institutionalised revenge in the sense described. Notwithstanding, victim participation of the kind I described above would not jeopardise any of the advantages of a good institutionalised system. This is what I argue in the section below.

6.4 Victim involvement and the dangers of revenge

Having unregulated, raw revenge in mind, rather than legal institutionalised forms of it of the kind I have been defending, Russell talks about revenge as “a very dangerous motive.”³⁹⁸ Hume similarly cautions against it with the words,

Who sees not that vengeance, from the force alone of passion, may be so eagerly pursued as to make us knowingly neglect every consideration of ease, interest, or safety?³⁹⁹

More recently, Wallace has emphasised that in unregulated revenge “there is always the danger of vendettas developing and spirals of violence” and that “when personal relations have deteriorated so much that violent revenge is on the agenda proportionality is unlikely.”⁴⁰⁰ Wallace does recognise, however, that when there is no system of law, or when the law is “discriminatory by failing to offer equal protection to all citizens,” or

³⁹⁸Russell 1961, p. 561.

³⁹⁹Hume 1975, Appendix II, p. 302.

⁴⁰⁰Wallace 1995, p. 374.

when some class of harms are “legally ignored no matter who suffers them, ... revenge, with all its drawbacks, is then the only form of retribution available.”⁴⁰¹ But because of the dangers involved in unregulated revenge, Wallace argues that

injustice is the probable outcome. So even if it is sometimes possible to justify harmful revenge, a general right to be revenged cannot be sustained.⁴⁰²

I agree that any form of retaliation or punishment which is outside a formalised and fair institutionalised framework, such as unregulated, raw revenge and vigilante punishment, for example, can have many dangers and ills. The wrongs and the losses being very close to home, the need to “even the scores” in revenge is likely to be driven by powerful retributive emotions, such as resentment, indignation and anger. In the heat of the moment, judgement and control can be easily affected and, therefore, more can be inflicted in retaliation, for example, than what, in all fairness, would be due. Such factors are bound to increase the dangers of escalation also. Therefore, even though there is nothing *inherently* wrong with non-institutionalised acts of revenge which observe the requisite principles of justice, the need for structural controls by means of institutionalisation and formalisation of the processes which deal with offences requires no further argument. What remains at issue is the question whether the same considerations also give good grounds for making the processes of the justice system non-personal by excluding victims from the processing and resolution of their cases. In this section I argue that there are no justifiable reasons for such victim exclusions. There are several reasons why this is so and I shall cover them briefly.

To start with, the claims and requirements of victim rights and victim justice outlined in the previous section alone are sufficient to cast serious doubt on the defensibility of the current marginalisation and disempowerment victims experience with the judicial processing of their cases. It is increasingly recognised that more attention must be paid to the rights of the wronged individuals in question and those closely related

⁴⁰¹Wallace 1995, p. 374.

⁴⁰²Wallace 1995, p. 375. Similar views expressed in Primoratz 1989, p. 84.

to them. One of the latest developments in this regard is the recently announced law reforms in New South Wales (Australia) to establish a statutory charter of victims' rights. The new legislation will also give victims of crime "the legal right to influence a court's sentencing of people convicted of crime ... [and also] a say in parole and temporary leave for serious offenders."⁴⁰³ To reiterate an earlier point, given that it is the victims who in fact have been wronged and that they are the ones most affected by the criminal activity in question, justice must be primarily *theirs*, not society's; and it must be at least as much theirs as it is the criminal's. As Tom Campbell suggested,

those who object to victim participation on the grounds that it distracts attention from doing justice to the accused, may simply be begging the crucial questions against victim justice.⁴⁰⁴

Furthermore, there are no reasons to think that victim involvement would jeopardise any of the principles of offender justice, if it all took place within a fair, properly structured institutionalised framework. For, notwithstanding victim involvement, the usual safeguards of a well-constructed system in guarding against violating principles of justice with respect to offenders would continue to be present. In other words, there are no reasons to think that victim justice and justice to offenders are incompatible with each other under either of the two institutionally based models I have described in the previous section. Moreover, the fear of vengeful victims being unreasonable and unduly harsh on their offenders has no rational basis. Apart from the fact that the power of victims is still very limited under both models discussed and, therefore, there is virtually no chance that they can abuse their enhanced status to bring about unfair outcomes for their offenders, research into attitudes to sentencing consistently shows that "victims of crime are no more punitive than others."⁴⁰⁵

⁴⁰³Victims of crime get rights in sentencing. *The Canberra Times*, Thursday, October 5, 1995, p. 5.

⁴⁰⁴Campbell, 1984, p. 359.

⁴⁰⁵Hough and Moxon 1988, p. 147.

The message from studies in several countries is that first-hand experience of crime as a victim does not, in general, fuel a desire for heavy sentences. Furthermore, victims have not been found to be more punitive than the general public.⁴⁰⁶

Lastly, under the two models I have covered, the danger of escalation would also be contained. Because of the institutionalised nature of the system there would be an in-built finality which ends the process. Thus, the danger of conflict escalation which is ever-present in raw revenge, for example, is no longer a problem and, therefore, it cannot serve as an objection.

6.5 Conclusion

In this chapter I defended the moral justifiability of institutionalised revenge and argued for the importance of victim restoration by means of engendering a feeling of satisfaction in victims regarding the legal processing and resolution of their cases. I argued that such processes and resolutions must take into account not only victims' needs for adequate restitution and compensation, but also their legitimate personal need for an apology and/or retributive justice. This is especially the case, I suggested, in cases of serious crime, such as murder.

As part of my strategy to support my position in favour of victim restoration through greater involvement and empowerment in the punishment process, I first provided an account of moral retributivism - which I defined as the thesis that offenders may be punished within limits set by the principles of justice for retributive reasons alone, quite independently of consequentialist or functionalist considerations. I then went on to provide a justification of this thesis by an appeal to the notions of human flourishing, life within a moral community and personal accountability. I also showed how the central claim of this argument, namely that retributive responsibility for one's

⁴⁰⁶Erez 1991, p. 4. See also Gardner 1990; Walker & Hough 1988; Kelly 1990. Shapland, Willmore & Duff 1985 (especially pp. 72 - 74).

actions is a necessary part of one's full humanity, finds support in the philosophy of Kant, Hegel and Lewis.

