

Remorse and Retribution:

Justifying Mitigation at Sentencing

Hannah Maslen, New College, Oxford

Submitted for examination for the degree of Doctor of Philosophy in Law
(Criminology)
September 2011

Word count: 90, 100

Abstract

Remorse can be a powerful source of mitigation at sentencing. However, there is a lack of formal justification for this practice and a paucity of theoretical literature engaging with this issue. Addressing this gap, this thesis offers a comprehensive justification for why an offender's remorse should mitigate the punishment he receives. It begins by discussing the emotion of remorse – its nature and value. With reference to broadly-retributive theories of punishment, it then considers various arguments that could be offered to justify the mitigating effect of remorse on the offender's sentence. It rejects two arguments: either remorse constitutes some of the offender's deserved punishment or remorse reduces the seriousness of the offence. Instead, it develops a justification inspired by philosophical work distinguishing blameworthiness and blaming. The thesis argues that, in the context of sentencing, a broadly-conceived dialogical model of censure is the most legitimate. Remorse, as the offender's ideal input into the dialogue about the offence, modifies the subsequent censure required. If censure seeks a response, and this response is already forthcoming, to nonetheless continue to seek this response as if it were absent devalues the censure. Von Hirsch and Ashworth's assertions that censure appeals to the offender as a rational moral agent, and their adherence to certain quasi-retributive values, are shown to provide further support for these arguments. If the deserved censure is mitigated, then so is the corresponding punishment communicating this censure. The thesis next explores how this justification for mitigation compares with 'mercy' justifications, arguing that the justification offered in this thesis operates more internally to deserved censure, and is more principled, so is preferable on these grounds. In conclusion, the thesis considers the implications of its arguments for sentencing practice and whether it is a concern that they are valid only within 'censure' theories of punishment.

Contents

Chapter One: Introduction	6
The importance of remorse to sentencers and the general public	6
The need for theoretical justification	7
The problem of establishing genuineness	10
Focus on retributivism	10
Outline and aims of the thesis	12
<u>Section One: Remorse</u>	
Chapter Two: Conceptualising Remorse.....	17
Overview	18
Socio-linguistic influences on the categorisation of emotional experience.....	20
Some preliminary comments on cognitions in emotions.....	24
Remorse	25
Shame.....	33
Shame and remorse	46
Guilt feelings.....	48
Guilt feelings and remorse	52
Regret.....	54
Regret and remorse	57
Clarifying remorse	57
Return to social construction and the reality of emotional experience.....	59
Final thoughts	62
Chapter Three: The appropriateness of remorse as a response to moral wrongdoing	64
Overview.....	64
An appropriate distinction.....	65
Aristotle and virtue ethics	69
The value of emotions in general.....	70
Objections to the appropriateness of remorse.....	73
Bittner’s objection: remorse is not reasonable.....	75
The fleeting objection	93
Concluding Remarks.....	94
Chapter Four: Can punishment constitute or substitute for some of the appropriate remorse?	96
Is experiencing remorse phenomenologically similar to the experience of being punished?	97
Even if the phenomenology of remorse and punishment are not the same, are the elements we value in remorse also present when an offender experiences punishment?	103
Is there <i>anything</i> punishment can achieve as a substitute for remorse? Are the (more tangible, utilitarian) consequences of remorse also achievable through punishing?.....	110
Perhaps punishment can only significantly substitute for the penitentiary behaviours motivated by remorse: Is punishment a substitute for making amends?.....	115
What is left out when there is punishment but no remorse?	119

Is substituting for remorse what we are trying to do when we punish anyway? ...	120
Conclusion	121
Section Two: Remorse and Retributive Punishment	
Chapter Five: Mitigation and retributive punishment.....	123
Overview.....	123
Ambiguity surrounding the concept of mitigation.....	124
Possibility one: exculpation	125
Possibility two: yields lesser offence.....	126
Possibility three: a determinant of offence seriousness	128
Possibility four: an adjuster of offence seriousness.....	131
Possibility five: relevant external to offence seriousness	132
Possibility six: motivates leniency ungrounded in principle	134
Mitigation in summary.....	136
Mitigation as a particularly retributive concern.....	138
Retributive punishment.....	141
The possibilities for remorse.....	146
Chapter Six: Does remorse constitute some of the deserved punishment?	149
Natural punishment.....	149
Punishment proper	151
The subjective vs. objective metrics of punishment debate.....	152
Theory one: benefits and burdens.....	155
Theory two: Duff and the ‘penance perspective’.....	167
Theory three: Von Hirsch and Ashworth: censure and sanction	179
Conclusion	185
Chapter Seven: Can remorse reduce the seriousness of the offence?	187
<u>Section One: Harm</u>	
Harm as the relevant factor	188
Von Hirsch and Jareborg’s ‘living standard’ analysis of harm.....	190
Conclusion	203
<u>Section Two: Culpability</u>	
Overview.....	204
‘Immediate repentance’	205
Quasi-retributive grounds for mitigation	208
The ‘time delay’ argument.....	208
Time, change and personal Identity	209
Time, change and blameworthiness.....	218
Conclusion	231
Chapter Eight: Developing the positive thesis: remorse and censure in a dialogical context.....	232
Overview.....	232
The extent to which we should blame the offender	233
Censure as central to communicative retributive theories	239
Communication as dialogue.....	242
Remorse as the offender’s ideal, relevant contribution to the dialogue.....	250
Further support from quasi-retributive grounds and von Hirsch and Ashworth’s time-delay argument	251

Mitigation of punishment is necessitated by the change in the meaning of the censure	254
An argument against the plausibility of thinking of punishment in dialogic terms.....	255
Summary	268
Conclusion	270
Chapter Nine: Comparing my account with others.....	272
Comparison one: an account with similarities	272
Comparison two: dialogical censure and the exercise of mercy	277
A closer look at the internal-external distinction	282
Further reasons to prefer my account.....	293
Why we need a principle rather than merciful discretion	294
Conclusion	296
Chapter Ten: Conclusion	298
How can such an abstract theory be translated into everyday practice?.....	300
What do your arguments mean for the quantum of mitigation?	300
Doesn't your justification require acceptance of a censure-based desert theory? .	303
Bibliography	305

Chapter One

Introduction

If someone commits an offence and subsequently demonstrates profound remorse, should this have any bearing on the severity of the sentence he or she¹ receives? This thesis attempts to answer this question, exploring the nature and value of a remorseful response to wrongdoing, and why it might justify mitigating the punishment the state imposes on the offender.

The importance of remorse to sentencers and the general public

Remorse is one of the most powerful sources of personal mitigation. In borderline cases, it often tips the balance towards a non-custodial sentence (Tombs and Jagger 2006). In the US, it can make the difference between a sentence of life or death (Sundby 1998). Recent research in England and in Scotland has demonstrated that sentencers consider remorse an important mitigating factor. Jacobson and Hough (2007: 17) reported that, within the sample of judges they interviewed, ‘intense remorse’ ranked 4th out of 13 personal mitigating factors in the extent to which they influenced the sentence imposed. The research also showed that remorse was considered a very important mitigating factor, *despite* the difficulty, on occasion, of establishing whether it was genuine. One judge is quoted as saying that while he might be ‘conned’ by a defendant pretending to be remorseful, ‘I’d rather make that mistake than find a [genuinely] repentant, reformed offender and reject his pleas ...’ (ibid.: 49).

¹ Throughout the thesis I shall mainly use the pronoun ‘he’ to denote hypothetical individuals as an arbitrary simplification device.

Commitment to the importance of remorse also emerged from Tombs' and Jagger's (2006) research with judges and sheriffs in Scotland. One sheriff is quoted as saying: 'The nature of the response of the accused is important. Even in relatively serious cases, where you must take a very serious view in the public interest, evidenced contrition should be taken into account. The way in which contrition is carried through can indicate that another option [to custody] is possible' (2006: 816). Indeed the article concluded that 'signs of remorse' and 'signs of hope' frequently tipped the balance for these sentencers in favour of a non-custodial sentence in borderline cases.

Remorse is also considered important by the general population. Research shows that members of the public take remorse into account when making hypothetical sentencing decisions: they reduce the severity of the sentences they recommend when the offender is depicted as remorseful (Robinson and Darley 1995; Scher and Darley 1997). By way of illustration, a study carried out in England and Wales found that 21% of the sample believed that an offender showing remorse 'should result in more lenient sentence in all or most cases'. A further 56% believed that an offender showing remorse 'should result in more lenient sentence in some cases' (Roberts *et al.* 2009: 777).

The need for theoretical justification

That something happens in practice and generally finds support is not always enough to justify it. This is particularly the case in matters of justice, where we are dealing with an institution burdened with the serious task of getting punishment 'right'. As far as possible, sentencing practice should emerge from a coherent, principled approach. It follows that part of this task is to justify the operation of mitigating factors,

explaining how they are compatible with the key principles on which practice is based. As was seen above, remorse is an important mitigating factor in some jurisdictions. In England and Wales it appears in the sentencing guidelines as a factor in personal mitigation. However, whereas some factors have well-established justifications (for example: involvement due to coercion, intimidation or pressure is justified as a partial excuse, as it indicates reduced culpability) others have received little theoretical attention. The need to address the paucity of positive justification is made even more pressing when literature on the subject provides arguments *against* remorse being justified as a mitigating factor.

Arguments denying a role for remorse

The arguments against remorse being a mitigating factor tend to fall into two categories. Either, they focus on the difficulty of establishing whether displays of remorse are genuine, arguing that too many mistakes would be made, or they argue that allowing remorse to mitigate is inconsistent with a particular sentencing philosophy.

Ward (2006) argues that remorse should not be relevant in criminal sentencing because it is an ‘increasingly ambiguous concept’, resulting in its application by sentencers being completely subjective (ibid.: 131). Further, he argues that allowing remorse to be considered at sentencing does not necessarily advance any theory of punishment. He suggests that for those who see punishment as a mechanism intended to remould the defendant into an accepted and productive member of society, the use of remorse for purposes of punishment is unfair and self-defeating. In relation to retributive theories, he argues that remorse is neither relevant to the blameworthiness of the offender, nor does it repair the harm caused. Thus, he concludes, remorse does not assist in determining the appropriate severity of the sentence. In addition to identifying theoretical

arguments and expressing concern over the lack of clarity of concept, Ward (2006) also suggests that the practical difficulties involved in sorting the truly remorseful defendant from the unrepentant but savvy defendant should further weigh against giving remorse a role at sentencing.

Bagaric and Amarasekara (2001) argue that remorse is irrelevant to sentencing, irrespective of which of the main theories of punishment one adopts. Similarly to Ward, with regards to retributivism, they argue that remorse does not affect the severity of the offence so has no bearing on the deserved punishment. With respect to utilitarian theories they argue that neither specific deterrence nor rehabilitation can be achieved through a system of state-imposed sanction, and that therefore any sentencing factors that derive their relevance from their association with these ideals must also be disregarded. This second argument is as much dissatisfaction with utilitarian aims of punishment as with remorse.

Murphy (2006) presents both the epistemic problem that remorse can be disingenuously simulated and further theoretical arguments. He suggests that general deterrence might be undermined if we allow remorse to mitigate. His argument is that if it becomes known that one way to avoid serious punishment is to express remorse, then people may be more inclined to commit offences that they would otherwise have seen as involving too high a potential cost. From a retributive perspective, he argues that a concern with remorse would require measuring desert against the offender's character, rather than the offence. Sentencing on character is seen as a very unattractive approach by most desert theorists.

In the face of such arguments, the burden is placed on those who want to retain remorse as a mitigating factor to offer a justification. Providing such a

justification is what this thesis aims to do. However, the scope of the thesis will be limited in two ways, which are set out below.

The problem of establishing genuineness

It cannot be denied that there are elements of uncertainty involved in assessing whether an offender is genuinely remorseful or simply simulating remorse in an attempt to secure a reduction to his sentence. This being said, and as the judicial comments above suggest, many judges do claim to be able to fairly confidently assess sincerity. Further, recent psychological research (ten Brinke *et al.* 2011) has made important steps in analysing facial, verbal and body language behavioural differences in displays of genuine and falsified remorse. Such research could provide advice to help sentencers feel even more confident about their assessments.

However, this thesis is not going to deal with the problem of genuineness. The prior question is whether remorse is theoretically relevant in the first place – this must be established before we turn to seeing what problems might be involved in putting it into practice. Further, if it can convincingly be argued that remorse *should* have a mitigating effect then it falls to practitioners to try to work out how best to deal with any practical problems. In the face of a firm justification, not to do so would be to fail to act in the interests of justice. So, I put aside the problem of how we might judge the sincerity of an offender's remorse and focus on the question that logically precedes it.

Focus on retributivism

The other restriction to the scope of the thesis is that it will confine the search for justification to (broadly) retributive theories of punishment. Establishing whether remorse is of relevance to utilitarian theories – theories concerned with maximising

the positive consequences of punishment for society – would require engaging with empirical research into remorse and deterrence, rehabilitation and reform. To date, there has not been the empirical research conducted to enable any strong conclusions to be drawn, and the evidence that does exist is itself inconclusive (see Cox 1999).

Further, versions of ‘just deserts’ theories (the more contemporary versions of retributivism) and an emphasis on proportionality underlie the sentencing policies of many Western states. In England and Wales, just deserts and the proportionality principle have a major role to play in sentencing (Ashworth 2010: 104-5). The jurisdictions of Finland, Sweden and Minnesota have also fully embraced proportionality theory (ibid.: 89). ‘Just deserts’ has been adopted in the United States more generally (Tonry 1996) and the U.S. Supreme Court has said that retribution is a legitimate purpose of criminal punishment (see Finckenaaur 1988: 81).

Punishment is more fundamentally linked to a retributive rationale. People may be punished and yet fail to be deterred, rehabilitated or reformed; and deterrence, rehabilitation and reform can be achieved without punishment. However, punishment is necessarily imposed as a response to something someone has done, which is at the heart of retributivism. Without theoretical commitments to desert and proportionality, criminal justice systems would look very different. Given the prominence of retributive concepts in Western jurisdictions, and the absence of the empirical research necessary to explore the relevance of remorse to consequentialist aims of punishment, this thesis will be dedicated to the project of justifying remorse-based mitigation within retributive theories of punishment.

It should be noted that jurisdictions that adopt multiple aims for sentencing will still find a deserts-based justification applicable if considerations of desert and

proportionality are amongst those multiple aims. In reality, proportionality is rarely the sole principle or value pursued by a jurisdiction's sentencing policy.²

Outline and aims of the thesis

The overall aims of this thesis are twofold. First, I will explore the nature and value of remorse so that we emerge with a clear conception of what it involves and its potential relevance to sentencing. Second, I will explore whether remorse can be justified as a mitigating factor: I will explain why some potential justifications do not succeed and finally develop an account which I argue does justify the mitigating effect of remorse. Thus, the aims can be summarised as: understanding remorse and its relevance to sentencing. The thesis will proceed as follows:

Section One: Remorse

Section One comprises three chapters (Two, Three and Four) which focus on remorse as a response to wrongdoing.

Chapter Two explores the nature of remorse: its affect, cognition and behaviours. It compares it with other, related emotions. It suggests that, of these emotions, remorse should be the response to wrongdoing of relevance to sentencing. It also considers the practical difficulty of establishing which precise emotion is being experienced (often emotions overlap and vacillate) and what should therefore be minimally present to be of justifiable interest to sentencers.

Building on the suggestion that remorse is the emotion relevant to sentencing, Chapter Three positively argues that remorse is the appropriate response to moral

² For example, although proportionality is predominant in sentencing in England and Wales, courts must have regard to the following purposes of sentencing: a) the punishment of the offenders, b) the reduction of crime (including its reduction by deterrence), c) the reform and rehabilitation of offenders, d) the protection of the public, and e) the making of reparation by offenders to persons affected by their offences. (Section 142 of the Criminal Justice Act 2003).

wrongdoing. It deals with objections to the contrary that find no value, or even disvalue, in remorse.

Chapter Four serves as a bridge to Section Two, introducing punishment to the discussion in order to complete the exploration of the nature and value of remorse. It asks whether, given the value I have ascribed remorse, punishment can constitute or otherwise substitute for remorse, where it is not forthcoming. Whilst the answer is for the most part that it cannot, I also question whether substituting for remorse is anyway what we attempt to achieve by punishing.

Section Two: Remorse and Retributive Punishment

Section Two (Chapters, Five, Six, Seven, Eight and Nine) is concerned with relating remorse to retributive punishment and finding a justification for a mitigating role in this context.

Chapter Five introduces the concept of mitigation in the sentencing process, separating out the different mechanisms through which mitigation can operate, and highlighting which of these might be possibilities for remorse. This chapter also explains why mitigation is a particularly retributive concern and introduces the three retributive theories that will be used to test my arguments.

Chapter Six explores the first possible justification: that remorse might constitute or otherwise substitute for some of the deserved punishment. It argues that none of the three theories are amenable to this justification, although Duff's (2001) theory will be argued to have to concede some mitigation if the remorseful offender were to have already imposed upon himself a particularly burdensome penance.

Chapter Seven asks whether remorse can reduce the seriousness of the offence. This involves exploring whether remorse can reduce the harm of the offence or the

offender's culpability. The longer discussion of culpability raises issues of personal identity and explores two conceptions of 'blameworthiness' which remorse might serve to diminish. However I ultimately argue that remorse can convincingly reduce neither the harm of the offence nor the blameworthiness of the offender.

Chapter Eight introduces my positive thesis. Starting from a distinction between the offender's blameworthiness and the extent to which we should blame him, it develops a dialogical account of sentencing which is argued to be an improvement on a 'one-way' communication model. Dialogical censure, I argue, is more legitimate than that which universally assumes no understanding and no moral transformation. Correspondingly, it is suggested that the offender's remorse is a relevant input into the dialogue and that the censure the state delivers should be sensitive to this input: the nature of the censure that is appropriate is changed by the offender's communication of his remorse. Reducing the censure (and hence proportionate punishment) responds to the offender in a way consistent with having 'heard' his remorse. Von Hirsch and Ashworth's assertion that censure appeals to the offender as a moral agent capable of deliberation, and their adherence to certain quasi-retributive values, are shown to provide further support for these arguments.

Chapter Nine extends my account by comparing it with two others. The first involves similar ideas but, in pointing out the differences, my account is honed and argued to be preferable. The other comparison is between my account and mercy accounts which also claim to justify mitigation on the grounds of remorse. I show why the justification – based on the intimate link between remorse and censure – depended on in my account is preferable to the more diffuse concern for the offender's wellbeing involved in mercy accounts.