I then considered directly the question of greater victim involvement and empowerment in the processes of the justice system. I argued that the interests of victim rights and victim justice in terms of material, psychological and social restoration would be enhanced significantly as a result of such involvement. I argued that, especially from a social-psychological point of view, empowering victims and according them a role which is commensurate in significance to their status as the persons who in fact have been wronged and whose lives have been most affected and disrupted by the offence in question, is a necessary first step towards enabling them to put the past behind them and return, in so far as it is possible, to the quiet enjoyment of their lives. Finally, in the last section I suggested that there are no defensible reasons for the continued marginalisation and disempowerment which victims currently experience under entrenched judicial processes and practices. In particular I argued that, regardless of whether retributive considerations played a role or not, victim participation under the two models I described would not jeopardise any of the advantages of a good institutionalised system. Important goals, such as ensuring justice to offenders and the prevention of escalation would not be endangered as a result of such victim participation.

Summary and Conclusions

In the Introduction to this thesis I posed the question of whether revenge remains a relevant consideration to justice in situations where the quest for justice is confined to the exclusive domain of legal institutionalised systems. I suggested that the concept of revenge requires a proper analysis as a precondition for an informed evaluation of its normative status and claimed that, in spite of a strong anti-revenge paradigm in industrialised Western cultures, revenge remains highly relevant to contemporary concerns of justice. Through subsequent analysis and argumentation I sought to lend support to this claim and lay the foundations for a new paradigm of justice which gives as much importance to doing justice to victims as it does to doing justice to offenders. I argued that the requirements of justice can be satisfied from the point of view of all parties involved provided that we can avoid lop-sided paradigms in which attention is focused too much on either the victim or the offender.

Seeking to neutralise the gravitational pull of a deeply entrenched anti-revenge paradigm, Chapter 1 was devoted to examining critically the theoretical and philosophical adequacy of commonly held beliefs and dogmas about revenge. I sought to establish in the first main section that the charges levelled against revenge by Lewis, that is, that revenge loses sight of the end in the means, that it is a perversion, that it is a nasty vindictive passion and an evil, are unjustified. Following this, in the next main section I challenged suggestions by Nozick and Glover that revenge is characterised by a morally repugnant emotional tone: pleasure in the suffering of another. Apart from questioning whether there would be anything objectionable about taking pleasure in the deserved suffering of wrongdoers, I argued that there are more plausible ways of identifying the emotional tones which may typically accompany revenge and that the notion of

satisfaction would especially be appropriate and plausible in this regard. In the last main section of the first chapter I explored the question of whether revenge can be properly regarded as a form of punishment and argued in support of an affirmative answer. I argued that revenge has all the morally relevant features to qualify as punishment since it is a punitive response to injustices and wrongs perceived to have been committed against us and those close to us. I also argued that the dominant emotions which typically accompany revenge are the retributive emotions of anger, indignation, but especially resentment, which quite properly arise in us in response to injustices and wrongs.

In the second chapter I clarified the notion of retribution and considered its relationship to the principles of justice and revenge. First I considered the concept of retribution and argued that the best way to define it is by reference to the notion of negative repayment. I then showed that principles of justice which set the limits of just punishment are not retributive in character. In the final main section of the second chapter I focused on the relationship between retribution and revenge. I considered the two rival views on this issue, namely the view that revenge and retribution are identical and the view that the two fall into quite different moral categories. I argued that both views are unsatisfactory but noted that the two concepts are significantly related, evidenced by their frequent, albeit mistaken, identification.

A solution which could accommodate all the above considerations was offered in the third chapter where I argued that revenge is a form of retribution of a personal sort. Thus, while revenge is not identical with retribution, their closeness consists in a one-way conceptual link. As a way to prepare the ground for such an analysis, I began the third chapter by discussing the modalities and the process of punishment, pointing out the moral issues in both areas. I then went on to provide a typology of the reasons for punishment. I showed that, whatever the reasons for punishment, the process of punishment can occur in institutionalised as well as non-institutionalised settings and that punishment can be administered with personal or non-personal involvement of the wronged persons in question. I then argued in the fifth section of that chapter that, in

relation to reasons for punishment, revenge is retributive and that, with respect to the forms of punishment, revenge is personal. These arguments furnished the basis for my claims that revenge is personal retributive punishment and that it can take institutionalised, as well as non-institutionalised forms. I sought to elucidate my arguments and definition by locating revenge on a visual representation of the logical space of punishment. I defined revenge as personal retributive punishment which is typically accompanied by feelings of indignation, anger and resentment for wrongs suffered in one's personal domain of care and concern. I then went on to argue that the retributive element in revenge remains compatible with instrumental considerations of one's interests. I drew attention to the fact that in certain cases, such as the feud, instrumental considerations of honour and social standing can become the primary considerations in the direct reasoning leading to a decision to take revenge. However, I also pointed out that this phenomenon of strong instrumental focus does not disconfirm my analysis of revenge in terms of retribution because the rationale of revenge is retribution even in such cases. The retributive idea of negative reciprocity, or repayment (payback) is an essential element in any satisfactory analysis of the feud because the most basic point of justification for retaliation is not the need to maintain honour and social standing but the retrospective responsibility the targeted person bears for the offence in question and, correlatively, a deontologically based moral duty of the offended party to materialise such retrospective responsibility through appropriate retaliation. Honour acts merely as an additional stake and serves to reinforce the importance of such responsibilities and duties.

Following this account of revenge and clarification of its relationship to other important elements of punishment, I went on to consider its normative defensibility from rational and moral perspectives. In Chapter 4 I discussed the notion of rationality and why it is important that revenge be rational. Firstly, I showed that rationality is important because, in general, rational behaviour has the best chance of enabling us to pursue and achieve our short-term and long-term goals. I then discussed rational choice theory and

argued that if it is to have normative strength, the purely ends-means conception must be supplemented by appropriate evaluations of the agent's goals and beliefs. Following this, I distinguished between direct and indirect rationality. Behaviour is directly rational if, as a direct result of a rational choice calculation, it is the best means of achieving the agent's goal. Behaviour is indirectly rational if it forms part of a rational habit, disposition, attitude, custom or social norm. I further argued that it is not significant from a normative point of view whether one's behaviour is directly or indirectly rational. What matters is not that we constantly engage in cost-benefit analyses, but that we do not jeopardise our short- and long-term interests. Having rational habits, dispositions, attitudes, customs and norms is enough to ensure that this does not happen. I also made the suggestion that problems encountered by rational choice theory in situations, such as the prisoner's dilemma, can be satisfactorily overcome only by a theory of indirect rationality.

I then went on to consider the collective rationality of revenge. With reference to the collective rationality of particular group actions, I argued that, given the importance of *mana* for the traditional New Zealand Maori, and given the general conditions prevailing in pre-European New Zealand, there is good evidence that in a great many situations Maori group revenge was perfectly rational in the required sense of the word. I argued that traditional Maori revenge behaviour satisfied the requirements of both direct and indirect rationality, as well as rationality's formal and substantive requirements. I suggested that this was so regardless of whether the range of proper goals was restricted to strictly material interests in the way preferred by Elster or not. I then considered the collective rationality of revenge practices as social institutions. I argued that social institutions such as the feud served the same kind of functions in societies with no centralised state as legal justice systems do today and suggested that victim-based institutions of justice amounting to practices of revenge were not obviously inferior choices, given the combined aim of promoting deterrence, honour and retributive justice at economically competitive cost.

In Chapter 5 I focused on the arguments put forward by Elster against the individual rationality of revenge behaviour. I gave a detailed presentation of Elster's overall position and identified four problem areas for his arguments. First I argued against his claim that revenge is backward-looking and that, therefore, it cannot be rational. Following this, I disputed his claims that honour is universally present in revenge and that it is the key to understanding revenge. I argued that when honour is an essential part of people's sense of identity in the epistemic sense, it is a legitimate and defensible value and goal for them to have from the point of view of a substantive, normative rationality. Further, I pointed out that, when combined, the remaining two claims are suggestive of a dilemma: Either revenge is rational because it is in the material interests of the agent, or it is rational because it is in the non-material interests of the agent. According to Elster, the former is not the case, since revenge can never be in the material interests of the agent, while the latter case, if true, is trivial and uninformative. I argued against both these claims and showed that revenge can be in both the material and non-material interests of the agent. With respect to material interests, in certain cultural contexts there is a real threat of severe social sanctioning when the individual fails to carry out the duty of revenge. In such instances revenge is individually rational in the direct sense. I also argued that revenge can be rational in the indirect sense because a revengeful attitude or disposition is in the long-term material interests of the agent. This is so because, as part of a resilient TIT FOR TAT pattern of behaviour, it is the best way of eliciting cooperation from others, and thus, a better way of doing well for oneself than otherwise. I also showed that revenge can maximise agents' non-material interests, such as the preservation of their honour, and that this is not a trivial or uninformative claim.

Having shown that revenge can be rational and that it often is, I moved on to consider in the final chapter its defensibility and contemporary relevance to institutions of justice from a moral point of view. More specifically, I defended the moral justifiability of institutionalised revenge and argued for the importance of victim restoration through the engendering of a feeling of satisfaction in victims regarding the legal processing and

resolution of their cases. I argued that such processes and resolutions must take into account not only their need for adequate restitution and compensation, but also their legitimate personal need for an apology and/or retributive justice. This is especially the case, I suggested, in cases of serious crime such as murder. As part of my strategy to support my position for victim restoration through their greater involvement and empowerment in the legal justice system, I first provided an account of moral retributivism - which I defined as the thesis that offenders may be punished within limits set by the principles of justice for retributive reasons alone, quite independently of consequentialist or functionalist considerations. I then went on to provide a justification of this thesis by an appeal to the notions of human flourishing, life within a moral community and personal accountability. I also showed how the central claim of this argument, namely that retributive responsibility for one's actions is a necessary part of one's full humanity, finds support in the philosophy of Kant, Hegel and Lewis.

I then considered directly the question of greater victim involvement and empowerment in the processes of the justice system. I argued that the interests of victim rights and victim justice in terms of material, psychological and social restoration would be enhanced significantly as a result of such involvement. I argued that, especially from a social-psychological point of view, empowering victims and according them a role which is commensurate in significance to their status as the persons who in fact have been wronged and whose lives have been most affected and disrupted by the offence in question, is a necessary first step towards enabling them to put the past behind them and return, in so far as it is possible, to the quiet enjoyment of their lives. I argued that doing justice to victims requires a legal justice system which, instead of silencing, marginalising and disempowering victims, acknowledges and respects their legitimate need for retributive justice, empowers them to speak up for themselves and encourages them to participate in the legal resolution of the offences committed against them. Giving such considerations to victims is necessary for their social-psychological restoration and is in part what it means to treat them with justice. Finally, in the last section of the

chapter I suggested that there are no defensible reasons for the continued marginalisation and disempowerment which victims currently experience under entrenched judicial processes and practices. In particular I argued that, regardless of whether retributive considerations played a role or not, victim participation under the two models I described would not jeopardise any of the advantages of a good institutionalised system. Important goals, such as ensuring justice to offenders and the prevention of escalation, would not be endangered as a result of such victim participation and empowerment. What must be especially stressed in these regards is that, apart from the fact that victims of crime are generally no more punitive than others, the legal institutionalised nature of the models I have considered can ensure that victims do not abuse their enhanced status to bring about unfair outcomes for their offenders and that there is an in-built finality which ends the process for purposes of preventing escalation of the conflict.

If this is right, and if so too are my arguments for victim justice in terms of victim restoration, personal involvement, empowerment and satisfaction with the legal processing and final resolution of cases, - which, as I have argued, must also allow for the legitimate claims of retributive justice - there is a very strong case for abandoning Bacon's advice while maintaining a legal institutional perspective on justice, and for substituting his advice with the following: *Instead of trying to eradicate revenge, the law should only seek to tame and weed the wildness out of it.* For, to repeat the theme of the preface, justice is too important to be left solely to the judiciary.