The thesis therefore offers a justification for accepting remorse as a mitigating factor at sentencing. This will be an important contribution – both to academic theoretical discussions of sentencing, and to the practical project of developing coherent sentencing policy and guidelines. In England and Wales, there is currently no formal justification for allowing remorse to play the mitigating role that it so often does. Appeals to improvement in the offender’s character are seen as untenable from the perspective of modern desert theories – sentencing on character is to be avoided. My thesis offers what I believe to be a robust alternative, and argues that the principal modern retributive theories are committed to tenets and values that given them reason to concede my arguments.

Section One:

Remorse

Chapter Two

Conceptualising Remorse

The violator of the more sacred laws of justice can never reflect on the sentiments which mankind must entertain with regard to him, without feeling all the agonies of shame, and horror, and consternation. When his passion is gratified, and he begins coolly to reflect on his past conduct, he can enter into none of the motives which influenced it. They appear now as detestable to him as they did always to other people. By sympathizing with the hatred and abhorrence which other men must entertain for him, he becomes in some measure the object of his own hatred and abhorrence. The situation of the person, who suffered by his injustice, now calls upon his pity. He is grieved at the thought of it; regrets the unhappy effects of his own conduct, and feels at the same time that they have rendered him the proper object of the resentment and indignation of mankind, and of what is the natural consequence of resentment, vengeance and punishment. The thought of this perpetually haunts him, and fills him with terror and amazement. He dares no longer look society in the face, but imagines himself as it were rejected, and thrown out from the affections of all mankind. He cannot hope for the consolation of sympathy in this his greatest and most dreadful distress. The remembrance of his crimes has shut out all fellow-feeling with him from the hearts of his fellow-creatures. The sentiments which they entertain with regard to him, are the very thing which he is most afraid of. Everything seems hostile, and he would be glad to fly to some inhospitable desert, where he might never more behold the face of a human creature, nor read in the countenance of mankind the condemnation of his crimes. But solitude is still more dreadful than society. His own thoughts can present him with nothing but what is black, unfortunate, and disastrous, the melancholy forebodings of incomprehensible misery and ruin. The horror of solitude drives him back into society, and he comes again into the presence of mankind, astonished to appear before them, loaded with shame and distracted with fear, in order to supplicate some little protection from the countenance of those very judges, who he knows have already all unanimously condemned him. *Such is the nature of that sentiment, which is properly called remorse*; of all the sentiments which can enter the human breast the most dreadful. It is made up of shame from the sense of the impropriety of past conduct; of grief for the effects of it; of pity for those who suffer by it; and of the dread and terror of punishment from the consciousness of the justly provoked resentment of all rational creatures. [Smith 2002: 98-99]

In the above passage, Adam Smith provides a vivid description of the thoughts and feelings that an individual who has violated ‘the more sacred laws of justice’ may

experience.³ I quote him at length to illustrate the multifaceted – and to some extent ‘messy’ – nature of emotional experience; and the challenge involved in fully capturing or defining such experiences. The emotional experience Smith describes is complex. Objects of the individual’s attention include: the attitudes of others towards the individual and his crime, his own conduct, the plight of the victim, the consequences of the crime more broadly, his social and ontological status in society, and the punishment that is likely to be imposed on him. The range of emotions precipitated includes: shame, horror, consternation, self-hatred, pity, grief, regret, terror, amazement, distress, fear, and dread. Smith states that this ‘sentiment’ – the complex experience of all these thoughts and feelings – is remorse. He also suggests that it is the most dreadful sentiment a human being can experience.

Overview

If we are going to ask whether remorse can be justified as a mitigating factor in sentencing, we need a clear understanding of what we mean by ‘remorse’. I will eventually argue that the ‘messiness’ of emotional experience realistically precludes consistent application of sharply-delineated descriptions of emotions. This being said, there is still value in characterising the paradigmatic instance of remorse and in attempting to discern what is essential to any instance. Moreover, we must do this if we are to orientate ourselves with regards our key concept.

Remorse is not the only emotional response an individual might have to his wrongdoing. Shame, guilt and regret are other possible responses. Moreover, these emotions may well overlap in aspects of their phenomenology or be experienced concurrently. In the above passage, Smith’s description of remorse involves shame

³ I would add that an individual’s non-criminal moral wrong could also beget similar if not identical thoughts and feelings. Indeed, the ‘most sacred laws’ to which Smith refers are likely to be those covering the particularly morally-reprehensible.

and regret. Empirical studies carried out by Davtiz (1969) and Shaver *et al.* (1987) both revealed a substantial overlap in aspects of the experiences of remorse, shame, regret and guilt.

With this in mind, my aim in this chapter is to comprehensively analyse various descriptions and characterisations of remorse and the other closely-related emotional reactions to one's wrongdoing: shame, guilt and regret. I will draw attention to what about these emotions is agreed or disagreed upon. Particularly where there is disagreement, I shall examine the positions and argue that one is the most plausible. I shall explore the differences between these emotions enabling, as far as possible, a clear understanding of remorse as a distinct moral emotion. Although most of the literature I shall be discussing will be philosophical, I shall also examine relevant empirical – particularly psychological – literature.

Philosophical and empirical modes of enquiry

Something must be said here about the relationship between the prescriptive enterprise of defining concepts somewhat *a priori*, and the descriptive (or exploratory) enterprise of empirical research. There is a danger that, with thoroughly-considered and precisely-delineated concepts worked out, the philosopher might disregard contradictory empirical research on the basis that lay people do not fully understand the concepts being explored. The philosophical project then is liable to becoming a game of internal coherence, with no assurance that the result corresponds to reality and general experience. As Deigh (1996: 39) says with regards to the philosophical study of psychological concepts, 'philosophers cannot live by intuition alone'. This being said, many people – and hence many people who take part in psychological studies – will have not spent much time analysing and cataloguing their various moral

emotions and may indeed be subject to imprecision or conceptual confusion. We thus need to be aware not to assume the priority of one method of producing knowledge over the other.

The chapter will proceed thus: I will begin by outlining what I believe to be an instructive framework for thinking about the conceptual analysis of emotions. I will argue that, to some extent, emotion labels are influenced by culture and language, and that we should be sensitive to the fact that our vocabulary may not map one-to-one onto naturally independent mental phenomena. Then, following a preliminary characterisation of remorse, the emotions of shame, guilt and regret will be discussed, enabling comparison and clarification. Having identified, as far as possible, the essential characteristics of remorse, it will then be suggested that the reality of emotional experience may not correspond with neat theoretical distinctions. The implications of this will be discussed.

Socio-linguistic influences on the categorisation of emotional experience

Although my discussion of remorse will be principally philosophical – in that it will be concerned with conceptual analysis and moral evaluation – I wish to introduce a framework for thinking about remorse (and emotions in general) that is not prevalent in existing philosophical thought. This framework is derived from the socio-linguistic tradition of thinking of emotions as partially socially constructed and intimately linked to language.

Some of the philosophical discussions of remorse and related emotions treat these phenomena more like natural kinds than is perhaps warranted. Whilst there is no denying that there are observable physiological and psychological markers associated

with emotional experience, it is questionable whether emotions can be examined and neatly categorised in the same objective way as, for example, chemical elements.

Whilst there are likely to be some basic emotions – such as fear, sadness and anger – that are universal (see Ekman 1992), other, more social emotions – such as those explored in this chapter – will be influenced by culture and language. Human beings pick out and linguistically label what is important to them, including their emotional experiences about which they wish to communicate. Different labels suggest an at least perceived differences between the phenomena to which they refer.

This does not mean that different nations differ in their capacities for physiological and psychological arousal but, rather, that these patterns of arousal are sometimes interpreted and labelled differently. It also does not mean that emotions would not exist without language, but the ways in which they are distinguished and conceptualised is influenced by language.

In their study of the universal development of emotion categories in natural language Hupka *et al.* (1999: 248) articulate this view as follows: ‘emotion words are not simply referential labels for putative, universal, internal feeling states but, more importantly, are about social relations, particularly power relations, involving meaning-making practices of individuals engaged in ordinary conversation and interaction’. This does not mean that by exploring emotion words in different languages we can work out which emotions really ‘exist’ in an ontological sense, but instead that we can understand emotions best with reference to their socio-linguistic contexts.

Importantly, Hupka *et al.* (1999: 248) stress that they ‘agree with R. Brown (1958) and Beeman (1985) that languages are neither advanced nor primitive. They fulfil the communication needs of their speakers.’ I would also add that the number of

labelled ‘moral’ emotions in a language should not invite judgements about the moral integrity of a language’s speakers.

Languages do fulfil the communicative needs of their speakers. For example, Hupka *et al.* (1999) report that, due to particular needs, and in contrast to English, the Karok (Bright, 1957) and Yurok (Robins, 1958) languages of the California Indians have approximately 20 words related to different conditions of acorns, such as acorn dough, acorn water, acorn flour, mouldy acorns, and leached acorns. In terms of emotion words, there are also differences between languages.

Russell (1991) provides evidence that words in the English language that are often assumed to denote natural basic categories of emotion have no equivalents in some other languages. Conversely, other languages contain commonly-used emotion words with no equivalent in English. Both of these scenarios can be illustrated by the following comparisons, relevant to the moral emotions that will be discussed in this chapter.⁴

The English distinction between *shame* and *fear* is not made by the Gidjingali aborigines of Australia (Hiatt 1978). Both are covered by the same word, *gurakadj*. The English distinction between *shame* and *embarrassment* is not made by the Japanese (Lebra 1983: 194). Levy (1973: 342) observed that the Tahitians have ‘no word which signifies anything like a sense of guilt’. A word for *guilt* is missing from the Sinhala language of Sri Lanka (Obeyesekere 1981: 79), from the Ilongot language of the Philippines (Rosaldo 1983: 139-40), from the Pintupi language of aboriginal Australians (Morice 1978: 93), and from the Samoan language (Gerber 1975). Guilt is subsumed under *metagu* (fear or anxiety) for the Ifalukians (Lutz 1980: 223).

⁴ All of the following examples are taken from Russell’s (1991) paper.

So here we have instances where emotions that we (English speakers) think of as distinct are, in different socio-linguistic contexts, conflated or thought to overlap in phenomenology. In others, there is no word for an emotion (guilt) that English speakers refer to on a regular basis. There are further examples, however, where there are more (or, at least, different) distinctions between emotional experiences than in the English language. Indonesian, which does not distinguish shame from embarrassment, does distinguish shame/embarrassment brought on by one's own deeds, *malu*, from shame/embarrassment brought on by someone else's deeds, *dipermalukan* (Keeler, 1983: 153).

The main point I am making is that the way that we divide and describe the emotions in the English language does not necessarily correspond to a naturally determined set of distinct experiences. We should not be surprised (or too concerned) if the emotions of remorse, shame, guilt and regret are not as sharply independent of each other as is sometimes conceptually expected.

However, I am not suggesting a wild relativism with respect to emotional experience where individuals from different cultures (or even within cultures) have greatly different emotional lives. Human beings will experience similar things in similar situations, but may think and talk about these experiences differently. Moreover, as Russell (1991: 434) reminds us: 'even if a language lacks a word for, say, guilt, there remains the possibility that guilt may be expressed in a phrase, or metaphorically, or even nonverbally'.

Having introduced the idea that emotional labels are to some extent dependent on language and culture, I will now review what theorists have taken our particular labels to denote. The different concepts to which these labels refer highlight important differences in the various ways we respond to wrongdoing. I will return to the socio-

linguistic perspective at the end of the chapter when I suggest that remorse as a labelled concept might have come about as a way of denoting the (generally conceived) ideal response to wrongdoing, and so by definition must be the response we should hope to witness in the offender.

Some preliminary comments on cognitions in emotions

In my discussion below, beliefs and thoughts will be prominent in the characterisations of the four emotions. This being so, I want to make clear my position on the cognitive elements of emotions and the importance it has for how we think about the emotions of offenders. Cognitivists argue that emotions involve propositional attitudes: beliefs, thoughts etc. Extreme views *identify* emotions with these judgements (e.g. Nussbaum 2001). Others, however, have argued that emotions do not require cognitions at all (e.g. Deigh 1994). Whilst it is possible to conceptually isolate the affective element of an emotion – it feels hot or cold; pleasant or unpleasant – to understand emotions as only involving this element, I argue, would leave us with an incomplete and emaciated account.

Instead, I take what might be called a ‘weak’ cognitivist view, whereby, at minimum, emotions must require the world being presented in a certain kind of way to the individual experiencing the emotion. Thus, I agree with Charles Taylor (1985: 71) when he says that ‘experiencing an emotion essentially involves seeing that certain descriptions apply’. So, whilst a feeling cannot be one of remorse unless there is a sense of my having done wrong, I might not yet be able to fully articulate what is

wrong about what I have done or whether I assent to my actually having done wrong. The emotion prompts further articulation.⁵

If explicit beliefs are not *necessary*, one might object before we have even started that any emotions felt by offenders tell us nothing about what they think about what they have done and so are inconsequential to how we should attend to them when sentencing. At best, they might say, nominal value can be found in the emotion's showing that the agent has some sort of awareness of, and unease at, others' displeasure at him. However, whilst it is possible that in feeling remorse initially, an offender divests himself of it by articulating to himself the belief that his sense of wrong-doing is unfounded, it is still the case that his situation is presented to him in a certain way by his emotion.

Moreover, 'mistaken' emotions are not the rule and it is crucial to acknowledge that emotions prompt one to understand oneself: the articulation an emotion precipitates can reveal to the offender that he has done wrong, even if the explicit beliefs did not accompany the basic feeling initially. Perhaps an assertion of belief in wrongdoing is required for deep remorse, even if explicit cognitions are not essential for a less intense affective precursor. In the following, therefore, I shall discuss the thoughts and beliefs that accompany the emotions with the caveat that they may not be fully articulated in emergent or misplaced affect.

Remorse

I am going to argue for a set of characteristics that are essential to remorse. The characteristics are as follows: First, the agent must perceive himself to have done something that he considers to be morally wrong. Most often this involves having

⁵ This feature of affective experience will become particularly important in the discussion of the value of remorse conducted in Chapter Three.

harmed another person in some way. Second, the agent must assume responsibility, humbly seeing himself as blameworthy. Third, the focus of the agent's attention is on the reality of the individual he has wronged, as well as on himself as the one who wronged the other. Fourth, there is a desire to atone, which is likely to manifest itself in apology and reparative activity. Fifth, there is at least an intention not to repeat the harmful action. It will be my task in the remainder of this chapter to explain why each of these characteristics is essential to remorse and why this set in its entirety is unique to remorse.

Gaita on Remorse

Raymond Gaita provides one of the richest descriptions of remorse and its connection to morality. He says little that contradicts, and much that expands on, other analyses of remorse in the literature. For this reason, I shall begin with his conceptualisation and then go on to consider the main points of agreement and any points of disagreement found in other accounts.

What Gaita emphasises in particular about remorse are the ways in which the reality and value of other people (the victim(s)) are revealed in an experience of remorse, and its capacity to fundamentally isolate the remorseful individual. Gaita (1999: 34) explains:

Our sense of the reality of others is partly conditioned by our vulnerability to them, by the unfathomable grief they may cause us. It is also conditioned by our shocked and bewildered realisation of what it means to wrong them. Remorse is that realisation...It is the perspective in which the meaning of what one has done, what one has become through doing it, and what one's victims have suffered through doing it, are inseparable.

Remorse, then, is a pained realisation of the reality of others and what it means to wrong them: 'Pained bewilderment is the most natural expression of remorse. "What

have I done? How could I have done it?” (ibid.: 31). These questions, that illustrate the intense anguish of remorse, demonstrate a focus both on the wrongdoing and on the self: a shocked realisation of the harm done to the victim, in all his individuality, and horror at the realisation that *I* was capable of doing it. These are inseparable because, as Gaita notes, ‘that *we* did it is internal to the character of *what we did* and...to what the victim suffered’ (1991: 54).

Gaita argues that we learn about what is really ‘evil’ – the harming of other people – through studying remorse. In this way he urges a resistance to Kantian-style conceptions of morality that might depict ‘evil’ in terms of rule violation and transgression of moral codes. When we consider a remorseful murderer, for example, we find he is not haunted by the principles he betrayed or by the Moral Law he transgressed; he is haunted by the particular individual he killed. Thus: ‘Reflection on remorse takes us closer ... to the nature of morality and of good and evil, than reflection on rules, principles, taboos and transgressions can’ (1999: 32). Any moral theories that say that a murderer discovers in her remorse how terrible it is to become someone who broke a certain rule or principle are clearly inadequate. According to Gaita, we learn something about the seriousness of murder through the way the murderer becomes haunted by his victim: ‘remorse is an awakening to the terribleness of what was done’ (Gaita 1991: 52).

The second of Gaita’s particularly instructive insights into remorse is what he describes as its rendering one ‘radically singular’. He explains that remorse is the only kind of human suffering that can find no consolation from human fellowship. Those who grieve can find comfort from others who grieve, knowing that they are not alone in their suffering. In remorse, however, there can only be corrupt consolation in the knowledge that others are as guilty as we are. In contrast to the consolation that the

grieving or fearful or frustrated may find in human fellowship, 'it is different with guilt. It should be no consolation if what we did was also done by the best of people. That is not pride because remorse does not focus on what kind of person we are. Its focus is on what we have become only because we have become wrongdoers' (Gaita 1991: 50).

Gaita explains that as well as remorse going hand-in-hand with understanding the wrong one has done, it also goes hand-in-hand with reparation. He argues that someone who often and tearfully expresses remorse but is never prepared to make reparation has a 'desperately thin moral understanding' (1999: 100). Conversely, he claims that attempts to make good the damage done whilst feeling no remorse and seeing no need to apologise are equally unsatisfactory. In fact, he suggests that relief of material and psychological damage caused does not count as reparation unless the spirit in which that relief is given is informed by a recognition of the wrongs the victim suffered: 'That is part of what we *mean* by "reparation" and it is why we distinguish reparation from other actions which would bring the same material benefits to those who have been wronged' (Gaita 1999: 101).

So, on Gaita's account, remorse occurs when one realises that one has harmed another human being, becoming aware of his reality and what having harmed him means about the self. It renders one radically singular, in the sense that one cannot enter into consoling fellowship over one's remorse. Remorse inspires reparative efforts which have a particular quality that is not shared by other kinds of compensation unaccompanied by remorse.