Bibliography

- Abel, C. F. and Marsh, F. H. 1984. *Punishment and Restitution: A Restitutionary Approach to Crime and the Criminal*. Greenwood Press. Westport.
- Acton, H. B. (Ed.) 1969. *The Philosophy of Punishment*. Macmillan. London.
- Adams, R. M. 1985. Involuntary Sins. *The Philosophical Review*, 94: 3 - 31.
- Agassi, J. 1987. Theories of Rationality. In Agassi and Jarvie (Eds.) 1987.
- Agassi, J. and Jarvie, I. C. (Eds.) 1987. *Rationality: The Critical View*. Martinus Nijhoff. Dordrecht.
- Allen, C. K. 1958. *Aspects of Justice*. Stevens & Sons. London.
- Anderson, T. L. and Hill, P. J. 1979. An American Experiment in Anarcho-Capitalism: The Not So Wild, Wild West. *The Journal of Libertarian Studies*, 3: 9 - 29.
- Antonucci, T. C. and Jackson, J. S. 1990. The Role of Reciprocity in Social Support. In Sarason and Pierce (Eds.) 1990.
- Armstrong, K. G. 1968. The Retributivist Hits Back. In Acton (Ed.) 1969.
- Arrow, K. J. 1977. Current Developments in the Theory of Social Choice. *Social Research*, 44: 607 - 622. Also reprinted in Barry and Hardin (Eds.) 1982.
- Atkinson, M. 1974. Interpreting Retributive Claims. *Ethics*, 85: 80 - 86.
- Axelrod, R. 1984. *The Evolution of Cooperation*. Basic Books. New York.
- Ayers, E. L. 1984. *Vengeance and Justice: Crime and Punishment in the 19th-Century South*. Oxford University Press. Oxford.
- Bacon, F. 1985. *The Essays*. (J. Pitcher, Ed.) Penguin Books. England.
- Bar-Elli, G. and Heyd, D. 1986. Can Revenge be Just or Otherwise Justified? *Theoria*, 52: 68 - 86.

- Barry, B. and Hardin, R. (Eds.) 1982. *Rational Man and Irrational Society?* Sage Publications. Beverly Hills.
- Becker, S. and Stigler, G. J. 1974. Law Enforcement, Malfeasance, and Compensation of Enforcers. *The Journal of Legal Studies*, 3: 1 - 18.
- Benedict, R. 1977. *The Chrysanthemum and the Sword*. Routledge and Kegan Paul. London.
- Berger, P. 1984. On the Obsolescence of the Concept of Honour. In Sandel (Ed.) 1984.
- Berkowitz, L. 1962. *Aggression: A Social Psychological Analysis*. McGraw-Hill. New York.
- Berofsky, B. 1987. *Freedom from Necessity*. Routledge & Kegan Paul. New York.
- Berscheid, E., Boye, D. and Walster, E. 1968. Retaliation as a Means of Restoring Equity. *Journal of Personality and Social Psychology*, 10: 370 - 76.
- Best, B. 1952. *The Maori As He Was*. R. E. Owen. Government Printer. Wellington.
- Bird, O. A. 1967. *The Idea of Justice*. F. A. Praeger. New York.
- Bishop, J. 1985. Review of Elster 1983. *Australasian Journal of Philosophy*, 63: 245 - 48.
- Bishop, J. 1986. Review of Schick 1984. *Australasian Journal of Philosophy*, 64: 238 - 39.
- Black, M. 1982. Why Should I Be Rational? *Dialectica*, 36: 147 - 168.
- Black-Michaud, J. 1975. *Cohesive Force: Feud in the Mediterranean and Middle East*. Blackwell. Oxford.
- Block, M. K. and Lind, R. C. 1975. Crime and Punishment Reconsidered. *The Journal of Legal Studies*, 4: 241 - 247.
- Boehm, C. 1984. *Blood Revenge*. University Press of Kansas. Lawrence.
- Bourdieu, P. 1977. *Outline of a Theory of Practice*. (R. Nice, Trans.) Cambridge University Press. Cambridge.
- Bradley, F. H. 1927. The Vulgar Notion of Responsibility. In *Ethical Studies*. (2nd ed.) Clarendon Press. Oxford.

- Braithwaite, J. 1979. *Inequality, Crime and Public Policy*. Routledge and Kegan Paul. London.
- Braithwaite, J. 1989. *Crime, Shame and Reintegration*. Cambridge University Press. Cambridge.
- Braithwaite, J. 1993. What is to be Done about Criminal Justice? In Brown and McElrea (Eds.) 1993.
- Braithwaite, J. 1995. Corporate Crime and Republican Criminological Praxis. In F. Pearce & L. Snider (Eds.) *Corporate Crime: Ethics, Law and State*. University of Toronto Press. Toronto.
- Braithwaite, J. and Pettit, P. 1990. *Not Just Deserts*. Clarendon Press. Oxford.
- Brown, B. J. and McElrea, F. W. M. (Eds.) 1993. *The Youth Court in New Zealand: A New Model of Justice*. Legal Research Foundation Publication No. 34. Auckland.
- Buss, A. H. 1961. *The Psychology of Aggression*. John Wiley. New York.
- Butler, J. 1896. Upon Resentment (Sermon VIII); Upon Forgiveness of Injuries (Sermon IX). In *The Works*. (Vol. 2. W. E. Gladstone, Ed.) Clarendon Press. Oxford.
- Buunk, P. B., Doosje, B. J., Jans, L. G. J. & Hopstaken, L. E. M. 1993. Perceived Reciprocity, Social Support, and Stress at Work: The Role of Exchange and Communal Orientation. *Journal of Personality and Social Psychology*, 65: 801 - 811.
- Calhoun, C. 1992. Changing One's Heart. *Ethics*, 103: 76 - 96.
- Campbell, T. 1984. Compensation as Punishment. *UNSW Law Journal*, 7: 338 - 61.
- Campbell, T. 1988. *Justice*. Macmillan. London.
- Carthwright, D. 1985. Revenge, Punishment and Mercy: The Self-Overcoming of Justice. *International Studies in Philosophy*, 17: 17 -26.
- Carthwright, D. E. 1984. Kant, Schopenhauer, and Nietzsche on the Morality of Pity. *Journal of the History of Ideas*, 45: 83 -98.
- Clark, M. S. and Mills, J. 1979. Interpersonal Attraction in Exchange and Communal Relationships. *Journal of Personality and Social Psychology*, 37: 12 - 24.
- Consedine, J. (Ed.) 1995. *Restorative Justice: Healing the Effects of Crime*. Ploughshares Publications. Lyttleton, NZ.