Remorse elsewhere in the literature

Having outlined one comprehensive and persuasive account of remorse, we can now examine the views of other theorists, noting agreement and any disagreement. Remorse is generally agreed to be a particularly painful emotion. Indeed, according to Adam Smith, it is the ‘most dreadful’ sentiment a person can experience. Tudor (2001) describes the ‘horror’ and the ‘bitter and deeply painful’ regret that remorse involves. Deigh (1996) refers to the ‘inner torment’ suffered. Such emotive descriptions make sense given the etymology of the word ‘remorse’. The word comes from the Latin root ‘mordere’ (to bite), with the prefix ‘re-’ denoting repetition (Gorelick, 1989). Thus, we get the sense of a relentless biting pain. Roberts (2003: 222), however, suggests that remorse is not always so intense, saying the emotion can range from a ‘pithless momentary urge to atone to an enduring burning pain that can reorient an entire life’. On the whole, though, it seems that the feelings involved in remorse are most characteristically painful and intense.

It is generally agreed that remorse is felt following a ‘sin’, ‘moral wrong’ (Taylor 1985) or ‘evildoing’ (Deigh 1996). It is also agreed that remorse involves a judgement that one is to blame for the action, and acceptance of responsibility for the wrongdoing (Thalberg 1963; Soloman 1976; Tudor 2001; Taylor 1985). Paradigmatically, this wrongdoing involves harm to another person. On Gaita’s account, it would seem that harm to another human being is necessary for an experience of remorse. He says, ‘in remorse we respond to what it means to wrong another, which involves a new and terrible shock at their reality’ (Gaita 1999: 102). If this is exhaustive of what remorse is and, indeed, is internal to its grammar, then remorse cannot be felt over victimless wrongdoing.

However, other writers do not agree that remorse is this limited in its extension. Is it not intelligible that remorse might be felt over animal cruelty, for example? Further, what about when there is no conceivable suffering at all that results from the wrongdoing? In order to expand the scope of remorse, Deigh (1996) cites a more abstract ‘destruction of value’ as central to remorse. The need for a wider conception of what can appropriately evoke remorse is illustrated by the following scenario.

If, on a countryside walk, a person gratuitously stomps on, and destroys, the last flower of a particular species it may seem intelligible that he feel remorse over the irreversible destruction of something of value.⁶ It would be unsatisfactory to analyse the wrong in this case in terms of the harm done to others (although, no one can now enjoy this particular flower). Rather, the moral concern in the flower-stomper’s remorse is that something valuable⁷ has been irreversibly destroyed. It may be that the locus of this value is most often other people and that remorse felt over the harm done to others begets the most intense and authentic instances of remorse. The other possibility is that instead of remorse, the flower-stomper should feel guilt. It will be argued below that some instances of guilt and remorse are hard to differentiate and may come down to linguistic preference and disagreement.

There is some disagreement about the objects of remorse. For Taylor, the object of remorse is the deed and, explicitly, *not* the ‘agent as he who has done the deed’.⁸ She thus denies that remorse is an emotion of self-assessment. Similarly,

⁶ This example occurs in Thomas (1999)

⁷ Diversity in nature? The beauty of this particular species?

⁸ Taylor intends ‘deed’ to cover the other who has been harmed. This must be the case as Taylor relies on the ‘other-regarding’ nature of remorse to support her argument that remorse is more moral than guilt or shame. However, as our discussion of guilt will show, ‘deed’ can be separated from the harm to the victim as an object of attention, resulting in consequences for the moral value of the emotion.

Dilman (1999) argues that in remorse the pain of what one has done to others is foremost in consciousness, so there is no room for thoughts about oneself. Tudor (2001: 127), on the other hand, argues for two objects of remorse: first, the 'Other whom one has wronged' and 'the self as the one who has wronged the Other'. Importantly, these are not distinct objects of remorse but 'two aspects of one remorseful mode of attention'. Tudor argues that self-reflectiveness is vital to remorse otherwise what the agent feels is not remorse but a distorted form of compassion. Since an awareness of the self's part in the other's suffering is crucial to remorse, it would seem that (contra Taylor 1985: 1) the self's 'standing in the world' must be judged as having altered, thus rendering remorse an emotion of self-assessment.⁹

Tudor's analysis coheres with Gaita's (1991; 1999) argument that the fact that *I* did this is internal to the understanding of what was done, of which remorse is the shocked realisation. Thus, one does not fully understand what the victim has gone through unless one is aware of oneself as crucial to the description. However, he is keen to remind that, although an object of attention in remorse, the victim is the primary focus: 'In remorse, we respond to what it means to wrong another, which involves a new and terrible shock at their reality. Far from being intrinsically self-indulgent, lucid remorse makes one's victim vividly real' (1999: 102).

There is also disagreement with respect to the actions motivated by remorse. In addition to Gaita, many argue that the remorseful person is motivated to offer reparation (Thalberg 1963; Taylor 1985; Tudor 2001). Tudor emphasises that this reparation is not simply a way of discharging a debt. Instead, the remorseful response is 'communicative'. It symbolises a deep concern for the victim and an attempt to redeem oneself through confession, apology and penance. However, others have

⁹ This alteration of the self's standing in the world is the 'identificatory belief' of someone experiencing an emotion of self-assessment, according to Taylor (1985: 1).

argued that since remorse is only felt over the irredeemable, one does not seek to repair, knowing that this is futile. Instead, akin to an experience of grief, one is paralysed (Deigh 1996; Rosthal 1967).

However, we have already acknowledged that there is a significant symbolic element to the reparative behaviour of the remorseful and here it might be useful to draw a distinction between material and moral reparation (confession, apology, penance etc.) – the remorseful response involves the desire to offer both.¹⁰ Moral reparation may still be possible even if material reparation is not.

Moreover, even when material reparation is impossible (the harm was too great), one may still want to communicate one's desire to repair in this way. Indeed, as discussed above, Gaita (1999) believes that it is this communicative element, borne of remorse, that makes reparation reparative at all. This being said, it is possible that one might be overwhelmed by remorse. This is not incompatible with wishing one could make good and I would argue with Gaita that a complete absence of desire to repair (even if one has no idea how to act on this desire) would make the presence of genuine remorse debatable.

Also essential to genuine remorse is the disposition to act differently in the future (Thalberg 1963).¹¹ The prevalence of this disposition in the remorseful has been empirically substantiated (Shaw 1988).

It has been suggested that in genuine remorse we find 'a person prepared to react to infliction of pain and forgiveness with acceptance and absence of resentment and defiance' (Morris 1971: 432). This echoes Smith's description of the individual

¹⁰ Cordner (2007: 353) draws this distinction between the reparation one makes that is 'expressive of one's sorrow' and reparation that is 'simply the equivalent of paying a fine to the victim'. He argues that the former is required for remorse.

¹¹ Except, perhaps, in cases of 'tragic-remorse' in which the action was the lesser of two evils and thereby 'unavoidable' for the agent (see De Wijze 2004).

who acknowledges that the consequences of his actions have ‘rendered him the proper object of the resentment and indignation of mankind, and of what is the natural consequence of resentment, vengeance and punishment.’ In remorse one does not fear the person he has wronged. Instead, the remorseful is motivated to turn to the victim and try to ascertain how to repair (Tudor 2001). These suggestions characterise the attitude one has in remorse and the way one is inclined to respond to others’ dispositions towards one.

Remorse seems to have moral value: the remorseful humbly accepts responsibility, demonstrates deep concern for the victim, is disposed to repair and redeem, condemns oneself as the one who has, and is capable of, harming the victim, and resolves not to inflict such harm again. In addition, moral value seems to reside in the more abstract ontological realisation of the reality of another and the relationship the self can have to this reality (Tudor 2001; Teichman 1973; Gaita 1999). Again, as Gaita (1999: 34) puts it: our sense of the reality of others is partially ‘conditioned by our shock and bewildered realization of what it means to wrong them. Remorse is that realisation’.

Shame

Shame most basically involves the feeling of being exposed, of becoming diminished or dishonourable, or in some sense of being at a disadvantage (Williams 1993; Tudor 2001). This is thought to require some notion of an audience – of being held in the gaze of another (Williams 1993; Taylor 1985). Roberts (2003) classifies shame as an emotion of ‘defect’ and it is widely accepted that in shame a global self-assessment is made that one is flawed or has not lived up to expected standards (Dilman 1999). That

one makes such a negative global self-assessment is substantiated by empirical research (e.g. Niedenthal *et al.* 1994).

In shame, the primary focus is on the self, but also on the defects, flaws or weaknesses one perceives there. Experiences of shame can have triggers that are non-moral in nature, such as bodily functions, one's appearance, or one's emotions (Scheff, 1995). It has also been suggested that it is possible to feel ashamed of one's involuntary actions as well as the actions of others (Thalberg, 1963). Whilst I would agree with the former of Thalberg's two suggestions, I would argue that in some – if not all cases – what precipitates one's shame in the latter cases of 'vicarious' shame are not the actions of others *per se* (for example, my child's behaviour) but rather it is one's association with, or responsibility for, the others' actions that occasions a feeling of shame. I feel ashamed because of what my connection says about me, or simply because I am connected. However, I shall not dwell on these sorts of instances of shame, because what I intend to do in this section is characterise experiences of shame precipitated by moral wrong doing.

There is general agreement in the literature that in an experience of shame the focus is on the self and that the notion of exposure is central. However, I have identified three main areas of disagreement about shame in the literature. Importantly, the conclusions one comes to in these areas have implications for the moral value and character of shame. I shall therefore consider these disagreements in detail and take – and argue for – a particular position on each issue.

The three points of disagreement are: first, what role, if any, an 'audience' has to play in an instance of shame and what the exact nature of this 'audience' is; second, whether the agent has to share the view of the audience, both in terms of the audience's description of the agent's situation and the evaluative judgement the

audience makes about this situation; third, whether it is intelligible to expect positive change from the agent feeling shame and, if so, whether this is likely.

The 'audience' in shame

Shame is often thought to involve the agent being held in the gaze of another. He feels ashamed because another person has seen him do something that reveals a flaw or weakness in his self. In being diminished in the audience's eyes, he is diminished in his own eyes. However, it is widely agreed that it is not necessary for an actual audience to be present to occasion shame. Indeed, empirical research has found that participants recalled experiencing shame when they were alone (Tangney *et al.* 1996).

However, I wish to argue that the suggestion that some notion of audience is required is often overstated. Instead, I argue, the appropriate model should not make central the concept of audience, but rather should allow for mutual observation – a kind of comparative context in which one might measure one's relative worth against one's fellow human beings, as well as attending to their actual or potential measuring of one. Thus, I agree with Morris (1976: 61) when he says that 'the whole focus in a shame scheme where relationships are valued is the question, "Am I worthy of being related to the other"'. Not dissimilarly, Dilman (1999: 320-1) argues that in shame 'one has alienated oneself from the moral community to which one belongs: one can no longer hold one's head high among them'. Thus, shame retains its fundamentally social character, without requiring an actual or imagined audience.

The focus perhaps should therefore not be so much on the agent as *observed* but rather on the agent being *prompted to self-reflect*. This prompting might be precipitated by the presence of an actual or imagined audience but alternatively might

involve the *agent* measuring himself up (morally) against others instead of attending to their actual or potential measuring of him.

My proposed revision of the notion of audience is supported by examples of shame in which the ashamed agent himself is more appropriately described as the audience of a more worthy other. For example, the agent might be observing another's generosity and suddenly become acutely aware of his own lack of generosity, precipitating shame. Deflating, comparative thoughts such as 'I never behave like that' initiate the self-judgement. Here, the agent is measuring himself up against the other who might be completely unaware of the agent and unperturbed by any lack of generosity in others.

Interestingly, in these cases where the other is observed by the agent it does not have to be the case that the agent, prior to his experience of shame, considered himself particularly generous.¹² In fact, the agent could be completely aware that he never gives of his time or money. Moreover he could be completely at ease with this fact. Unlike the agent in the first case, the reality of what he is *is* transparent to him but he does not think he need be any better. Such an individual may still come to feel shame on encountering the generous other.

The logic of such a possible experience of shame differs to that of the standard audience paradigm. Instead of it being revealed to the agent through taking another perspective that he is not what he assumed or hope to be, what is revealed to the agent is that what he thought he was content being he is no longer content being. His *standards* change rather than his perception of himself relative to these standards. Importantly, this alternative structure may well be instantiated in the experiences of some who commit serious moral wrongs. Particularly in cases of recidivism, it is

¹² This somewhat puts into question the often-held view, represented by Lynd (1999: 24), that the sense of degradation in shame is excited by the 'consciousness of having done something unworthy of one's previous idea of one's excellence'. In my example, there is no previous notion of excellent.

unlikely that the agent would be unaware of what his character is capable of; he just might not be bothered by it. What is possible is that in becoming aware of others' worthiness, he painfully re-assesses this capability as a flaw.

In both structures, the agent painfully discovers a flaw but in the former the agent *learns something about his character* which implies failure to meet his standards whereas in the latter the agent *learns that his standards are other than he thought* and thus his judgement of his character, in light of them, changes.

The perspective one takes on the role of the audience in shame is of consequence as it informs whether the agent can disagree with the audience's description of the situation and/or the negative judgement the 'audience' makes whilst still feeling shame. Shame must involve some perspective that is not one's own if one is to be in disagreement with it.

Whether the agent agrees with the negative judgement of his actual situation has implications for the moral nature of shame. If one can feel ashamed following wrongdoing without judging what one did to be wrong then shame will be less morally valuable than it would be otherwise. We shall now turn to the assessments made in shame.

The judgements in shame

I am going to argue that in order for the discomfort one feels to be shame, one must accept the *description of the situation*: one must really be doing what one is thought to be doing or be as one is thought to be: one must actually be stealing (rather than mistakenly be perceived to be stealing) or must actually display racist views (rather than having been slandered), for example. If one believes that the observer is mistaken, then what one might feel, I will argue, is embarrassment, but not shame. I

will argue that shame might be felt even if the agent does not agree with the *evaluation* the observer makes about the agent or his behaviour (stealing is wrong; racist views are abhorrent), but that this will be a sort of phantom shame that would not persist once the agent is unobserved.

Taylor (1985) argues that the agent must agree with both the description and a negative evaluation in order to feel shame. This follows from her position on ‘audience’, whereby all that is required is that the agent shifts perspective to that of a detached observer. Although the agent may have ‘discovered’ the fittingness of the description of his situation by taking a detached-observer view it is, no less, one that *he* now would apply to himself. It is essential that there be a ‘revelation’ to the agent of himself as ‘inferior to what he believed, assumed or hoped to be’ (ibid.: 68). If he did not accept the possible description, he would not experience a truth about himself as being revealed.

Harré (1990: 199) also argues that the agent must accept the description of her situation and the evaluation the audience makes about it, asserting that shame is ‘caused by the realisation that others have become aware that what one has been doing has been a moral infraction, a judgement with which I, as an actor, concur’.

Alternative positions are represented by arguments from Crozier (1998) and Calhoun (2004). Contra Taylor’s position, in which the agent must accept both the description of his situation *and* the negative evaluation made about it, Crozier argues that the agent need not accept the audience’s description of the situation (but must accept the negative evaluation that would have applied were the description to be true). Calhoun differs again, arguing that the agent does not have to accept the negative evaluation (but must accept the description).

The description of the situation

Crozier discusses examples of potentially-shame-inducing situations in order to make his point that the agent need not accept the audience's description of his situation. He presents as similar in all relevant respects the following three examples.

1. Castlefranchi and Poggi's (1990) example of a man who becomes aware of being observed by a passer-by whilst giving a woman mouth-to-mouth after saving her from drowning. He suggests that the man might feel shame if he realises it might appear that he were taking advantage of the woman.
2. His own example of being left in charge of a lady's bag which he looked through to check if it contained anything too valuable to allow him to safely leave it for a few moments. On her sudden return, he felt 'discomforted' at the thought that she might think he was looking for money or taking a 'prurient interest in her personal belongings'.
3. G. Taylor's (1985) example of a life-drawing model who experiences shame on realising that the artist no longer regards her as a model, but lustfully as a woman. (Importantly, it is a potential (second-order) observer of her as object-of-lust who might judge this situation negatively.)

Crozier uses these examples to try to argue that shame can occur when 'the actor behaves in good faith and there is no self-evaluation relative to standards' (1998: 276). However, I am going to argue that the first two examples would more realistically lead to embarrassment rather than shame and that the third, despite potentially precipitating shame, is importantly different.

I suggest that the first two examples would only precipitate shame if there was at least some truth in the possible description of mal-intent. Otherwise, it is more likely that embarrassment would be experienced, accompanied by a strong drive to

explain to the observer that he or she had misunderstood the intentions. The desire would be to convince the observer of the correct description of the situation. Indeed, Crozier uses the word 'discomfort' to describe the emotion experienced in his example. This discomfort is precipitated in him because 'in searching the contents of the bag, I am behaving exactly as someone would who is actually stealing, or worse, and *I cannot think of a way out of this predicament*' (ibid.: 279, emphasis added). I would suggest that it is this anxiety that is most prominent for the agent, not any deep shame.

Such an interpretation coheres with G. Taylor's account of embarrassment in which she argues that a demand for some response is central: 'The tension and confusion so typical of embarrassment are due to his seeing the situation as creating a demand to which he is unable to respond' (1985: 69). So, in being seen rifling through the lady's bag, a demand is imposed on the man to explain himself, to make sure his actions are not misperceived. Intuitively, if he had not been 'caught', or if he managed to convince the lady of his good intentions, then there would be no embarrassment (and certainly no shame), since the demand does not emerge or is dissolved. This nervous anxiety is importantly different to the shock at what is revealed to oneself about oneself in shame.

We might further distinguish embarrassment from shame by noting the more intermittent nature of embarrassment – embarrassment seems more related to situation (which can change frequently), shame to disposition (which does so much less often, if at all). Embarrassment *does* seem to more straightforwardly require an audience, explaining why it does not endure in the same way as shame often does. Once the audience disperses, embarrassment begins to fade.