- Copeland, J. (Ed.) 1996. *Logic and Reality: Essays in Honour of Arthur Prior*. Oxford University Press. Oxford.
- Cottingham, J. 1979. Varieties of Retribution. *The Philosophical Quarterly*, 29: 238 - 46.
- Davidson, D. 1980. *Essays on Actions and Events*. Clarendon Press. Oxford.
- de Sousa, R. 1987. *The Rationality of Emotion*. The MIT Press. Cambridge, Mass.
- Del Mar, F. 1924. *A Year Among The Maoris*. Ernest Benn. London.
- Dickie, P. 1989. *The Road to Fitzgerald and Beyond*. University of Queensland Press. Brisbane.
- Dimsdale, T. J. 1866. *Vigilantes of Montana*. Montana Post Press. Virginia City.
- Doolan, M. P. 1993. Youth Justice - Legislation and Practice. In Brown and McElrea (Eds.) 1993.
- Drapkin, I. and Viano, E. (Eds.) 1975. *Victimology: A New Focus. (Vol. 3) Crimes, Victims and Justice*. Lexington Books. Massachusetts.
- Ellis, A. 1995. Recent work on punishment. *The Philosophical Quarterly*, 45: 225 - 333.
- Elster, J. (Ed.) 1986. *Rational Choice*. Basil Blackwell. Oxford.
- Elster, J. 1983. *Sour Grapes*. Cambridge University Press. Cambridge.
- Elster, J. 1990. Norms of Revenge. *Ethics*, 100: 862 - 85.
- Erez, E. 1991. Victim Impact Statements (Issue paper No. 33). In Grabosky, P. (Ed.) *Trends and issues in crime and criminal justice*. Australian Institute of Criminology. Canberra.
- Evans, D. A. 1977. *Retributive Theories of Punishment Justification*. Thesis. Wayne State University. Detroit.
- Ezzat, A. F. (Ed.) 1986. *From Crime Policy to Victim Policy: Re-orienting the Justice System*. Macmillan. Basingstoke.
- Ezzat, A. F. (Ed.) 1989. *The Plight of Crime Victims in Modern Society*. St. Martin's Press. New York.
- Fawcett, S. B., Seekins, T., Whang, P. L., Muiu, C. and de Balcazar, Y. S. 1984. Creating and Using Social Technologies for Community Empowerment. In Rappaport *et al.* 1984.

- Feibleman, J. K. 1985. *Justice, Law and Culture*. Martinus Nijhoff Publishers. Dordrecht.
- Feinberg, J. 1970. *Doing and Deserving*. Princeton University Press. New Jersey.
- Foote, P. 1970. *The Viking Achievement*. Sidgwick & Jackson. London.
- Foucault, M. 1977. *Discipline and Punish*. Pantheon Books. New York.
- Foucault, M. 1980. On Popular Justice: A Discussion with Maoists. In Gordon C. (Ed.) *Power/Knowledge, Selected Interviews and Other Writings 1972 - 1977*. Pantheon Books. New York.
- Frankfurt, H. G. 1969. Alternate Possibilities and Moral Responsibility. *Journal of Philosophy*, 66: 829 - 39.
- French, P. A. 1984. *Collective and Corporate Responsibility*. Columbia University Press. New York.
- Friedman, D. 1973. *The Machinery of Freedom: Guide to a Radical Capitalism*. Harper & Row. New York.
- Friedman, D. 1979. Private Creation and Enforcement of Law: A Historical Case. *The Journal of Legal Studies*, 8: 399 - 415.
- Friedman, D. 1984. Efficient Institutions for the Private Enforcement of Law. *The Journal of Legal Studies*, 13: 379 - 397.
- Garcia, J. L. A. 1986. Two Concepts of Desert. *Law and Philosophy*, 5: 219 - 235.
- Garcia, J. L. A. 1989. Deserved Punishment. *Law and Philosophy*, 8: 263 - 277.
- Gardner, J. 1990. *Victims and Criminal Justice*. Office of Crime Statistics. South Australian Attorney-General's Department. Adelaide.
- Gaus, F. G. 1990. *Value and Justification*. Cambridge University Press. Cambridge.
- General Assembly. 1985. *Resolution Adopted By The General Assembly. 40/34. Declaration Of Basic Principles Of Justice For Victims Of Crime And Abuse Of Power Annex. Declaration Of Basic Principles Of Justice For Victims Of Crime And Abuse Of Power*. Document No: A/RES/40/34. 11 December 1985. United Nations.
- Gensler, H. J. 1986. Ethics is Based on Rationality. *The Journal of Value Inquiry*, 20: 251 - 264.
- Glover, J. 1970. *Responsibility*. Routledge & Kegan Paul. New York Humanities Press. London.

- Golding, M. P. 1975. *Philosophy of Law*. Prentice-Hall. Englewood Cliffs. New Jersey.
- Gragas. 1980. *Laws of Early Iceland: Gragas*. (A. Dennis, P. Foote and R. Perkins, Trans.) University of Manitoba Press. Winnipeg.
- Greenspan, P. S. 1988. *Emotions and Reasons*. Routledge. New York.
- Griffin, J. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Clarendon Press. Oxford.
- Hamlin, A. P. 1991. Rational Revenge. *Ethics*, 101: 374 - 81.
- Hampton, J. and Murphy, J. G. 1988. *Forgiveness and Mercy*. Cambridge University Press. Cambridge.
- Hardin, R. 1991. Trusting Persons, Trusting Institutions. In R. Zeckhauser (Ed.) *Strategy and Choice*. The MIT Press. Cambridge, Mass.
- Harding, C. and Ireland, R. W. 1989. *Punishment: Rhetoric, Rule and Practice*. Routledge. London.
- Hare, R. M. 1981. *Moral Thinking: Its Levels, Method, and Point*. Clarendon Press. Oxford.
- Harsanyi, J. C. 1977. Morality and the Theory of Rational Behavior. *Social Research*, 44: 623 - 656.
- Harsanyi, J. C. 1986. Advances in Understanding Rational Behavior. In Elster (Ed.) 1986: 82 - 107.
- Hart, H. L. A. 1968. *Punishment and Responsibility*. Clarendon Press. Oxford.
- Hasluck, M. 1954. *The Unwritten Law in Albania*. Cambridge University Press. Cambridge.
- Hedenius, I. 1973. The Concept of Punishment. In *Modality, Morality and Other Problems of Sense and Nonsense: Essays Dedicated to Soren Hallden*. Gleerup. Lund.
- Hegel, G. W. F. 1942. *The Philosophy of Right*. (T. M. Knox, Trans.) Oxford University Press. Oxford.
- Heider, F. 1958. *The Psychology of Interpersonal Relations*. John Wiley & Sons. London.
- Heil, J. 1992. Believing Reasonably. *Nous*, 26: 47 - 61.

- Heller, A. 1987. *Beyond Justice*. Basil Blackwell. Oxford.
- Hobbes, T. 1962. *Leviathan*. J. Plamenatz (Ed.) Collins. London.
- Holmstrom, L. L. and Burgess, A. W. 1975. Rape: The Victim Goes on Trial. In Drapkin and Viano (Eds.) 1975.
- Honderich, T. 1984. *Punishment: The Supposed Justifications*. Penguin Books. Harmondsworth.
- Hough and Moxon, 1988. Dealing with Offenders: Popular Opinion and the Views of Victims in England and Wales. In Walker & Hough 1988.
- Houston, J. 1965. *Maori Life in Old Taranaki*. A. H. & A. W. Reed. Wellington.
- Hume, D. 1975. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. (3rd ed.) L. A. Selby-Bigge (Ed.) Oxford University Press. London.
- Husak, D. N. 1994. Is Drunk Driving a Serious Offence? *Philosophy and Public Affairs*, 23: 52 -73.
- Jackson, F. 1987. Group Morality. In Pettit, Sylvan and Norman 1987: 91 - 110.
- Jackson, M. 1988. *The Maori and the Criminal Justice System, a New Perspective: He Whaipaanga Hou*. Department of Justice. Wellington.
- Jackson, M. 1992. The Colonisation of Maori Philosophy. In Oddie and Perrett (Eds.) 1992: 1 - 10.
- Jacoby, S. 1983. *Wild Justice: The Evolution of Revenge*. Harper & Row Publishers. New York.
- Johansen, L. 1987. The Theory of Public Goods: Misplaced Emphasis? In F. R. Forsund (Ed.) *Collected Works of Leif Johansen*. (Vol. 2) Elsevier Science. Amsterdam.
- Johnson, C. D. 1975. Moral and Legal Obligation. *The Journal of Philosophy*, 72: 315 - 333.
- Kafka, F. 1961. In the Penal Settlement. In *Metamorphosis and Other Stories*. Penguin Books. Middlesex. England.
- Kahn, A. S. (Ed.) 1984. *Social Psychology*. W. C. Brown Publishers. Dubuque, Iowa.
- Kamler, H. 1985. Strong Feelings. *The Journal of Value Inquiry*, 19: 3 - 12.

- Kant, I. 1965. *The Metaphysical Elements of Justice*. (J. Ladd, Trans.) Bobbs-Merrill. Indianapolis.
- Kant, I. 1974. *The Philosophy of Law*. Augustus M. Kelley Publishers. Clifton. New Jersey.
- Kaufmann, W. A. 1950. *Nietzsche: Philosopher, Psychologist, Antichrist*. Princeton University Press. New Jersey.
- Kaufmann, W. A. 1959. *The Owl and the Nightingale*. Faber and Faber. London.
- Kelly, D. P. 1990. Victim Participation in the Criminal Justice System. In A. J. Lurigio, W. A. Skogan & R. C. Davis (Eds.) *Victims of Crime: Problems, Policies and Programs*. Sage. Newbury Park, Calif.
- Kelly, J., Kitchin, J. and Cahill, R. 1993. *Report on Victims of Crime*. Report No. 6. Community Law Reform Committee. Australian Capital Territory.
- Kelsen, H. 1957. *What is Justice?* University of California Press. Berkeley.
- Kennedy, L. H. C. 1989. The Advantages of the Inquisitorial over the Adversary System of Criminal Justice. In Whitton 1994 (Appendix I).
- Kenney, D. J. 1987. *Crime, Fear and the New York City Subways*. Praeger Publishers. New York.
- Kleinig, J. 1973. *Punishment and Desert*. Martinus Nijhoff. The Hague.
- Korner, S. (Ed.) 1974. *Practical Reason*. Basil Blackwell. Oxford.
- LaGaipa, J. J. 1977. Interpersonal Attraction and Social Exchange. In S. Duck (Ed.) *Theory and Practice in Interpersonal Attraction*. Academic Press. San Diego.
- Landes, W. M. and Posner, R. A. 1975. The Private Enforcement of Law. *The Journal of Legal Studies*, 4: 1 - 46.
- Lave, L. B. 1962. An Empirical Approach to the Prisoners' (sic) Dilemma Game. *The Quarterly Journal of Economics*, 76: 424 - 436.
- Lewis, C. S. 1949. The Humanitarian Theory of Punishment. *Twentieth Century: An Australian Quarterly*, 3: 5 - 12.
- Lewis, C. S. 1957. *The Problem of Pain*. Fontana Books. England.
- Lovibond, S. 1990. True and False Pleasures. *Proceedings of the Aristotelian Society*, 15: 213 - 230.

- Lucas, J. R. 1980. *On Justice*. Clarendon Press. Oxford.
- Luce, D. and H. Raiffa. 1957. *Games and Decisions*. Wiley. New York.
- Lycan, W. G. 1986. Moral Facts and Moral Knowledge. *The Southern Journal of Philosophy*, 24 (Supplement): 79 - 94.
- Lynd, H. M. 1958. *On Shame and the Search for Identity*. Routledge & Kegan Paul. London.
- MacIntyre, A. 1981. *After Virtue*. University of Notre Dame Press. Notre Dame, Ind.
- Mackie, J. L. 1985. Morality and the Retributive Emotions. In J. & P. Mackie (Eds.) *Persons and Values*. (Vol. 2) Clarendon Press. Oxford.
- Magnus, B. 1978. *Nietzsche's Existential Imperative*. Indiana University Press. Bloomington.
- Malony, J. N. 1980. *I am Ned Kelly*. Allen Lane. Australia.
- Margolis, J. 1966. *Contemporary Ethical Theory*. Random House. New York.
- Marongiu, P. and Newman, G. 1987. *Vengeance, The Fight Against Injustice*. Rowman & Littlefield Publishers. New Jersey.
- Mawby, R. I. and Gill, M. L. 1987. *Crime Victims: Needs, Services, and the Voluntary Sector*. Tavistock Publications. London.
- McCarthy, T. 1994. Victim Impact Statements: A Problematic Remedy. *Australian Feminist Law Journal*, 3: 175 - 195.
- McElrea, F. W. M. 1993. A New Model of Justice. In Brown and McElrea (Eds.) 1993.
- Mellema, G. 1985. Shared Responsibility and Ethical Dilutionism. *Australasian Journal of Philosophy*, 63: 177 - 187.
- Milgram, S. 1974. *Obedience to Authority*. Tavistock Publications. London.
- Mill, J. S. 1962. *Utilitarianism*. M. Warnock (Ed.) Collins. London.
- Miller, W. I. 1990. *Bloodtaking and Peacemaking: Feud, Law, and Society in Saga Iceland*. The University of Chicago Press. Chicago.
- Morris, C. W. 1991. Punishment and Loss of Moral Standing. *Canadian Journal of Philosophy*, 21: 53 - 80.
- Mundle, C. W. K. 1954. Punishment and Desert. In Acton 1969.