Crozier's model example – the third of those above – is importantly different to the other two, but he does not acknowledge this. With regards this example, I agree that the model may experience shame. However, here, it is not the case that the description of the situation by an imagined observer is wrong. The naked model really is being looked at lustfully by the artist. Whereas Crozier's handbag guard would not be lying if he said he was not looking to steal money, the model would be lying if she said that the man was not looking at her lustfully. Indeed, it is her acceptance of the description of the situation (and evaluation that it is bad) that makes it a plausible instance of shame. Her motivation here would not be to defend an alternative description of the situation but, rather, to denounce any responsibility: 'I did not know that he was going to look at me lustfully – that was not something I anticipated'. (However, as soon as she is aware of the artist's mode of attention, barring coercion, she becomes somewhat complicit in her being observed lustfully if she allows the situation to persist.)

What the model example demonstrates is that there need not be a self-attribution of responsibility in order for shame to be experienced.¹³ What it does not demonstrate is that an agent can feel ashamed for something that he is not actually doing or being. Rather, what he might feel in such situations, as we saw, is embarrassment.

Whilst we have established that the agent must accept the description of his behaviour in order to experience shame, it still could be argued that he does not have to judge the behaviour as bad or wrong in order to experience shame. We shall now assess such an argument.

¹³ Although it is likely that the most intense experiences of shame will occur in people who do believe themselves to be in some sense responsible for what is revealed about them.

Accepting the negative evaluation

Calhoun (2004) argues that it is not necessary that the person experiencing shame accepts the negative evaluation of those before whom she is ashamed. Calhoun's position results from her argument that the power to shame is a function of the agent sharing a 'moral practice' with the 'shamer' (which results from co-participating in some form of social practice: academic; sporting; neighbourhood...) and recognising that the 'shamer's' opinion expresses a *representative viewpoint* within that practice. However, the agent need not endorse that view point. That Calhoun does require the agent to agree with the description of his situation is evident from her discussion of someone who has 'misinterpreted the facts' as having a gaze that 'lacks the power to shame' (ibid.:141).

If an agent is able to experience shame without accepting the judgement of the shamer, then this would put into question its moral worth. What would it tell us about the ashamed's evaluation of himself or behaviour if it could be aroused simply by others' disapproval? Perhaps the social nature of shame is powerful enough that it can trigger a sort of phantom shame even when one does not believe that one has done anything shameful. However, we have discussed that central instances of shame involve a reassessment of one's worth. It is unlikely that unless one at least suspected that there was some truth to the 'representative view point' that any deep running reassessment of one's identity would take place. Most likely, the phantom shame would dissipate once unobserved. When one believes that one has done something that makes one less worthy in relation to others, however, one will likely continue to feel the diminishing experience of shame even if one isolates oneself from others.

Consequences

If an offender could feel shame without judging himself negatively or even accepting the description of his actions as correct then we might question what value shame has other than its showing that the agent has some sort of awareness of and unease at others' displeasure at him. However, given that shame can occur in scenarios that do not involve wrongdoing, we must be careful not to define morality into shame. Instead, we might emphasise that characteristics such as accepting that one is legitimately seen as having done wrong increase the moral praiseworthiness of particular instances of shame. When one (appropriately) judges oneself negatively in shame – as has been argued is the case in any deep experience of shame – value can be found in the agent's aspiring to standards and in his awareness of himself as a moral agent with variable worth. What such aspiration and awareness might lead to is considered next.

The hope for positive change

Ideas concerning the capacity that shame has for motivating positive change in the agent are conflicting. G. Taylor (1985) has a pessimistic view on the cognitive and behavioural consequences of a person's shame, although this is not stated explicitly. She describes the experience of shame as revealing to the agent who he really is, as if 'who he really is' is an unchanging fact that was made known to the agent through his awareness of his behaviour. The person feeling shame about the way he lives his life '[sees] it as being just the sort of life a person of [his] sort would lead' (ibid.: 92).

Taylor invokes the sociologists' concept of secondary deviance to further explain the potential self-perception of the ashamed. In secondary deviance, the person commits offences because he self-identifies as, say, a burglar or vandal. This

self-perception is different from that in primary deviance, which applies to cases in which the person accepts that he has done wrong – has committed a burglary or vandalised a property – but does not see this as affecting who he really is. Indeed, his behaviour may seem alien to him.

This consequent perception of determinism – identity dictates behaviour – makes intelligible Taylor’s argument that repayment (conceived broadly as ‘making up for’ the wrong) is not an appropriate expectation of the ashamed. She puts the question thus: ‘...how can I possibly make up for what I now see I am? There are no steps that suggest themselves here. There is nothing to be done...’ (ibid.: 90).

For Taylor, the only options for the person experiencing shame (all these being ways of ridding oneself of shame) are realizing that one does not actually accept the standards by which one was judged, changing one’s standards, or conceiving of one’s ideal standards as unattainable ideals and thus not as an appropriate measuring stick. However, these conclusions rely on Taylor’s supposition that the ashamed perceives her behaviour as determined by her unchangeable character. There are two beliefs here: first, that we have no control over the *expression* of our natural impulses and second, that these natural impulses themselves – our character traits – are entirely resistant to adjustment. Both of these beliefs can be challenged, particularly the former.

I would argue that there can be a legitimate sense in which the ashamed can try to ‘make up for’ the behaviour resulting from ‘who he is’. We cannot change our genetic makeup but, through self-awareness, self-discipline and perhaps behavioural therapy we can cultivate less shameful character traits and certainly control better their expression. Indeed, I think that embarking on such ‘self-improvement’ would be

a natural response for many instances of shame, and it certainly would not be inappropriate for others to expect this sort of commitment from the ashamed.

For example, we can imagine a case in which a person becomes violent towards his partner. On one occasion he catches himself in the mirror during a violent episode. In seeing himself, the reality of what he is doing suddenly strikes him and he is overwhelmed by shame: 'I am an awful person capable of terrible things'. According to Taylor's conception of shame, the man would then think, 'but that's who I am and I can't do anything about who I am.' I suggest that this is not the only response shame might precipitate. Rather, on having what he is capable of revealed to him, he might become determined to change. He might seek out an anger-management course. He would resolve to become a better person.

Whether the above possibility would constitute 'making up for' what he revealed himself to be in the violent episodes is mostly a matter of what one means by 'making up for'. But, the important point is this: it seems likely that in instances of shame there may be motivation to *do something*: to try to be a better person and rectify one's shameful flaws. This contrasts with Taylor's account.

In line with my view, Tudor (2001) suggests that the person shamed into exile has to work at *changing himself*, at proving himself *worthy* in order to be let back in. Similarly, Dilman (1999) suggests that 'The response of shame which presupposes such a split within oneself, however, is at the same time the first step in bridging that split, in coming together and *regaining self-mastery*'. (p.318, emphasis added). However, Dilman's notion of regaining control is a selfish concern, focussed on internal integrity.

Williams (1993), on the other hand, incorporates both the notion of internal integrity and one's concern with one's worth when he talks of shame facilitating the

‘rebuilding of the self’. He says that shame is suited to this function as it ‘embodies conceptions of what one is and of how one is related to others’...‘*shame may be expressed in attempts to reconstruct or improve oneself*’ (p94).

However, that shame often results in less honourable behaviour is supported by empirical research. Shame may fuel the defensive blaming of others and motivate hostility towards them or oneself (Lewis 1971; Nathanson 1992; Tangney 1999). It may also inhibit empathy for others through excessive negative focus on oneself (Tangney 1991). Shame-proneness was found to be positively correlated with irritability, suspiciousness, resentment, anger arousal, and externalisation of blame (Tangney *et al.* 1992). Interviews conducted by Lindsey-Hartz (1984) revealed that the desire to hide or escape was characteristic of shame.

Despite evidence for these defensive reactions in shame, it is not illogical to hope for reparative action in some sense from the ashamed. Importantly, I would point out, hiding away and resolving to change are not mutually exclusive. Indeed, one might wish to hide away until one has rebuilt oneself in such a way that one again feels worthy of being related to others. Repairing *oneself*, however, seems to be the likely function of any ‘reparative’ action.¹⁴

Shame and remorse

In shame one is focused on oneself and one’s negative assessment of one’s very worth in relation to humanity. In remorse, one is focused primarily on the other whom one has wronged, but also on the self as the one who wronged the other. Taylor (1985) cites shame’s not being other-focused as reason to see shame as selfish and as less moral than remorse. However, as Tudor (2001: 185) notes, shame need not wholly

¹⁴ Although, if one repairs oneself *for* someone else, then this might count as other-focused reparation.

ignore victims if there are victims, but it does not attend to them in the direct and fundamental way needed for ‘adequate acknowledgement of their moral reality’. Remorse does attend to them in this way.

Despite Taylor’s (1985: 77) evaluation of shame as less moral than remorse, she allows that it is rightly described as moral if morality is taken to ‘include personal morality, a person’s own view of how he ought to live and what he ought to be’. Thus, we might say that shame precipitated by wrongdoing involves awareness of one’s *own* moral reality, where as remorse primarily involves awareness of the moral reality of *others*. Remorse, however, does not neglect the self’s relation to this reality. Recall Gaita (1999: 34): Remorse is ‘the perspective in which the meaning of what one has done, what one has become through doing it, and what one’s victims have suffered through doing it, are inseparable’.

More generally, it might be more appropriate to feel shame than remorse for simple deficiency in virtue, when no harm has been done. Malicious thoughts or drunken idiocy, for example, might more appropriately precipitate shame, where the focus is on one’s own flaws, not harm done to others. However, given that criminal offences often do involve harm to others, for our purposes we should emphasise that remorse responds more thoroughly to this sort of wrongdoing – attending to the reality of the harmed victim as well as one’s capacity to harm.

Acceptance of responsibility and of one’s status as blameworthy is crucial for remorse but less so for shame. This has implications for the attitudes towards others exhibited: the remorseful displays humble acceptance whereas the ashamed (who has not accepted the description of his situation or behaviour) can display hostility and resentment. However, I argued that acceptance of responsibility would make the experience of shame deeper and, where not misplaced, more valuable.

The action tendencies are also different: whereas one is disposed to make reparation in remorse, one might hide away in shame. Where some sort of positive activity is motivated in shame, this is likely to be focussed on rebuilding the self rather than the motivation to reduce the harm caused and to redeem the self through apology and penance in remorse.

Guilt feelings

Discussions of guilt in the literature seem to conform to two approaches. Either guilt is not distinguished from remorse – the terms are used interchangeably – or guilt is sharply contrasted with remorse by those who see the former approach as inappropriately conflating two quite distinct emotions. I am going to argue that although there may be some differences to be teased out between the two emotions, there is more overlap than the latter approach allows for.

Where guilt is seen as entirely distinct from remorse, it is construed as an emotional response that focuses on rule violation and predominantly comprises fear at the real, imagined or internalised anger of an ‘enforcer’. Most abstractly, it is rules and commands that are salient in experiences of guilt (Tudor 2001). Psychoanalytic theories – most commonly that of Freud – are appealed to in order to explain the process by which external enforcement can be internalised, and fear at the enforcer’s anger becomes fear of justified retribution for wrongdoing (Thomas 1999). Words such as ‘transgress’, ‘violation’, ‘forbidden’, ‘burden’ and ‘debt’ pervade these accounts of guilt feelings. They are often presented as being closely connected to a Kantian style of morality, concerned with abstract rules and principles.

In these sorts of accounts, the feelings seen to characterise guilt are fear, anxiety, self-pity, the dread of punishment, and the sense of having a heavy burden

imposed upon one (e.g. Tudor 2001; Williams 1993; Taylor 1985). Further, it is claimed that fear can lead to attempts to justify acts carried out or to deny responsibility. However, this denial is often motivated by the drive to protect oneself, rather than by a firmly held conviction that one is not to blame. Those feeling guilt have a painful awareness of responsibility for an act without necessarily openly and consciously acknowledging this responsibility (Gobodo-Madikizela 2002).

The language of burden and debt – the ‘conceptual scheme of obligations and entitlement’ (Morris 1976: 61) – might seem at odds with the notion of concerned repair and reparation. However, empirical literature has shown reparation to be part of guilt (Baumeister *et al.* 1994; Tangney 1990, 1995) and the ‘cynical’ guilt theorists do not deny this. However, they remain cynical as the claim is that any reparation is not well-motivated: in guilt, we are told, the agent thinks that he has put himself in a position where punishment or repayment is due, and his aim is solely to rid himself of this burden. As Williams (1993: 89-90) suggests: ‘what the offender may offer in order to turn this [the anger, resentment or indignation of others] away is reparation; he may also fear punishment or may inflict it upon himself’. Thus, what may appear as a concern to alleviate a victim’s suffering or redeem oneself is often a means of ridding oneself of a painful burden. In other words, reparation or penance are not ends in themselves in these accounts of guilt feelings, but are means to the end of reduced personal suffering.

Those who characterise guilt in this way are understandably critical of it. An emotional response in which the agent neglects the immoral and harmful quality of what he has done, sees the situation merely as rule violation, and experiences primarily fear for himself does not strike one as a morally mature response to wrongdoing. Selfishly-motivated reparation seems equally as suspect. However, I

would argue that, although the above characterisation does hold for some forms of guilt, it is not the complete picture. In fact, I would argue that the above depicts a particular, corrupt form of what otherwise can be a more morally valuable response to wrongdoing. In other discussions of guilt, we do see concern for the victim and genuine desires to repair for its own sake (e.g. Morris 1971; 1981; 1999, Dilman 1999; Murphy 1999).

It is not just academic commentary that often conceives of guilt in this broader way. The breadth of the range of experiences that can be identified as guilt is revealed in everyday discourse. We can take the following literary example as a basis for reflection: In Dickens' *David Copperfield* (2003), David Copperfield ferociously bites the hand of Mr. Murderstone (who is in charge of his lessons) when he holds David in place so that he can strike him:

He beat me then, as if he would have beaten me to death...How well I remember, when my smart and passion began to cool, how wicked I began to feel! I sat listening for a long while, but there was not a sound. I crawled up from the floor, and saw my face in the glass, so swollen, red, and ugly that it almost frightened me. My stripes were sore and stiff, and made me cry afreash, when I moved; but they were nothing to the guilt I felt. It lay heavier on my breast than if I had been a most atrocious criminal, I dare say. [2003: 59-60]

Here we have a self-report of guilt in which there is a deep concern with the 'wickedness' David perceives in himself. This is not, therefore, an instance of the corrupt (Dilman (1999) calls it 'persecutory') guilt in which the 'fear which dominates one's consciousness distracts one from appreciating the character or moral significance of what one has done' (Dilman 1999: 325). Indeed, David's 'punishment' is over, yet in the quiet of the aftermath, he comes to feel guilty. He is not angry or defiant: he perceives himself as wicked and this matters to him.

We also see cases of guilt in which the notion of breaking a rule does not even meaningfully apply, and deficiency in virtue is at issue instead. For instance, someone

might feel guilt over having not donated to charity for a long time. It seems more natural to say that such an individual feels guilty because he has not been charitable rather than that he feels guilty because he has violated the (personal?) rule that one should donate at least X pounds to charity every X months. Murphy's (1999) analysis of guilt does much to convince on this score. He presents further examples of paradigm cases where guilt is likely to arise, including the following from *Incline Our Hearts*, by A.N. Wilson (1990: 143–144):

It is only on those whom I have loved that I have ever knowingly inflicted pain. The guilt of it remains forever, my words selected with such malice and the startled expression on the victim's face as the effect went home. These are the faces which return during nights of insomnia, forever hurt in my memories, and inconsolably so . . . Sometimes in spells of profound depression, it is these moments alone which surface in the memory. Everything else is a bland, misty background against which these figures stand out sharp and clear – women in tears, or my uncle, drawing back the corner of his lips and sticking a pipe in his mouth, trying to conceal the extent to which I was hurting him. [Quoted in Murphy 1999: 332]

I agree with Murphy that to analyse such a case as this in terms of disobedience to authority or rule violation would be to radically distort it. Interestingly, the central element of what haunts the character seems to be, *à la* Gaita, the reality of the victims and their individuality: their hurt faces haunt him; everything is hazy except the sharp, clear figures replete with their idiosyncrasies. Murphy concludes that 'although I would not deny that disobedience of authoritative rules is an important element in some instances of guilt, I would not myself stress its centrality... What is central in many cases – and what tends to cause the most painful pangs of guilty conscience in morally mature persons – is not just wrong but *wrongful injury to others*' (1999: 335).

When considering examples such as the above and many instances of everyday discourse I would not want to argue that they demonstrate misunderstanding

of guilt or misinterpretation of emotional experiences. Instead, I would agree with Morris (1987: 221) when he says ‘I am sceptical about any claim of widespread misuse of terms for emotional states, and I am generally disposed to accept first-person reports as accurate’. Similarly, Williams (1993: 91): ‘what people’s ethical emotions are depends significantly on what they take them to be’. What Gaita may call remorse, Murphy may prefer to call guilt.

When guilt is portrayed by the ‘cynical’ theorists as only applying to a corrupt form, remorse is then frequently invoked as the morally-superior alternative. Thus, the cynical portrayals of guilt are perhaps more of a theoretical technique to ensure the superior value of remorse. In line with the socio-linguistic perspective introduced at the beginning of the chapter, it may be that there is no truth of the matter about exactly what guilt is: it may just be more or less what people take it to be.

Labels aside, what we do know for certain is that we do not find particularly valuable the response that focuses only on one’s disobedience in fear of another’s wrath. But, in acknowledging that it seems natural to describe some more virtuous responses as guilt, we do not thereby devalue what we have described as remorse. All we do is concede that it is not so easily dissociated from guilt. As Gaita (1999: 31) says: ‘If we understand guilt-feeling to be a pained acknowledgement of the wrong one has done, then there is no significant difference between guilt-feeling and remorse’.

Guilt feelings and remorse

Although attempts have been made to sharply distinguish remorse from guilt, I have argued that these attempts focus only on one specific, and arguably corrupt, form of guilt. Whilst it does seem appropriate to call the corruption guilt, and not remorse, this

should not lead us to conclude that guilt never involves anything morally praiseworthy. Indeed, guilt has aptly been conceived as sometimes describing the emotional state of an agent who ‘grieves the misery of which he has been the cause’ and ‘apprehends the wrong one has done painfully in one’s concern for the person one has wronged and for the values one has violated’ (Dilman 1999: 326).

That being said, we can imagine situations where someone might feel guilty but not remorseful, where what is felt and what is not felt are importantly different. We might imagine the animal tester who feels guilty about experimenting on monkeys but not remorseful. His staying in his job speaks to his not feeling remorse but is compatible with experiencing some guilt. Also instructive is that whilst ‘guilty pleasure’ is an intelligible concept, ‘remorseful pleasure’ is not. I might gain pleasure from the car I have stolen whilst also feeling some guilt. If I were to instead feel remorse, it is impossible that I would at the same time be able to enjoy the car.

Thus, guilt is much broader in its extension than remorse, incorporating much less desirable responses to wrongdoing. Guilt can apply to the emotional experiences of those who do not attend to the victim, who do not accept responsibility for what they have done, who do not take reparative steps, and who do not wish to undo their actions. Moreover, it can be a response to one’s bad *intentions*, whilst these do not seem an appropriate object for remorse. Guilt also seems fitting for minor transgressions, where remorse might be seen as an overreaction. What seems minimally sufficient for guilt is a certain anxiety in response to wrongdoing. This *may* be accompanied by morally-valuable aspects but need not be to nonetheless aptly be labelled guilt.

Remorse, although liable to corruption (self-indulgent remorse, for example), seems an inappropriate label once we have identified that there is a corruption. We

would be inclined to say the distorted emotional response ‘was not really remorse’, whereas we would still permit that corrupt guilt was, nonetheless, guilt.

The above positive moral aspects that guilt can do without cannot be done without in order for an emotional reaction to be remorse. Moreover, whilst anxiety seemed central to many instances of guilt, remorse seems more naturally to command sorrow as its principal affective component (cf. Dilman’s (1999: 326) discussion of ‘depressive guilt’, as continuous with remorse). Thus, there are elements necessary for remorse that are not necessary for guilt. However, as has been argued, just because these elements are not necessary for guilt, does not mean that they are always absent.

So, whilst remorse might naturally describe only the most morally valuable responses to wrongdoing, guilt’s extension to more corrupt variants does not preclude it applying appropriately to emotional responses with some moral value.

Regret

In an experience of regret ‘something unfortunate is construed in terms of what might have been’ (Roberts 2003: 240). Despite incorporating painful feelings (Landman 1993; Rorty 1980) regret also involves a high degree of cognitive appraisal, and thus is ‘an experience of felt-reason or reasoned-emotion’ (Landman 1993: 36). Roberts categorises regret as an emotion of ‘loss’, and this sense of ‘loss’, accompanied by frustration is seen to characterise the affective element of regret.

It is agreed in the literature that regret does not require personal responsibility, although it is liable to weigh more heavily when one is responsible. Bernard Williams’ (1981) well known distinctions between spectator- and agent-regret demonstrate such a supposition. Williams presents the case of a lorry driver and his passenger who, completely by accident, knock down a small child, causing his death.

Simple regret or spectator-regret involves the wish that something had not happened when one feels no responsibility (regardless of whether one was involved in some non-culpable sense). This is the regret of the passenger. Agent-regret is the thought that it matters, despite not being legally or morally culpable, that it was *I* who was at the wheel. Agent-regret is what the driver experiences. Whether agent-regret would fully capture the experience of a culpable lorry driver is questioned by Tudor (2000) who argues for a further distinction between remorseful agent-regret and non-remorseful agent-regret.

What is also generally agreed upon is that regret involves counterfactual thinking (Kahneman and Tversky 1982) in which, at the most abstract level, one believes the world to be less satisfactory than it might be, and imagines the way(s) in which it could have been, or could be, better. Regret can thus be described as an ‘if only...’ emotion (Solomon 1976). What writers have disagreed about is whether mental ‘undoing’ is a necessary feature of regret. Although empirically the wish to undo has been found to be an important characteristic of regret (Zeelenberg *et al.* 1998), it has been argued that it is not a necessary feature (Taylor 1985; Rorty 1980). G. Taylor suggests that it is possible to ‘regret an action but accept it as the thing to do under the same description’ (1985: 99). Further: ‘perhaps regret always implies acceptance of what has been done’.

What she means by ‘acceptance’ is not quite clear. ‘Acceptance’ could mean that the decision made or action taken was, in the agent’s mind, the correct course of action. Or it could simply mean that the agent acknowledges that a particular path has been taken, and that this precludes any alternative path: a choice has been made. That the latter must be true for regret seems obvious. If the agent does not think of the decision or action as in the past and irreversible then he is unlikely to experience

regret since a less-than-satisfactory possible world has not yet been realised. Things could still go one of many ways. So, in the sense that the agent must perceive the action or event to be a fact about what happened, he must accept what has been done. He must come to terms with some truth of the matter.

However, that the agent must perceive the regrettable course of action as the correct or only course of action is not clear. I would argue that although it is possible that one can regret something that one would not change (for example, when an action is the lesser of two evils), this sort of ‘acceptance’ is certainly not characteristic of regret. Moreover, I would argue that at some level, regret always involves a wish to undo in the sense that one would at least undo the circumstances that necessitated a regrettable action if one could. In this sense, one always wishes things were otherwise – would undo if one could – even though undoing on such a cosmic scale is impossible. Indeed, Solomon (1983) defines regret as the response to circumstances beyond the agent’s control.¹⁵ This does not, however, mean that the agent accepts the consequences of these circumstances in any positive sense.

Ameliorative action may be motivated by regret (Rorty 1980). However, this will depend on the impact the agent believes he can have. Attempts may be made to ‘undo’ the action or to improve things overall, given that the action has occurred. Notably, such amelioration may often be intended to improve things for the agent in situations in which no one else is affected, or even despite others being affected (Gilovich and Medvec 1995).

¹⁵ Although such regret is possible, I do not think it is the paradigmatic case. The features of regret such as feeling like kicking oneself, feeling like correcting one’s mistake and wanting a second chance supported by empirical studies (Roseman *et al.* 1994; Zeelenberg *et al.* 1998) suggest that agents tend to berate themselves precisely because they could have done things differently.

Regret and remorse

The possible objects of regret are wide-ranging. Most broadly, regret is felt over an unsatisfactory state of affairs. Regret, unlike remorse, can be felt over morally neutral actions, even morally virtuous actions, and for circumstances that are not a consequence of one's own actions. Regret, unlike remorse, does not involve judgements about the self and does not have to involve concern for others. In regret there might only be concern for one's actions in so far as they affected future personal gratification.

What regret and remorse do share is the presence of an intention not to behave in the same way in the future. However, whilst the motivation for this in regret might be that it 'did not work out well for me last time', a firm resolve not to commit the same act again in remorse is motivated by a concern to avoid causing harm, particularly to other human beings.

Since regret involves a desire for things to be different and, commonly, the desire to undo a particular action, remorse might be said to involve regret. Indeed, both Tudor (2001) and Dilman (1999) say that remorse involves a 'bitter and deeply painful' or 'acute' sense of regret.

Clarifying remorse

We can now return to remorse to clarify its essential features and defend as plausible the proposition that remorse is the emotional response relevant to sentencing. Recall the characteristics I introduced towards the beginning of the chapter. First, the agent must perceive himself to have done something that he considers to be morally wrong. Most often this involves having harmed another person in some way. Second, the agent must assume responsibility, humbly seeing himself as blameworthy. Third, the

focus of the agent's attention is on the reality of the individual he has wronged as well as on himself as the one who wronged the other. Fourth, there is a desire to atone, which is likely to manifest itself in apology and reparative activity. Fifth, there is at least an intention not to repeat the harmful action. Having explored shame, guilt and regret, we can now clearly explain why this set of characteristics is unique to remorse.

Acceptance of responsibility for having done something morally wrong (characteristics one and two) is unique to remorse. One might feel shame over something for which one was not responsible, although responsibility is likely to invoke stronger shame. This 'something' does not have to be morally wrong. In guilt one perceives that one has done wrong but not necessarily morally wrong. One may defensively try to deny responsibility. In regret no perception of having done wrong need be present other than the potential perception of a less-than-optimum-action as 'wrong' in some prudential sense. These differences in acceptance of responsibility lead to different dispositions. In remorse the agent humbly accepts blame and punishment without defiance. In contrast, both shame and guilt have been shown to have the potential to lead to hostile dispositions.

Remorse is the only of the four emotions to focus on the victim of the wrong doing – on the harm caused to him (characteristic three). In shame one focuses mainly on the self (although this focus may be honourable in its concern with being a worthy human being and involves one's relation to humanity). In guilt one focuses on the self as one who is likely to receive punishment or is obliged to pay a debt. If present, the explicitly outward-looking aspects of shame and guilt are towards the disappointed/disgusted other and the angry other, respectively. In regret one might only be concerned with one's personal gratification.

Remorse is the only of the four emotions in which reparative desire, as an end in itself (characteristic four), must be present. In shame one might be motivated to improve oneself, although this is not a necessary feature. In guilt one might use reparation as a way of discharging one's debt but, again, this is not essential to guilt. Regret may lead to ameliorative action, but in some cases this is aimed only at making things better for the agent himself.

Remorse and regret are the only of the four emotions to essentially involve the intention not to repeat the action (characteristic five). Where there is resolve in shame, guilt or regret, this may (but not necessarily) be so as to avoid experiencing the emotion and the personal consequences in the future. In remorse, although one may want to avoid having cause for feeling remorseful in the future, the intention has to be more outwardly focused on wanting to insure no one else suffers as a result of one's actions and will also involve commitment to moral values for their own sake.

It would therefore seem that remorse is at least as morally valuable as shame, guilt and regret, if not more so. Being so, we are justified in focusing on remorse as a response to wrongdoing worthy of attention at sentencing. However, I wish to raise a question that will be of importance to considerations of the actual emotional responses of offenders. This question is whether reality consistently matches theory.

Return to social construction and the reality of emotional experience

It could be argued that actual experiences of emotions are not as 'neat' as the analysis in this chapter has suggested. Indeed, it might be the case that these emotions can overlap and that, even if they do not, it is difficult to tell which of them is being experienced, not least for the offender himself. What this analysis has provided might

be described as prototypes of these emotions – a depiction of what they consist in when they are most developed and at their most different from each other.

Hoyt's (1983) psychoanalytic research suggests that remorseful people experience cycles of both shame and guilt. Elsewhere in the empirical literature, research has provided evidence that a sense of shame underlies remorse (Shabad 1989) and that regret and guilt feature in remorse (Stern, 1989; Hauck 1989). In the theoretical literature, Murphy (1999) has argued that shame can 'creep through' guilt, generating the archetypal moral response to wrongdoing. He argues that shame is required in order for guilt to become sufficiently self-punishing to deserve characterization as a full-blown bad conscience. This is because it is not just rule violation and harm that play a role in a full account of the sufferings of bad conscience, but there is a vital role played by the pain of negative assessment of one's very worth.

Further, if we return to Smith's description presented at the beginning of the chapter we see that he evokes shame and regret. He also describes fear of the sentiments of others and of punishment, which seemed of questionable value in some forms of guilt. Yet, we would not, I think, see this as devaluing the response of Smith's remorseful person overall. Further, emotions are not a matter of all-or-nothing: there might be some intelligible concept of partial remorse where the agent experiences some of the feelings, has some of the thoughts, and so on. We may still find value in such partial emotional responses.

As I suggested at the beginning, we should not expect emotional experience to be as neatly divisible as emotion labels might suggest. Of course the labels refer to and pick out important differences amongst our reactions to wrongdoing, but these reactions can be as sprawling as that described by Smith and substantiated in

psychological research. Further, elements of shame were seen to be integral to remorse ('How could *I* have done this?') and some of the more morally mature accounts of guilt were seen to be indistinguishable from accounts of remorse.

Perhaps there might be some offences for which remorse, as I have characterised it (as the all-consuming horror at what one has done to another), is not necessarily the response we should anticipate. Fairly non-serious offences, like the theft of a small amount of money, might be more likely to engender a more proportionate sense of guilt. Shame might be the response to expect of the person charged with being drunk and disorderly in a public place - not remorse. These responses in these situations still suggest an understanding of what one has done.

Why Remorse?

Why remorse, then? I wish to propose a hypothesis, following from the socio-linguistic perspective considered at the beginning of the chapter. It may be that the theoretical concept of remorse involves all the aspects that are valuable from regret, guilt and shame, brought together in a model of the ideal response to wrongdoing. In finding the co-instantiation of these aspects valuable, society wanted a label to attach to the response.¹⁶ It is also plausible that the long-standing interest of the court in 'remorse' has had some effect on what it means: If the court is willing to place significance in this response to wrongdoing, then it must be the most morally desirable response.

If such a hypothesis were correct, the question 'what is remorse' might be subverted to become 'what should remorse be?' My hunch is that the two would more

¹⁶ This may explain why it seems unnatural to retain the label 'remorse' if the actual response has been revealed as corrupt (e.g. when better construed as self-pity): an emotional experience is only remorse when it constitutes the ideal response to wrongdoing. (*Misplaced* remorse is a meaningful concept but, even in these instances, the same underlying 'ideal' moral structure will be present.)

or less match up (the concept of remorse has been around for a long time) but, given the interest in peoples' responses to their wrongdoing, it would make sense that we have some way of referring to that which constitutes the optimum.

The veracity of this hypothesis is not crucial to my project as it does not hinge on something that is called remorse being exactly one particular way, but rather on there being a particular set of feelings, cognitions and behaviours that can be argued to justify mitigation. Remorse is just a label, albeit (it has been argued) a fitting one. Given that the reality of emotions is that they can elude easy categorization, we might consider what is sufficient for an emotional response worthy of judicial consideration. What we have done in defining remorse above is to present the idealised reaction to wrongdoing. From this we can extract the most worthy features. It seemed that (where appropriate) the focus on, and awakening to, the victim was central to the value of remorse. The offender's humble acceptance of responsibility and blame also seems particularly desirable, as do voluntary attempts at reparation, apology, and resolve to avoid offending in the future. Even if the conceptions of the emotions discussed have been somewhat stipulative, it is still the case that the set of characteristics of remorse that we have delineated are instantiated in some people's reactions to wrongdoing. We are therefore justified in using the term 'remorse' to denote such reactions.

Final thoughts

One might worry about what we lose from the concept of remorse when we strip it down in this way. Indeed, Gaita (1991: 59) warns that 'to describe remorse as a feeling or attitude to the fact that we did what we (morally) ought not to have done, harnessed to a resolve not to do it again, is seriously to underdescribe it.' He emphasises that, in addition, remorse is a disciplined remembrance of the moral

significance of what we did – the moral significance involving what it means for *me* to have harmed another, in all his individuality. The transformative nature of remorse cannot, however, be captured easily in a list of criteria which are intended to add up to justify mitigation of punishment. This chapter has been an attempt to understand remorse in all its depth and complexity, and listing concrete markers of remorse is a pragmatic necessity if it is to be identified and taken into account in a judicial context.

Having explored the concept of remorse, and acknowledged that its instantiations may not always be categorical, I will now turn to the task of further defending the view that remorse is the appropriate response to wrongdoing.

Chapter Three

The appropriateness of remorse as a response to moral wrongdoing

Remorse is often regarded as an emotion that we should value (particularly morally so) and encourage when experienced in appropriate circumstances. It is seen as a fitting – or even as the ideal – response, on the part of the wrongdoer, to his moral wrongdoing. Adam Smith (2002) describes remorse as the safeguard of society. Thomson (1989) claims that remorse is *the* factor that differentiates moral beings from people who are motivated only by reward and punishment. In this way, remorse is the hallmark of morality. The conceptualisation of remorse which emerged from Chapter Two confers *prima facie* plausibility on the idea that the appropriate response to one's moral wrongdoing is the experience of remorse. However, not everyone agrees that a remorseful response is what we should hope for from the wrongdoer. We therefore need to examine in depth the question of whether remorse is an appropriate response to moral wrongdoing, and assess the arguments attempting to discount the value of remorse.

Overview

I will begin by drawing a distinction between two ways in which an emotion can be 'appropriate' – a distinction that will be of importance throughout the chapter. I will then present a general account of why some emotions might be considered appropriate or inappropriate, drawing on Aristotle's virtue ethical approach. Having argued that emotions, in addition to actions, can be a locus of moral concern, I will move on to concentrating on remorse in particular and to addressing objections to its appropriateness. Focusing particularly on Bittner's (1992) claim that remorse is not

‘reasonable’, I will show how he considers only the instrumental *consequences* of experiencing remorse, to the exclusion of the *value* it might have in itself. I will also show how his confusion of the two types of appropriateness results in his argument having implications undesirable even for him. Finally, I will address an objection concerning the fleeting nature of our emotions and in so doing further explicate the nature of the appropriate emotional response to wrongdoing.

In considering the appropriateness of the remorseful emotional response, I am not denying that the *behaviour* that remorse motivates (apology, reparation etc.) might also have significance for the punishment the wrongdoer should receive. Indeed, the following chapters will consider such arguments. However, I wish to begin by making a case for the appropriateness of remorse as an emotional response to wrongdoing so that we can see what value this response might have *per se*, before we consider the more instrumental benefits resulting from its being experienced.

An appropriate distinction

We must begin by being clear about what is meant by ‘appropriate’. When we ask about the appropriateness of an emotion, we might be asking one of two things. Firstly, we might be simply asking if the emotion is a coherent response to the objects of its attention. That is, whether it has presented facts about the world to the agent correctly. Secondly, we might engage in value judgement when we ask whether an emotion is appropriate: we are then asking if it is prudential, or ‘right’, or ‘virtuous’ (we can make instrumental judgements and moral judgements in our evaluations).¹⁷

¹⁷ There is also a meaning of ‘appropriate’ that has to do with social propriety. This may overlap somewhat with the prudential appropriateness of *feeling* the emotion (experiencing the emotion might facilitate socially proper conduct), but achieving propriety results entirely from suitable *public display* of the emotion. Assessment of the propriety of public acts involves no consideration of whether any emotion they are thought to display is actually being felt. As long as certain words are said or actions

The two questions are importantly distinct. Williams (1973: 225) draws this distinction when he distinguishes between questions of ‘fact’ and ‘value’ in relation to emotions saying, ‘there are natural and indeed logical limits to the range of what objects a given emotion can take, and [limits to] what emotions a human being is expected to feel or, alternatively, to dispense with’.

An emotion can be coherent (logical) whilst being judged morally wrong or imprudent. For example, if jealousy is roughly a resentful and painful desire for another's advantages (material or immaterial) then we might say that John's jealousy of his wife's promotion above him is coherent – i.e. she really does have some plausible advantage: a superior job. But we might also assess his jealousy as wrong, particularly in the context of their marital relationship. If fear is roughly intense anxiety about the plausible negative effect something could have on one's wellbeing (broadly conceived), Jack's fear of his finals exams would be coherent. (The exams can affect his self-perception and other's perceptions of him; they can entirely determine his future work opportunities etc.) However, his friend would not be denying the fearsomeness of the exams by suggesting that it would be wise to try to abate his fear, else he will not be able to concentrate in the exam room. In this case fear, although coherent, is not expedient.