- Munson, M. E. 1985. Power and Virtue: Institutional Ascent or Individual Decline? *International Studies in Philosophy*, 17: 27 -31.
- Murphy, J. G. 1979. *Retribution, Justice and Therapy*. D. Reidel. Dordrecht.
- Murphy, J. G. and Hampton, J. 1988. *Forgiveness and Mercy*. Cambridge University Press. Cambridge.
- Nietzsche, F. 1966. *Beyond Good and Evil*. (W. Kaufmann, Trans.) Vintage Books. New York.
- Nietzsche, F. 1967. *The Will to Power*. (W. Kaufmann & R. J. Hollingdale, Trans.) Vintage Books. New York.
- Nietzsche, F. 1968. *Twilight of the Idols and The Anti-Christ*. (R. J. Hollingdale, Trans.) Penguin Books. Middlesex. England.
- Nietzsche, F. 1989. *On the Genealogy of Morals and Ecce Homo*. (W. Kaufmann & R. J. Hollingdale, Trans.) Vintage Books. New York.
- Nietzsche, F. 1993. *Thus Spake Zarathustra*. (T. Common, Trans.) Prometheus Books. Buffalo. New York.
- Nozick, R. 1974. *Anarchy, State and Utopia*. Basic Books. New York.
- Nozick, R. 1981. *Philosophical Explanations*. Clarendon Press. Oxford.
- Nyberg, D. 1981. *Power Over Power*. Cornell University Press. Ithaca.
- Oddie, G. 1996. The Consequences of Action. In Copeland 1996.
- Oddie, G. and Menzies, P. 1992. An Objectivist's Guide to Subjective Value. *Ethics*, 102: 512 - 33.
- Oddie, G. and Perrett, R. W. (Eds.) 1992. *Justice, Ethics, and New Zealand Society*. Oxford University Press. Oxford.
- Ohbuchi, K., Kameda, M. and Agarie, N. 1989. Apology as Aggression Control. *Journal of Personality and Social Psychology*, 56: 219 - 227.
- Olafson, F. 1961. *Society, Law and Morality*. Prentice-Hall. New Jersey.
- Oldenquist, A. 1986. The Case for Revenge. *The Public Interest*, 82: 72 - 80.
- Oldenquist, A. 1988. An Explanation of Retribution. *The Journal of Philosophy*, 85: 464 - 478.

- Parfit, D. 1986. Prudence, Morality, and the Prisoner's Dilemma. In Elster (Ed.) 1986: 34 - 59.
- Patterson, J. 1989. Utu, Revenge and Mana. *British Review of New Zealand Studies*, 2: 51 - 61.
- Patterson, J. 1992a. *Exploring Maori Values*. The Dunmore Press. Palmerston North, NZ.
- Patterson, J. 1992b. A Maori Concept of Collective Responsibility. In Oddie and Perrett (Eds.) 1992.
- Pears, D. 1984. *Motivated Irrationality*. Clarendon Press. Oxford.
- Peristiany, J. G. (Ed.) 1965. *Honour and Shame: The Values of Mediterranean Society*. Weidenfeld and Nicolson. London.
- Perrett, R. W. 1992. Individualism, Justice, and the Maori View of the Self. In Oddie and Perrett (Eds.) 1992.
- Perrett, R. W. and Patterson, J. 1991. Virtue Ethics and Maori Ethics. *Philosophy East and West*, 41: 185 - 202.
- Pettit, P and Braithwaite, J. 1990. *Not Just Deserts*. Clarendon Press. Oxford.
- Pettit, P and Sugden, R. 1989. The Backward Induction Paradox. *The Journal of Philosophy*, 86: 169 - 182.
- Pettit, P. 1986. Free Riding and Foul Dealing. *The Journal of Philosophy*, 83: 361 - 79.
- Pettit, P., Sylvan, R. and Norman, J. 1987. *Metaphysics and Morality*. Basil Blackwell. Oxford.
- Plott, C. R. 1976. Axiomatic Social Choice Theory: An Overview and Interpretation. *American Journal of Political Science*, 20: 511 - 527.
- Polinsky, A. M. 1980. Private versus Public Enforcement of Fines. *The Journal of Legal Studies*, 9: 105 - 127.
- Poundstone, W. 1992. *Prisoner's Dilemma*. Doubleday. New York.
- Primoratz, I. 1989. *Justifying Legal Punishment*. Humanities Press International. New Jersey.
- Quinton, A. 1954. On Punishment. *Analysis*, 14: 133 - 142. Also reprinted in Acton (Ed.) 1969.