It seems that an emotion's simply being coherent does not necessitate any obligation to feel it, and this is shown to be the case by the possibility of coherent but wrongful or inexpedient emotions. It is also shown if we consider neutral scenarios: for example, just because it would be coherent to feel disappointed about missing the film at the cinema, this does not mean that one *should* feel disappointed. No one will blame you if you do not. Just because remorse might be a coherent response to

performed, propriety can be satisfied. In this chapter we are concerned with the actual experience of remorse, so social propriety will not be discussed here.

wrongdoing this does not yet speak to its appropriateness in terms of rightfulness or expedience.

This view does involve a commitment. It requires that we hold that emotions can be evaluated as coherent or not. However, from a brief consideration of what the alternative position would look like, we can see that this commitment seems attractive and plausible. If we were to deny that emotions could be evaluated on grounds of their coherence, we would be left in a situation in which no instance of an emotion could ever be deemed irrational. A person's mere experience would validate it: this is scary because I am scared. Fear of trees would not be irrational (despite trees not plausibly affecting our wellbeing); regret over things not happened would not be irrational (the logic of regret requires that the object regretted is in the past); jealousy over something one already possesses would not be irrational (when actually it requires that one be at a disadvantage with respect to the object of the jealousy). In addition to accepting that no instance of emotion was ever irrational, we would have to give up concepts depicting objects or persons or situations as objectively worthy of some particular emotional response, since this leads to the possible assessment of coherence in subjective emotional responses. For example, academic plagiarism could not be shameful *per se*; people could merely experience shame in response. Great injustice could not be enraging *per se*, but people might happen to feel indignation when faced with it.

I argue that such results make the view that we cannot assess the logic, or coherence of emotional experience very unattractive. In saying that we can assess coherence we are not denying that irrational emotions can feel just the same as rational ones – phobias are often *more* intense than rational fears. We are, however, claiming that we can assess an emotion's appropriateness against its internal logic: is

there really a threat to my wellbeing? Has the undesirable outcome actually been borne out? Do I really want what he's got? ¹⁸

From the outset, then, we must be careful to be mindful of the distinction between an emotion being appropriate in the sense of its coherence, and being appropriate in the sense of being judged as what one should feel – for moral or prudential reasons. To infer from an emotion's wrongfulness that it is not coherent would be to commit what D'Arms and Jacobson (2000) call the 'moralistic fallacy'. Equally, to infer that an emotion must be what one should feel, on the grounds that it is a coherent response to one's situation is also fallacious.

In this chapter, I wish to argue that remorse is not only a coherent response to wrongdoing, but that it is also, in many cases, what one should feel – it is appropriate in both senses. Remorse does seem to be a coherent response to significant moral wrongdoing. Examples of incoherent remorse could be: remorse over something morally neutral or remorse over something for which one was not at all responsible. When one has culpably committed a significant moral wrong, however, remorse is a coherent moral response. What is much harder to establish is why remorse is appropriate in the second of the two senses explicated: why one is right to feel remorse in these circumstances. Let us begin by defending the view that some

¹⁸ There might be a further way of describing people's emotions which would apply to emotional experiences that sit somewhere between being appropriate (coherent) and being irrational. There might be circumstances in which an emotion is *understandable* yet not fully *appropriate*. A phobia of cotton wool is obviously irrational and thus both inappropriate and not understandable. However, fear felt when walking over the robust glass floor of a very high structure might be understandable yet not appropriate (there is no danger present). What makes it understandable is, for example, that those experiencing the fear perhaps feel uncertain about the robustness of the glass or imagine that it is in fact not structurally sound. If these beliefs *were* to be true then it would be coherent to feel fear, as there might be danger, thus making fear understandable for those that hold these beliefs. Perhaps the threshold for understandability has something to do with the plausibility of the beliefs that, if true, would make the emotion appropriate. Glass not taking one's weight is not very far-fetched. That cotton wool can harm, on the other hand, is a belief that requires ignoring or discounting overwhelming evidence to the contrary.

instances of emotion can plausibly be conceived as right or wrong. We shall take our lead from Aristotle, and the general position of virtue ethics.

Aristotle and virtue ethics

In his *The Nicomachean Ethics*, Aristotle (2007: 41) says that moral virtue is:

concerned with passions and actions, and in these there is excess, defect, and the intermediate. For instance, both fear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way, is what is both intermediate and best, and this is characteristic of virtue.

So, on this view, simply feeling an emotion in the right situation, to the right degree, has irreducible moral value. The idea here is that virtues are not simple dispositions to emotions, but instead dispositions to be *properly* affected by a certain range of situations, where ‘properly’ is also understood as having the right *amount* of some kind of emotion.

However, it could be argued that if we look closely at this passage, we find only an explicit regard for emotional coherence: it is bad to feel an emotion too much or too little, or to not feel it well. We can now, incidentally, consider further what makes up an emotion’s coherence: the fittingness of its objects, which was discussed before, and the fittingness of its intensity. Intensity is still a matter of coherence and not evaluation since it is also a part of its logic: great anger at a small slight is not rational.

But, it seems that Aristotle wants to argue that emotions can be appropriate and inappropriate beyond their coherence. We find Aristotle claiming that anger towards certain forms of injustice is *appropriate*. In contrast, envy and spite are thought to consist in having *inappropriate* feelings of hatred or pleasure. This is a

blanket value judgement of envy and spite, the consequence being that even if logically coherent, an emotion can be morally reprehensible.

Aristotle (2007: 46) says, ‘the man who is characterized by righteous indignation is pained at undeserved good fortune, the envious man, going beyond him, is pained at all good fortune, and the spiteful man falls so far short of being pained that he even rejoices [when someone suffers]’. So, here we find a claim that it is immoral or inappropriate to take pleasure in the suffering of others and inappropriate to be pained by the good fortune of those who deserve it. These are ethical statements about emotion. So, in Aristotle’s overall picture, not only correct motivation but also correct reactive or responsive feelings are constitutive of a person’s virtue. Thus, a person is assessable for his emotions as well as for his calculations and actions.

However, in order to make such statements about emotions, we need to be able to say what it is that makes some emotions appropriate (or not) in the latter of the two senses. I shall put forward a thesis about emotions generally and then move on to consider remorse in particular.

The value of emotions in general

That we judge people’s emotions as well as – or, indeed in spite of – their actions can be demonstrated with some examples. We might say of the overprotective mother who inappropriately restricts her teenage son’s freedom that it is ‘borne out of love’. Although we assess her actions as wrong, we recognise the value in her emotional state. We think less of a person who feels pride over his charitable donation than the person who donates the same amount humbly. We see the woman who responds with sympathy to the tragic accident in the news in a different light to the one who watches

coldly, unmoved. It matters to us whether some action ‘came from a good place’, regardless of whether it was the right thing to do or not. It matters to us that the person who we think of as our friend *feels* a bond of friendship, as well as whether he seems to treat us kindly. It matters to us whether the person telling us she loves us really feels it. In all these examples, we are valuing some emotions (or lack of) and casting blame on others, independently of any activity. What is it about some emotional responses that make them morally valuable or morally repugnant?

Roberts (1991) considers three ‘wicked emotions’: envy/pride, contempt, and anger. I would argue that not all instances of these emotions are wicked (e.g. an appropriate degree of anger at injustice) but let us assume that what Roberts says applies to less controversial cases. What is it, then, that makes these emotions inappropriate, or ‘wicked’? His strongest argument concerning all three is that in feeling these ways, we misuse a person in an attitudinal sense, as opposed to the behavioural sense. His argument rests on seeing all human beings as being ‘moral friends’ and thus having some of the same attitudinal moral responsibilities to people in general as one does to one’s friends. Aristotle said that friendship ‘is a virtue or implies virtue’ and involves ‘bearing goodwill and wishing well to each other’ (2007: 196). So, to be a moral friend to someone is constituted not just by appropriate behaviour, but also by proper attitudes which are subverted in Roberts’ wicked feelings. He says this is possible because:

such feelings are not sensations, but construals of persons whose grammar connects them with such moral concepts as equality, competition, dignity and oppression (envy), the dignity of persons, esteem or respect for persons (contempt), and culpable offense and punishment (anger). [1991: 22]

If we think about positive emotions, we might regard compassion and sympathy, for example, as being constitutive of moral friendship. Indeed, Blum (1980) argues that altruistic emotions such as sympathy and compassion have a value beyond their potency to bring about beneficent actions, and he connects this value with friendship. Remorse might be the appropriate response to having wronged a moral friend: showing him again, or for the first time, the respect which one's offence against him denied. One now wishes him well and bears him good will, and is distraught that one previously diminished this wellbeing, about which one now cares.

Roberts may be right that we have a duty to attitudinally (and therefore emotionally) treat our moral friends well. However, this does not get to the heart of our intuitions about the value of emotions. We do not think about the wrongdoer as having violated his attitudinal duty, we blame him for what his lack of emotion says about *him*. Schopenhauer (1995: 132) captures this intuition:

In some respects the opposite of envy is the malicious joy at the misfortune of others [Schadenfreude]. Yet to feel envy is human; but to indulge in such malicious joy is fiendish and diabolical. There is no more infallible sign of a thoroughly bad heart and profound moral worthlessness than an inclination to a sheer and undisguised malignant joy of this kind. The man in whom this trait is observed should be forever shunned.

Here, Schopenhauer not only ascribes extreme blame to the person feeling the malicious joy but compares its moral worth with the emotion of envy, which he judges to be less immoral. The reprehensible quality of the feeling, aside from any behavioural consequences, comes from its being the emotional activity of the thoroughly bad heart of a person of profound moral worthlessness. Disposition to emotional experience is part of our character and as such is an appropriate object of moral assessment. What remorse might involve is a profound concern for others and acute, accurate moral perception, revealing that the wrongdoer is not in possession of an entirely bad heart, nor is morally worthless.

So far I have argued that intuitions about the value or disvalue of emotions seem plausible. In constituting or revealing a more virtuous character, instances of remorse seem appropriate. However, we must now move from this more general discussion of the value of emotions to look at remorse in particular. I shall now attend to, and respond to, the most pressing objections to the value of remorse in order to make entirely clear the way in which I think remorse is an appropriate response to moral wrongdoing.

Objections to the appropriateness of remorse

Gaita (1991: 51) writes: 'the reasons for the hostility to remorse are various, ranging from a reductive functionalism about value, that focuses on the superficial thought that guilt serves no purpose (why should it?), to the most high-minded of them, which scorns remorse as a form of self-indulgence at the expense of a proper concern for the victim of our wrongdoing'.

I will first attend to the latter of these two arguments - that remorse is self-indulgent - and a third - that remorse is cowardly. These two arguments can be dealt with relatively briefly. Gaita's first reason for hostility - the questioning of the value of remorse - is much more problematic and I will consider a particular version of this argument in detail. Gaita's rhetorical 'why should it?' needs exploring.

The first two objections both misconstrue remorse: the former sees a distorted form of remorse as prototypical and the latter misinterprets core cognitions involved in an experience of remorse, thus devaluing it unfairly. First, Solomon (1976) argues that remorse is self-indulgent, precluding proper attention to the objects that we would hope a wrongdoer would come to have at the front of his mind. A kind of haze of self-pity obscures the view of any victims, or general consequences beyond the wrongdoer.

This would be a kind of narcissistic version of remorse, in which the wrongdoer is only concerned with his or her bad feeling and how he might now appear to others. However, we have already discussed that in a proper (or as Tudor (2001) calls it, 'lucid') experience of remorse the victim of one's wrongdoing, or (more abstractly) the value destroyed, is salient in the mind of the remorseful individual. Although self-indulgent distortions of remorse exist, where the focus is only on 'I as wrongdoer', this does not mean that we cannot rope them off and ask questions about the appropriateness of undistorted remorse.

The second criticism of remorse is that it implies a cowardly attitude to what one has done. This criticism is more interesting. Morris (1971) discusses the charge that feeling guilty (undistinguished from feeling remorseful) implies 'a cowardly attitude towards one's past deeds, an unwillingness to let go of the irretrievable and face the future. Instead of accepting that what is done cannot be undone, the guilty frantically attempt to undo. For the courageous there are no erasures.'

However, it seems to me that what needs to be made clear is what is meant by 'attempting to undo'. If the agent is trying to 'undo' in the sense that he would then be able to see himself as blameless or as if his actions never happened, then this indeed would be cowardly. His undoing would be motivated by the desire to rid himself of bad feeling. However, these would not be the intentions of the truly remorseful. This cowardly undoing is different to the 'undoing' often motivated by remorse. Here, one tries to undo the harm caused because one wants to reduce the victim's suffering or otherwise make up for what one did by way of reparation. This is perfectly compatible with knowing one cannot erase what one did and may be motivated by precisely this thought. That it would be courageous to say 'I'm not going to bother

trying to make it up to you because I can't change the fact that I did this to you' seems absurd and arrogant. Remorse, then, is not cowardly.

Bittner's objection: remorse is not reasonable

The third, and more serious, objection to the value of remorse is represented by Bittner's (1992) question: 'is it reasonable to regret things one did?' He intends regret to be understood as encompassing emotions in which the agent experiences painful feelings about something he did that he thinks was bad. He cites remorse, repentance, and feelings of guilt as specific kinds of phenomena falling under this description. Thus, according to Bittner, remorse is a species of regret. The consequence is that he must intend his argument to apply equally to remorse.¹⁹

Bittner explains that 'reasonable' is to be understood roughly as 'it's a good idea', 'it makes sense', 'it is recommendable'. We can note straight away that the question he poses seems to have a somewhat prudential orientation. At least, 'good idea' and 'recommendable' imply that a person may be choosing between the options of regretting or not regretting and that he is concerned about the consequences of this decision. 'Making sense', on the other hand, is concerned with intelligibility – would it be rational for the person to feel regret?

Here, we have a potential conflation of the two meanings of 'appropriate'. Under the same heading of 'reasonable', he includes value judgements and appeals to coherence. Bittner does make this error: as will be seen, he thinks that if he can show that remorse is not recommendable then it will cease to make sense. His idea is that emotions that cease to make sense will cease to be experienced, and that this is a good thing due to their not being recommendable. How this confusion affects parts of his

¹⁹ Indeed, Bittner specifically says that he has in mind regret over something taken to be wrong for moral reasons, and he is not concerned to demarcate what it is to regret things one did. So, that he also has remorse in mind is likely.

argument will be seen. Of note is that none of his descriptions of reasonableness attend to the potential for there being moral value to feeling regret.

Bittner starts from the idea that we might be able to re-programme our emotions, teaching ourselves to desist from experiencing those that our reason tells us make no sense (coherence appropriateness). An emotion's failing to make sense seems to be what Bittner thinks is required in order for the emotion to fall out of a person's repertoire of affect. If this is the case, Bittner must have to successfully argue that the emotions of regret make no sense. So, before we assess the particular elements of Bittner's argument, we should note that Bittner seems to be making three broad claims:

- 1) That remorse is not reasonable – has no value *nor* coherence.
- 2) That we *are able* to divest ourselves of unreasonable emotional responses.
- 3) That, since remorse is not reasonable, we *should* so divest ourselves.

In looking first at the question of the reasonableness of remorse, let us begin where Bittner begins: with his version of an argument from Spinoza.

Spinoza

Bittner first considers Spinoza's arguments about the desirability of repentance. In doing so, Bittner not only finds no reason to encourage repentance following wrongdoing, but goes as far as emphasising the disvalue in it. It is not just useless, it is damaging.²⁰ The argument proceeds thus: it is not reasonable, because one did something bad, to make things worse for oneself by suffering repentance. In other words, regret is double misery: the second (subjective) for the sake of the first (objective). Bittner paints the picture of the 'reasonable' wrongdoer who sees that

²⁰ Whether uselessness in itself confers disvalue is an interesting question, but not of particular importance here.

what he did was wrong and may be perfectly aware of the suffering he inflicted on others – he understands that he has put himself into an objectively ‘low’ state, he just does not grieve. That human beings would be sensible to avoid suffering pointless pain if they can is fairly uncontroversial. But the question, on Bittner’s terms, then becomes whether the ‘misery’ of the emotions of regret is pointless.

Bittner’s argument contrasts interestingly with the view espoused by Gardner (2005). In his paper, Gardner also draws a distinction between lives being objectively blemished by wrongdoing and subjectively blemished by the way we feel towards this wrongdoing. When the wrongdoing was faulty (i.e. when blame can be conferred on the person for the wrong), Gardner explains, the subjective blemish experienced can be described as the special kind of self-critical regret known as remorse. Faultless wrongdoing, on the other hand, fosters simple regret. Gardner argues that it is the continuing force of the undefeated reasons ruling against commission of the wrongful action that makes regret ‘rationally appropriate’.²¹ Here we have an assertion of coherence appropriateness.

Gardner (2005: 125) goes on to explain why he thinks the subjective experiences of regret and remorse are not only rational but, in a sense, deserved. He says that ‘our lives should be blemished subjectively because and to the extent that they are blemished objectively. More generally, our lives should feel as good, but only as good, as they are.’ Perhaps in the particular case of remorse, we might add that our feelings towards ourselves should be as good, but only as good, as we are. Gardner describes this doctrine as a doctrine of aptness or fittingness, saying that it is apt or fitting that people should suffer for their wrongs. Gardner is making a value

²¹ This contrasts directly with Spinoza who says in his *Ethica*, that ‘repentance is not a virtue, i.e. it does not arise from reason’ (2002: 348).

claim about regret and remorse, saying that such feelings are appropriate because they are deserved.

These notions of ‘aptness’ and ‘fittingness’ are at a conceptual distance from Bittner’s question of whether remorse is ‘commendable’. This contrast highlights Bittner’s somewhat consequentialist approach to the issues. Whereas he seems to be concerned with finding an instrumental value for remorse, Gardner is looking at whether remorse might be appropriate simply because it is deserved. Indeed, Gardner draws a comparison between remorse and punishment, suggesting the former might be a self-imposed version of the latter. Gardner is concerned with the alignment of subjective ‘goodness’ with objective ‘goodness’. In the event that this alignment is not satisfied by the experience of remorse, punishment can serve to correct the misalignment, giving people who do not feel remorse for their faulty wrong something extra about their faulty wrong to regret.

The ideas that punishment can substitute for remorse, or that remorse should be understood as a self-inflicted analogue of state punishment will be explored in detail later on (in chapters Four and Six, respectively) but, for now, what has been drawn out through comparing Gardner’s argument with Bittner’s is that we might find remorse appropriate (or reasonable) for reasons other than those to do with its usefulness. Rather, coherence appropriateness might be found more internally to the logic of remorse – in its reason-born fittingness – and value appropriateness in the result of deservedly aligning our subjective lives with our objective lives. However, although remorse may well be deserved, and in a similar way to retributive punishment, I do not think this is the only value remorse has, and not why we might attribute virtue to the remorseful.

Having set things up so that the burden is on the advocate of remorse to show why there is, after all, reason to be ‘twice miserable’, Bittner considers – in order to reject – a couple of possible arguments. The first being that regret is necessary for retaining one’s character and identity.

Remorse as necessary for retaining character and identity

Bittner presents an apology for regret constructed from the ideas of Bernard Williams. The argument can be summarised: when one fails to regret the consequences of one’s actions, one detaches oneself from these consequences. But, in so detaching oneself, one fails to retain one’s identity and character as an agent. The suggestion is that *who we are* is importantly constituted by a certain ownership in relation to our actions over-and-above the simple fact of our having performed them.

In claiming that one can detach oneself from what one did whilst still knowing that one was the doer, Bittner relies on a distinction. This distinction is between knowing that one was the doer of the deed and ‘identifying’ with what one did. What more is there when one ‘identifies’? Bittner rearticulates ‘identification’ as ‘positively taking on’ the action as one’s own and ‘incorporating it in what [one] is as an agent’ (1992: 269). It would seem that this notion of identifying with one’s deed involves seeing the deed as somehow contributing to how one sees oneself, and finding one’s self-image consistent with the actions in question. The implication of the distinction seems to be that in simply knowing what one did, one can see it as alien to oneself. It does not define one. In contrast, in identifying with what one did, one makes sense of who one is in terms of it, reconciling self-identity with behaviour.

I am going to argue that when experiencing remorse one does retain one’s character and identity but that this is not by way of ‘identifying’ with our deed in the

way that Bittner suggests. Bittner seems to assume that regret (and by extension, remorse) precludes renunciation of what the deed implies about one. In feeling regret one defines oneself by what one has done. The regretless person, on the other hand, is purportedly able to say ‘I am no more in it. You cannot find me there’ (ibid.: 271). Bittner thinks that such renunciation is not possible if one is, instead, remorseful.

However, I suggest that the experience of remorse does not have the identity-building, or -revealing, capacity in quite the way Bittner thinks. Indeed, if it did, it would be more akin to the experiences of shame (as discussed in the previous chapter) in which one’s flawed identity is revealed to one and where one sees this identity as an unchangeable fact which will go on to dictate further flawed behaviour. Rather, part of the pain of remorse occurs precisely because one cannot identify with the deed. The agent cannot find *himself* there. As Adam Smith (2002: 98) describes: ‘...he can enter into none of the motives which influenced it. They appear now as detestable to him as they did always to other people’. But yet, he still painfully recognises ‘the wrongdoing as having its origin in [him], as being *of*[him]’ (Norman 1971: 95).

These conflicting thoughts – this was me yet I do not recognise it as me – create a ‘fission’ or disharmony within oneself (Levinas 1991: 125).²² Thus there is a sense in which remorse, whilst involving acknowledgement of ownership of the deed, allows one to detach oneself from it: one renounces it. So, in remorse the agent does not maintain his character and identity by simply adding the deed to the list of information about himself from which he constructs and understands who he is.

²² Festinger’s (1957) psychological concept of cognitive dissonance may help explain this aversive component of remorse. Cognitive dissonance involves conflicting attitudes, beliefs or behaviours, which produce a feeling of discomfort. In remorse, one condemns – and dissociates oneself from – those vices that are involved in wrongdoing whilst at the same time observing that one has carried out wrongful actions. Indeed, Aronson reformulated Festinger’s theory by linking it explicitly to self-concept. He argues that ‘dissonance is greatest and clearest when it involves not just any two cognitions but, rather, a cognition about the self and a piece of our behavior that violates that self-concept’ (1992: 305).

Rather, in remorse the agent experiences inner turbulence, which actually requires a certain integrity or, as Tudor (2001: 144) puts it: ‘a certain integrity or self-collectedness, that is a certain self-understanding and self-awareness of all that one is, in order to *feel* the dissonance within one. One needs to have enough integrity to recognize the fissure in one’s soul’. Bennett (2008: 166) describes this painful experience as ‘self-dissociation’.²³

We can helpfully draw a distinction between two types of identification. Bittner’s notion of identification involves incorporating one’s deeds into one’s self-image, to which one assents. Let us call this ‘thick’ identification. If a man who stole something now self-identifies as a thief, he identifies with his deed in the thick sense. Identification, however, can also occur in a ‘thin’ sense. In such cases, one accepts agency for one’s deed, but does not see it as integral to one’s character. A man who stole something does not have to see himself as a thief in order to take full responsibility for his theft: he might despise thievery and be entirely surprised by his own actions. He would then be identifying with what he did in the thin sense. What I am arguing is that Bittner thinks that remorse necessitates thick identification whereas I suggest that actually remorse only involves thin identification. Moreover, it is this tension caused by identifying thinly but not thickly that motivates or reveals renunciation of the deed: one understands what one did whilst finding it irreconcilable with one’s conception of one’s entire self. One withdraws or dissociates from this part of oneself.

²³ Bennett (2008: 166) explains this idea within the following context: ‘If blame is a case of withdrawing from the wrongdoer, guilt is a case of withdrawing or dissociating oneself from (a part of) oneself. It is a painful splitting of oneself in two: rejecting or repudiating that aspect of oneself (one’s greed, laziness, insensitivity, selfishness, pride etc.) that brought the wrong about; and yet recognising that it is nevertheless part of oneself. Now feeling bad about oneself like this, though painful, is an essential part of “making things right”. If the retributivist is right that we ought to blame wrongdoers, then it must also be the case that wrongdoers should be subject to the pain of guilt. For the pain of guilt and what we might call self-dissociation will be an essential part of understanding that one has done something that merits blame and withdrawal.’

Thus, one's character and identity are maintained not because one positively (thickly) identifies with the deed, but because one painfully experiences a challenge to one's self image, through thinly identifying with the deed. Without the pain of remorse, one actually does *not* enjoy a 'particular unity of the soul', because one would be blind to particular aspects (Tudor 2001: 144).

This reconceptualisation of how the experience of remorse is linked to character and identity makes clearer why it might in part be the *uncomfortable* aspect of regret or remorse that allows maintenance of character and identity – maintains the unity of the self. In Bittner's argument, we do not see why 'grieving' is necessarily linked to accepting the deed as one's own. It is the grieving of regret that Bittner wants rid of and he seems to think that this would naturally flow from not identifying with the deed. However, I have argued that in remorse some of the discomfort stems from the fissure within one, which partly *results from* not identifying in this thick sense.

Remorse as necessary for deep understanding

However, Bittner might still say that this conception of maintaining character and identity does not make a difference to his argument that such maintenance is simply not obligatory or even recommendable. Whether this is the case or not is not of ultimate significance for the question of whether it is reasonable to regret the things one did because I think there are much stronger arguments. The first concerns how remorse might be necessary for full moral perception.

Bittner argues that 'it is not evident that one could not see in full clarity but without grief, what one did wrong'. I agree that one could have some notion that what one had done was wrong in the sense that one could know that one's action belonged

to a set of actions labelled 'wrong'. Yet, I would argue that such an intellectualised understanding of 'wrong' does not constitute a deep understanding. Nussbaum's (1990: 79) writings offer a similar position on other types of emotional experience.

She argues that

the agent who discerns intellectually that a friend is in need or that a loved one has died, but who fails to respond to these facts with appropriate sympathy or grief clearly lacks a part of Aristotelian virtue. It seems right in addition to say that a part of discernment or perception is lacking. This person doesn't really, or doesn't fully, see what has happened, doesn't recognise in a full-blooded way or take it in...The emotions are themselves modes of vision or recognition. Their responses are part of what...truly recognising or acknowledging consists in.

In the case at hand, remorse 'is a recognition of the reality of another through the shock of wronging her' (Gaita 1991: 52), and this 'shock' is painful. That Bittner overlooks this particular painful element is suggested from the words he uses to describe the pain: 'grief', 'hurt' 'torn' 'dejected'. All these are focussed on the self and, whilst I argued for a self-orientated discomfort experienced as a result of the fissure in one's identity, this is not the most important element for moral value. In experiencing the pain of remorse, we deeply understand what we have done.

Bittner's acceptance of his own argument might partially stem from his presumption that feelings of remorse are *just* painful sensations. He wants the agent to acknowledge responsibility and see the deed as wrong but not to then grieve in addition. Regret, for him, seems to simply be this grieving which the irrational tack on.

However, remorse and regret are better characterised as painful thoughts. Feelings are necessarily intertwined with cognition. As Teichman (1973: 345) says; 'remorse is a form of understanding' and what one understands in particular in remorse is the reality of the other. In painfully comprehending the other's moral reality and one's relation to it, one demonstrates care for the other. Indeed, according

to Morris (1971: 427), 'part of what it means to care for another is feeling pain in circumstances where another is pained. A mark of care both for oneself and others is the pain experienced when one hurts another or oneself.' Thus, a distinction can be made between knowing what one did and deeply understanding what one did. The former is possible without remorse but often not the latter. Remorse, then, becomes appropriate for this reason.

But is this just counter-assertion? I have already argued that Bittner does not acknowledge that affect and cognition might be interdependent when it comes to responses to wrongdoing. I will now make this position more robust. As I argued in Chapter Two, emotions reveal and make salient to us facts about the world and provoke evaluation of these facts. Stocker (1996: 123), for instance, argues that 'emotions can be organizationally dominant, helping constitute central thoughts and categorical preoccupations, interpretations of situations and...they focus attention and define what is salient'. This is similar to C. Taylor's (1985) argument that emotions reveal 'import' to the agent. We know what matters to us, and in what way, because we experience it through feeling. Importantly, there is interaction between the affective and non-affective elements of an emotion: our feelings incorporate a certain 'articulation' of our situation, that is, they presuppose that we characterize our situation in certain terms. But, at the same time, they admit of further articulation, permitting more penetrating characterisation. In turn, this further articulation can transform or deepen the feelings.

For example, it might be through experiencing a certain sadness that I realise that I miss a particular person. I might then think about why I miss her and about what she brings to my life. This then focuses my sadness. I then understand that I value her

and her part in my life more than I had previously realised. My feelings prompted my realisation, and further reflection intensified and sharpened my feelings.

Similarly, through such a process of emotional self-interpretation the remorseful person comes to really understand why she attributes so much disvalue to her wrongdoing. This co-refinement of moral understanding and affective response is summed up by Wiggins (1991: 196):

Finer perceptions can both intensify and refine [emotional] responses. Intenser responses can further heighten and refine perceptions. And more refined responses can lead to the further and finer and more variegated or more intense responses and perceptions.

Thus, on this view, without the affective element of remorse, self- and moral understanding would be restricted. Importantly, however, it is not *merely* the affective element of remorse to which we attribute value, it is remorse conceived as the response which also involves moral understanding, which the affective element helps to elucidate. We value remorse in an agent because it precipitates his insight into morally-relevant facts.

Establishing intrinsic value

Even if this argument is correct, and remorse does function to provide us with the deepest kind of moral understanding, it could still be argued that the value of remorse here comes from its function – it is a tool (even if a necessary one) for attaining moral understanding. Thus, its value is arguably instrumental. Even the ‘justice’ value remorse might have in aligning the subjective ‘low’ with the objective ‘low’ is instrumental, as it is in the proper alignment that the value ultimately lies. As long as the value of remorse is dependent on the value of something else, there may be ways to render it redundant: in the first case, we could perfect our purely-intellectual

capacities for moral understanding or, in the other, ensure that people experience a proportionate subjective loss through deployment of state punishment, so that in both cases the role of remorse is minimized or even eliminated entirely. In order to claim intrinsic value for remorse, we must value it independently of anything it might lead to or bring about: we must value it just in and of itself. We now return to my original argument at the beginning of the chapter – that remorse has intrinsic value because to experience it when coherent is morally valuable. This was shown by demonstrating that it is an appropriate object of moral assessment. We can now make this argument clearer.

In order to get to grips with the potential remorse might have for harbouring value intrinsically, we must generally ask the question whether, all else being equal, a world in which there are instances of remorse is a better place than a world in which there are no instances.²⁴ Some thought experiments will help us consider this. We might imagine a situation in which two people have committed exactly the same wrong and both apologise to the victim, providing reparation voluntarily. However, one of the wrongdoers does so remorsefully and the other does so begrudgingly (although he does not reveal this attitude). It seems that the former response retains some additional value, even though the victim may be left feeling the same in both cases. However, it may be argued that emotions can be seen to be part of a morally correct action: a remorseful apology and remorseful reparation are required as the most virtuous response. The value of feeling remorse is thus still in some sense parasitic, as it depends ultimately on its contribution to the full correctness of the action.

²⁴ This particular argument has been greatly influenced by those of Stark (2004). I will indicate where I make specific use of her language.

In order to really test for the value of remorse – value that comes simply from being appropriately felt – we might do better to consider a situation in which remorse is not part of an action but is simply felt in response to wrongdoing. Perhaps we might consider a situation in which it would actually be morally required that the wrongdoer does *not* apologise or try to make amends. Imagine an assault victim who expressly asks that the offender makes absolutely no contact in any way as he believes that any reminder of the assault would be too traumatic. He just wants to forget the whole thing. Here, the offender would actually be required to abstain from any reparative action. We might again imagine two offenders in this same situation, one who feels remorse and one who does not. It seems that the situation in which the offender nonetheless feels remorse is the morally better situation, suggesting that, where there is no morally-required action, moral assessment instead becomes focused on the wrongdoer's emotion. Stark (2004: 368) makes essentially this argument, concluding that 'moral emotions are sometimes valuable just in being properly felt. In cases where there is nothing to do, our emotions are fully and solely constitutive of our moral concern'.

To deny this, would be to deny that Schopenhauer's man, who indulges in malicious joy at the misfortune of his victim, is no morally worse than the man who is wracked with remorseful compassion. When there is nothing 'morally to do', it is these responses to which we look for evidence of virtue or vice. This does not mean that remorse, or moral emotions in general, are *only* intrinsically valuable when there is no morally required action. It is just that the enduring intrinsic value of remorse becomes most apparent in such situations.

It might be objected that we cannot allow such moral emphasis on emotions since people's capacities for emotional experience vary widely and so an unfair

element of moral luck would be brought in. It is true that emotional capacity varies. What might be morally required then, is that one feels the appropriate emotions to the extent that one can and that one attempts to, contra Bittner, cultivate coherent moral emotions where possible, thus moving closer to the ideal.

Still, it might be argued, even if we have the capacity for rich emotional experience, our emotional lives are busy and feeling the correct emotion at the correct time to the correct degree might often be too much to ask for. An example will illustrate this argument. Imagine that a woman mugs someone on the street. An hour later she hears that her son has died. Although she had begun to realise the wrongfulness of her actions and ordinarily would have had the capacity to feel an appropriate amount of remorse, the grief of her son dying leaves her emotionally drained. When she thinks of the mugging, although she recognises the wrongfulness and attempts to make amends, emotionally she is numb towards it. Is she therefore less virtuous than she would have been if she had been able to utilize her capacity for remorse? Perhaps in such situations, excuses might justifiably function for emotional response in the same way as they do for actions. An action is often less bad if performed under duress; an absence of emotional response might not be a vice to the same degree if understandably displaced by other extreme emotions.

We might also consider Gardner's (2005) view that lives should subjectively feel only as good as they are objectively. Perhaps the woman's intense grief actually brings her subjective state down to (or even beyond) a level that would be commensurate with the objective level of her life. Unless Gardner wants to specify exactly in what way a life should feel subjectively bad (e.g. pains of remorse rather than pains of grief), or wants to specify that there has to be a causal relationship

between the objective ‘blemish’ and subjective ‘blemish’ then this woman might actually not be out-of-kilter.

In connection with Gardner’s (2005) concept of desert in subjective feeling, we might draw a comparison between such negative emotional overload and the ‘totality principle’ in sentencing which requires that where an offender is being sentenced to multiple terms, or is otherwise to serve multiple sentences, then the sentencer should ensure that the total sentence remains ‘just and appropriate’ for the whole of the offending. This often means that sentences are understood as running concurrently and almost always results in there not being simple accumulation of sentences. So, in the case of the grieving woman, we might think along the same lines: her painful emotional state is already so intense (albeit for a non-offence-related reason) that to require extra suffering in addition is not fair – in a way her remorse might be conceived as running in parallel to her grief, despite not being independently identifiable. However, analogy with the totality principle breaks down when we consider that although we might want the emotionally-burdened woman to be excused her lack of remorse, we would not want remorse to actually be unprincipled. This being said, the notion that a person might already be ‘suffering enough’ emotionally does seem tenable and will be considered in Chapter Six.

Therefore, it would seem that in addition to the instrumental value remorse can have in enabling acute moral perception, it also has intrinsic value in simply being appropriately felt. Despite there being differing capacities for emotional response, and the possibility of busy emotional lives preventing full functioning of this capacity, our emotional responses are a locus of moral virtue and thus also a locus of value.

The possibility of emotional reprogramming, and what is at stake

Bittner might reject the perspective of virtue ethics and remain adamant that remorse is not reasonable as no one suffers if he does not feel remorse (so long as he offers reparation where possible and decides not to wrong anyone in this way again), whilst he *does* suffer if he does feel remorse. Bittner may still hold that, having intentionally hurt another person, it is sufficient to recognise that what he did belonged to a set of behaviour labelled as wrong – to say ‘true, I did that, I even did it intentionally. But I am no more in it. You cannot find me there. That is history. It does not cast a shadow over what I do now’ (Bittner 1992: 271). Yet, it is difficult to see how one could reprogram oneself to not feel remorse where one would have otherwise; not without denying the wrongfulness of one’s action and the harm caused. Let us return to Bittner’s instructions for emotional reprogramming. He suggested that the availability of a feeling is dependent on ‘how we see things’ and that we can divest ourselves of an emotion by coming to have a ‘different view’. Such a view is also held by Williams (1973: 224) who writes that ‘to feel a certain emotion towards a given object is to *see it in a certain light*; it may be wrong, incorrect, inappropriate to see it in that light, and I may become convinced of this. When I am convinced, the emotion may go away’.