- Rachels, J. 1993. *The Elements of Moral Philosophy*. (2nd ed.) McGraw-Hill. New York.
- Railton, P. 1984. Alienation, Consequentialism, and the Demands of Morality. *Philosophy and Public Affairs*, 13: 134 - 171.
- Raphael, D. D. 1955. *Moral Judgement*. Allen & Unwin. London.
- Rapoport, A. and Chammah, A. M. 1965. *Prisoner's Dilemma: A Study in Conflict and Cooperation*. University of Michigan Press. Ann Arbor.
- Rappaport, J., Swift, C. and Hess, R. 1984. *Studies in Empowerment: Steps Toward Understanding and Action*. The Haworth Press. New York.
- Rashdall, H. 1924. *The Theory of Good and Evil*. (Vol. 1) Oxford University Press. London.
- Rawls, J. 1955. Two Concepts of Rules. *Philosophical Review*, 64: 3 - 32.
- Rawls, J. 1984. The Right and the Good Contrasted. In Sandel (Ed.) 1984.
- Richards, D. A. J. 1971. *A Theory of Reasons for Action*. Oxford University Press. London.
- Rizzo, M. J. 1979. The Cost of Crime to Victims: An Empirical Analysis. *The Journal of Legal Studies*, 8: 177 - 205.
- Rook, K. S. 1987. Reciprocity of Social Exchange and Social Satisfaction Among Older Women. *Journal of Personality and Social Psychology*, 52: 145 - 154.
- Ross, W. D. 1930. *The Right and the Good*. Clarendon Press. Oxford.
- Russell, B. 1961. *History of Western Philosophy*. (2nd ed.) Allen & Unwin. London.
- Samuelson, P. A. 1948. *Foundations of Economic Analysis*. Harvard University Press. Cambridge, Mass.
- Sandel, M. J. (Ed.) 1984. *Liberalism and its Critics*. New York University Press. New York.
- Sandel, M. J. 1984. Justice and the Good. In Sandel (Ed.) 1984.
- Sarason, B. R. and Pierce, G. R. (Eds.) 1990. *Social Support: An Interactional View*. Wiley. New York.

- Satz, D. and Ferejohn, J. 1994. Rational Choice and Social Theory. *The Journal of Philosophy*, 91: 71 - 87.
- Schaefer, D. L. 1979. *Justice or Tyranny?* Kennikat Press. New York.
- Schick, F. 1984. *Having Reasons*. Princeton University Press. New Jersey.
- Schwartz, W. F. and Tullock, G. 1975. The Costs of a Legal System. *The Journal of Legal Studies*, 4: 75 - 82.
- Sen, A. K. 1974. Choice, Orderings and Morality. In Korner 1974.
- Sen, A. K. 1977. Rational Fools: A Critique of the Behavioural Foundations of Economic Theory. *Philosophy and Public Affairs*, 6: 317 - 44.
- Seneca. 1958. On Clemency. In Hadas, M. (Ed.) *The Stoic Philosophy of Seneca*. Doubleday Books. New York.
- Severin, P. 1995. *Report on the Visit to the United States of America and Germany*. Queensland Corrective Services Commission. Brisbane.
- Shapland, J., Willmore, J. and Duff, P. 1985. *Victims in the Criminal Justice System*. Gower. Aldershot.
- Singer, P. 1993. *Practical Ethics*. (2nd ed.) Cambridge University Press. Cambridge.
- Smith, A. 1971. *The Theory of Moral Sentiments*. Garland. New York.
- Smith, J. C. and Hogan, B. 1983. *Criminal Law*. Butterworth & Co. London.
- Solomon, R. C. 1989. *A Passion for Justice?* Addison-Wesley. Massachusetts.
- Solomon, R. C. 1990. Justice and the Passion for Vengeance. In Solomon and Murphy (Eds.) 1990.
- Solomon, R. C. and Murphy, M. C. (Eds.) 1990. *What Is Justice?* Oxford University Press. Oxford.
- Spencer, H. 1966. *The Principles of Ethics*. (Vol. 1) Otto Zeller. Osnabruck.
- Stambaugh, J. 1972. Thoughts on Pity and Revenge. *Nietzsche-Studien: Internationales Jahrbuch fur die Nietzsche-Forschung*, I: 28 - 35.
- Stocker, M. 1973. Rightness and Goodness: Is there a Difference? *American Philosophical Quarterly*, 10: 87 - 98.
- Sung Hee Kim and Smith, R. H. 1993. Revenge and Conflict Escalation. *Negotiation Journal*, 9: 37 - 43.

- Tahtinen, U. 1982. *Non-Violent Theories of Punishment, Indian and Western*. Motilal Banarsidass. Delhi.
- Tan, A. 1981. *Revenge Among the Tausug*. Thesis, The Australian National University Library.
- Taylor, A., Rosegrant, T., Meyer, A. and Samples, B. T. 1977. *Communicating*. Prentice Hall. New Jersey.
- Taylor, G. 1985. *Pride, Shame and Guilt: Emotions of Self-Assessment*. Clarendon Press. Oxford.
- Ten, C. L. 1987. *Crime, Guilt and Punishment*. Clarendon Press. Oxford.
- Tversky, A. and Kahneman, D. 1986. The Framing of Decisions and the Psychology of Choice. In Elster (Ed.) 1986.
- Vallentyne, P. 1987. The Teleological/Deontological Distinction. *The Journal of Value Inquiry*, 21: 21 - 32.
- Vallentyne, P. 1988. Teleology, Consequentialism, and the Past. *The Journal of Value Inquiry*, 22: 89 - 101.
- van den Haag, E. 1975. *Punishing Criminals*. Basic Books. New York.
- Vivas, E. 1963. *The Moral Life and the Ethical Life*. Henry Regnery. Chicago.
- Walker, N. 1980. *Punishment, Danger and Stigma*. Basil Blackwell. Oxford.
- Walker, N. 1991. *Why Punish?* Oxford University Press. Oxford.
- Walker, N. and Hough, M. 1988. *Public Attitudes to Sentencing: Surveys from Five Countries*. Gower. Aldershot, Eng.
- Wallace, G. 1995. Wild Justice. *Philosophy*, 70: 363 - 375.
- Whitton, E. 1994. *Trial by Voodoo*. Random House. Australia.
- Williams, G. 1983. *Textbook of Criminal Law*. (2nd ed.) Stevens and Sons. London.
- Wilson, B. R. (Ed.) 1974. *Rationality*. Basil Blackwell. Oxford.
- Wilson, R. W. 1988. *Labyrinth: An Essay on the Political Psychology of Change*. M. E. Sharpe. New York.
- Wiseman, R. 1995. The Megalab Truth Test. *Nature*, 373: 391.

Wolgast, E. H. 1987. *The Grammar of Justice*. Cornell University Press. Ithaca.

Zeid, A. A. M. 1965. Honour and Shame Among the Bedouins of Egypt. In Peristiany
(Ed.) 1965.