We can speculate as to how this might have worked for Bittner as he claimed to have lost the emotion of reverence for people through coming to have a different view of people who were important to him and of himself.²⁵ If to feel reverence for someone is roughly to respect and be in awe of them, one could teach oneself to always make sure one sought out and attended to others’ failings, thus precluding

²⁵ Incidentally, reverence seems a very reasonable emotion to me (if not felt excessively): it draws our attention to what we value (say, great musical ability, or unusual levels of wisdom) whilst instilling in us a healthy humility. Since reverence is less exclusively a moral emotion than remorse, it might not be as controversial to attempt to divest oneself of it. Nonetheless, I would be interested to see how Bittner would argue that reverence does not ‘make sense’.

idealising the other. Or one could always focus on how one is equally worthy of respect and admiration. William's (1973: 224) examples are equally instructive:

...my fear of the impending car journey evaporates on learning that Miss X is not in fact going to be the driver; or my reserve and suspicion towards this man dissolves when something shows that his manner does not mean what it appeared to mean; or my passionate loyalty to the partisan leader suddenly cracks when I am convinced that his actions can only mean betrayal.

However, in thinking about these examples, we come to see here that what Bittner's instructions suggest is that we must convince ourselves that a given object or type of object is not a proper object of that emotion – that is, we must deny its coherence. Sometimes this is appropriate, as in William's examples, when the object of the emotion really does change. However, the implications of Bittner's instructions for remorse would lead us down an unattractive path, and one I do not think *he* would really want us to take.

In the case of remorse, it is difficult to see how one could have a different view of one's wrongdoing and victim – see them in a different light – whilst retaining an appropriate level of moral understanding. I suggest that successful desistance in remorseful experience (in the way Bittner suggests) would require one to change how one sees things. Further, this would have to be to an extent that precluded both full understanding of wrongdoing and appropriate consideration of the consequences of the things one does. One would need to make the emotion incoherent. Perhaps one might play down one's responsibility or might think about how the victim will 'get over it'. But, Bittner wants to *retain* clarity of understanding; he just wants us to divest ourselves of the affective experience. However, it seems to me that divesting ourselves, following the guidance he gives us, would be at the expense of retaining this full understanding since we would need to change 'how we see things' so that they had no emotive force, denying the moral significance of what we had done.

Here, Bittner's conflation of the two ways in which an emotion can be appropriate results in his argument having an unattractive consequence, even for him. He thinks that since he has argued that remorse is not 'a good idea', it fails to make sense. This is not the case, and forcing it nonetheless into failing to make sense results in obscuring the reality of the wrongdoing to an entirely undesirable degree.

Of course, often there are emotions that we judge to be wrong or imprudent and so wish no longer to feel. Take jealousy. Perhaps we might try to convince ourselves of why our friend's promotion really is not something to be desired, although this may well involve a degree of self-deception. However, the more honourable alternative might be to admit that, 'yes, Jane's new job is fantastic – better than my job – but Jane is my friend and I should be happy for her in her achievement, not resentful'. This may involve will power and self-control, but not self-deception. However, I have argued that remorse is not something we should try to suppress in ourselves – when coherently felt, it is virtuous or revealing of virtue. Through self-deception we lose moral understanding, through suppression we become a less admirable person.

Perhaps Bittner's suggestions are possible in cases of simple regret. One could change how one sees things by adopting a pseudo-deterministic outlook on life. If everything was always going to happen as it has and there is nothing one could do to change it then kicking oneself and comparing other possible worlds stops making sense. One might feel disappointment but not regret. In fact, we might concede that simple regret is sometimes not reasonable, especially for non-moral actions. We are pushed and pulled by social and other forces all the time and should not berate ourselves when every decision we make is not the optimum: 'don't sweat the small stuff'. We should make a note about how to approach things if a similar situation

arises, but then move on. There is no further value in regret here. This cannot be true for remorse, however, since it is a response to moral wrongdoing and so is a locus for moral virtue when felt appropriately. Moreover, as I argued earlier, great instrumental value lies in its capacity for enabling the deepest moral understanding of our wrongdoing, and such understanding is intrinsically valuable.

The fleeting objection

But, it might be argued, Bittner does not have to work as hard as he does at arguing that we can teach ourselves to desist from experiencing remorse. The objector might continue: you have talked about remorse as the morally appropriate and valuable response to wrongdoing as if it is something categorical and permanent – one either has the appropriate response or one does not – when in reality people might experience remorse for an hour, a couple of hours, intermittently, or any number of ways that do not constitute a permanent response. Thus, Bittner's reprogramming becomes less important. If you want to affirm the significance of remorse for punishment, the objector challenges, you need to be clearer about what really is required for – what constitutes – this appropriate response to wrongdoing.

My response would be this: I have already discussed in the previous chapter how emotions are often not experienced categorically and might co-occur or shift. It is not here a matter of saying how soon and for how long a wrongdoer must experience the pain of remorse in order for it to fully be the appropriate response to his wrongdoing. Rather, it again highlights that the moral quality of remorse comes not only from the affective element, but also from the intertwined cognitions, the content of which involves understanding the wrong, concern for the victim and resolution never to act in such a way again. For the response to be worthy of the label

‘appropriate’ I would argue that the wrongdoer does not have to experience great pain for the rest of his days.

Rather, we might consider the notion of a ‘latent emotion’, which has the power to reignite the affect when the facts about the wrong are dwelt on, but more significantly has a sustained effect on how the wrongdoer thinks about his future actions in the shadow of his past wrong. There is, then, a distinction between a latent emotion and a non-existent emotion. The wrongdoer who feels pangs of remorse for a couple of hours and then puts the wrong out of her mind, never to feel remorse again, perhaps has not responded appropriately – certainly not ideally – to her wrong. Rather, a continued instantiation of latent remorse, not always affectively intense, but at least sometimes and initially so, may come closer to exemplifying the realistic and appropriate response to one’s wrongdoing.

Concluding Remarks

I have argued that remorse not only has great instrumental value in facilitating the deepest moral understanding, but that it is valuable *per se*. It was argued that our emotional responses to our wrongdoing are appropriate objects of moral assessment and that an experience of remorse constituted a virtuous – and thus morally valuable – response to wrongdoing. Despite varying capacities for emotional response and competing burdens on these capacities, remorse was seen to remain valuable when felt correctly. It was also argued that a staunch rejection of a virtue-based argument would be met with the challenge to demonstrate how one could desist from experiencing remorse without altering one’s conception of one’s wrong. On Bittner’s argument, an alternative conception of the facts and consequences of the wrong would

be necessary and thus one's hold on the (im)moral quality of what one had done would become weak at best.

But, perhaps, it was suggested, reprogramming is not even necessary, since emotional responses are often fleeting and certainly they vary in their characteristics. Since we might very rarely categorically feel remorse, the question then became whether value could ever be attributed to anything other than the full and enduring emotion. In response, I argued that a less categorical understanding of the remorseful response should be adopted and that the concept of a latent emotion was a helpful and more realistic one. One must also remember that virtue is not categorical either: it is not the case that one either never responds virtuously to the things one does or always does. Our degree of virtuousness can change minute by minute, but remorse still remains morally valuable every time it is experienced coherently.

Chapter Four

Can punishment constitute or substitute for some of the appropriate remorse?

In this section so far we have been focussing on developing a clear understanding of remorse. Chapter Two explored what experiencing remorse involves and suggested what should minimally be present in order to justifiably describe someone as remorseful. Chapter Three defended the view that remorse is the appropriate (and a valuable) response to moral wrongdoing. The second section of the thesis will examine the institution of state punishment, asking whether remorse should have any mitigating influence on the punishment an offender receives. However, before we move on to this project, which forms the remainder of the thesis, we can ask a preliminary question which will serve to expand further our understanding of remorse, whilst also introducing punishment into the discussion.

The question this chapter asks is whether, given that we value remorse, punishment can constitute or otherwise substitute for some of an individual's remorse, where it is not forthcoming. In Dante's *Inferno* we find an allegory that might lead to this question being posed. Canto 34 describes the 9th circle of Hell, which is reserved for the worst of the worst sinners. The ultimate punishment is here inflicted upon the three greatest traitors: Brutus and Cassius, who murdered Julius Caesar, and Judas Iscariot, the disciple who betrayed Jesus. These sinners each dangle from one of Lucifer's three mouths, being chewed to shreds for eternity without ever dying.

This portrayal of the ultimate punishment has close connections to the etymology of remorse. Recall that the root of remorse, 're-mordere', means to bite again and again. This idea of incessant gnawing is exactly what is inflicted on the

worst of Dante's sinners, suggesting a close semantic relationship between remorse and classical conceptions of the nadir of punishment. Thus, and in view of the line I have taken about the nature and importance of remorse, it might be reasonable to think of remorse as the basic phenomenon and punishment as a somehow artificial proxy for it.

When asking this question one could be asking a variety of things, and an affirmative (or negative) answer to one of the possible, distinct questions does not necessitate affirmative (or negative) answers to all the others. I am going to take each of the questions in turn, exploring the extent to which punishment and remorse overlap, and finally raising questions about the aims of punishment in relation to remorse. The question with which we begin is the most basic:

Is experiencing remorse phenomenologically similar to the experience of being punished?

I suggested above that relentless gnawing characterises both the etymology of remorse and one classical depiction of eternal punishment. Of course, neither remorse nor modern day punishment literally involves such an experience. We can ask, however, if there are any phenomenological similarities between experiencing remorse and the experience of being punished.

It is clear that at the most basic level we can assert that neither experiencing remorse nor being punished is pleasant.²⁶ Neither are experiences that people actively seek out. I am going to pick out some ostensible similarities between the experiences of remorse and punishment and explore the extent to which they really overlap.

²⁶ I will put aside the complicating issue of masochism and whether enjoyed punishment really constitutes punishment.

Feeling trapped

More contemporary literature provides us with one of the potential similarities of the two experiences: feeling unable to escape. In Ian McEwan's *Atonement*, Briony lives with the remorse she feels after the false testimony she gives resulted in her sister's boyfriend being wrongfully convicted of rape, and imprisoned:

The only conceivable solution would be for the past never to have happened. ... She longed to have someone else's past, to be someone else, like hearty Fiona with her unstained life stretching ahead, and her affectionate, sprawling family, whose dogs and cats had Latin names, whose home was a famous venue for artistic Chelsea people. All Fiona had to do was live her life, follow the road ahead and discover what was to happen. To Briony, it appeared that her life was going to be lived in one room, without a door. [McEwan, 2003: 272]

This description of living with the wrong one has done suggests that it has similarities with being locked up. One is trapped by one's past and cannot escape it for the rest of one's life. Of course, state punishment can consist in sanctions other than imprisonment. Fines and community service, although difficult to escape, do not have the same utterly restrictive quality. However, the sort of offences that we would expect (or hope) to beget a remorseful response in the offender are the sorts of offences that would receive prison sentences.

However, there is an important difference between being trapped by one's remorse and being literally trapped in prison. Remorse may persist much longer than punishment. Briony anticipates her remorse enduring indefinitely: a life sentence. In all but the very worst cases, punishment is finite – there is a pre-specified time after which it will cease. It could be argued, then, that the phenomenology of remorse involves the expectation of its persistence and, in many cases, its actual persistence even if sometimes it is latent (see Chapter Three). The sorts of punishments that

modern civilisations impose have time limits which are made clear to the offender and which go on to actually be adhered to.

So, although a long prison sentence might feel like ‘forever’ to the offender, the offender does know that it will come to an end (regardless of his feelings towards this fact). This knowledge is not part of an experience of debilitating remorse. We are familiar with remorseful people saying such things as ‘I’m going to have to live with this for the rest of my life’.

However, perhaps we should not be too quick to underestimate the long-term effects of having been convicted of a serious offence, and of having served a long spell in prison. Those with criminal convictions of a serious nature (or even not so serious nature) often, for example, find it very hard or impossible to find employment.²⁷ The offender may have to live the rest of his life with some degree of stigma when amongst those who know of the conviction. Also, the loss of motivation and optimism that often results from being incarcerated may well persist long after release and some offenders may never fully reintegrate into society. In these senses, there is no escaping their punishment. Both remorse and punishment, then, involve the feeling of being trapped.

A further, related difference can be seen when we think of the sense of isolation involved in being trapped in both remorse and punishment. In Chapter Two I discussed Gaita's claim that remorse renders one radically singular: 'remorse is the only form of suffering that cannot legitimately seek comfort in a community of the guilty' (1999: 33). Whilst punishment can be to some extent isolating - one is restricted in one's social engagement and in some cases offenders are literally solitarily confined - 'being punished' is a form of suffering for which one can gain

²⁷ However, most jurisdictions, including England and Wales, have criminal records legislation permitting offenders to declare that they have no convictions after a period of time.

comfort. Camaraderie and expressions of solidarity amongst offenders may ease the offender's burden and the mere knowledge that others are also having to undergo punishment will likely preclude a radical sense of isolation. Remorse, then, is the more thoroughly isolating.

Anger and sorrow

The phenomenology of both remorse and punishment can involve episodes of vacillation between anger and sorrow. However, the respective objects of these emotions are different in remorse and punishment. In remorse one is angry at, or disappointed in, oneself. Horrified at what one has done, one reprimands oneself again and again. If one is being punished, especially if not remorseful, one may (but not necessarily) be angry at any number of other people: prison officials; the judge who decided the sentence; the victim for going to the police; anyone who one can make central to the cause of one's predicament. Perhaps one simply feels a non-specific anger, borne of the very fact that one is being punished.

In remorse, one also feels sorrow – sorrow for the victim and what he or she has had to endure. Empathy may be heightened and one imagines again and again the humiliation or pain or fear the victim felt. When being punished, on the other hand (again, particularly if one is not remorseful), one might feel sorrow, but for oneself. One might wallow in self-pity, focussing on what a bad time one is having. One thinks of one's own humiliation, pain or fear.

Linked to these aversive emotional experiences is the extent to which one can ignore them. Whilst remorse, if genuine, is something one has to face (to ignore it would put into question whether one really accepted and understood what one had done) punishment is something to endure and a technique for doing this might be to

detach oneself as far as possible from what one is going through. We return to the etymological association of remorse with incessant gnawing. Remorse is insistent; it demands to be noticed. Punishment, originating from without rather than from within, does not of necessity have this unyielding demand for attention.²⁸ Thus, remorse does not only differ from punishment in its temporal extension, it also differs in its psychological insistence. In this way, remorse goes deeper than an experience of punishment.

Whilst it is conceivable that an offender could in some sense take himself away from prison in his mind, it is much harder to take oneself away from the experience of remorse, if it is genuine. Even if one does not manage to detach oneself from one's punishment situation, one can still be angry at, or defiant in the face of, one's punishment. The genuinely remorseful, on the other hand, have to accept and commit to the experience of their remorse. There is no mental escape from oneself.

Sense of desert

Another way in which remorse and punishment might be phenomenologically similar is that they both can feel 'deserved' (this is almost always true of remorse, and sometimes true of punishment). Meyer (2005: 9) considers this to be the case, saying, '[r]emorse in many ways parallels retribution. Like retribution, remorse is a boomerang of the crime. The offender gets his own law thrust back on himself'. Meyer's metaphor of a boomerang illustrates the causal connection between the wrongdoing and remorse and/or punishment: both of these experiences are as a direct result of offending. The remorseful offender believes he deserves to feel bad for what he did – that is part of fully understanding the reality of his actions. When it comes to

²⁸ Torture (the infliction of intense physical or psychological agony) would be closer to remorse in its psychological inescapability, but I am not extending my discussion of punishment to include torture.

experiencing punishment, even if an offender does not feel remorseful, he may still acknowledge that he was aware that what he was doing was illegal and that certain things would come as a consequence if he were to be caught. There might still be some sense in which he thinks ‘fair enough, I was warned, I took the risk’.

There is still a difference, though, between the experience of a sense of desert or inevitability that is involved in remorse and that which might be involved in punishment. The origins of the experiences are different and this tracks the distinction between a possible form of natural punishment and punishment proper (Teichman 1973).²⁹ Remorse comes from within the offender as a consequence of his realising the moral reality of what he has done, whereas punishment is imposed on the offender, whether he feels it is deserved or not. Teichman (1973: 344) explains: ‘remorse is a possible reaction of the agent himself to his action: punishment is (partly) a reaction of others. Thus each can be regarded as the consequence of an action.’

Noteworthy is that she goes on to say that ‘remorse can be appropriate or inappropriate ... Punishment too, can be fitting or unfitting’ (ibid.). Whilst these (appropriateness and fittingness) are objective assessments it is likely that they too operate in a subjective evaluation of the deservedness of one’s remorse and/or punishment. It is, however, probably very rare that one would lucidly experience remorse whilst also strongly believing it to be inappropriate (such an evaluation should (but perhaps does not always) serve to alleviate the emotion (see Chapter Three)). Thus, remorse (where present) probably nearly always feels deserved to the offender, even if others judge him to be ‘overreacting’. Punishment, on the other hand, may often be judged by the offender to be unfitting, even if he is aware that it was a

²⁹ The idea of remorse as natural punishment will be further explored in Chapter Six, where I ask the inverse to the question central to this chapter: Can remorse constitute or substitute for any of the deserved punishment?

likely consequence of his offending. In such cases, the sense of ‘deservedness’ will be absent from his experience of punishment.

It would seem then, that the main phenomenological similarities between remorse and punishment are a sense of being trapped and the pervasiveness of negative affect. It has been argued, though, that the objects and nature of this negative affect differ between experiences of remorse and punishment. Whilst the remorseful feels sorrow for the victim and anger (and even horror) at himself, punishment can precipitate self-pity and anger at others. Whilst an experience of being punished will for the most part involve intense frustration, genuine experiences of remorse are not experiences that one battles *against*, even if one must in some sense battle through.

The relative values of these various ways of ‘feeling bad’, precipitated by remorse and punishment, will be assessed below when I ask the next question:

Even if the phenomenology of remorse and punishment are not the same, are the elements we value in remorse also present when an offender experiences punishment?

Lucid understanding of the wrongdoing and the suffering of victim

In Chapter Three I argued that one of the most valuable elements of remorse is that it provides evidence of and to some extent constitutes lucid understanding of the wrong that was perpetrated and the harm done to the victim. In this way, remorse can be said to be revelatory. The experience of remorse both is the revelation and serves to further clarify the reality of the wrongdoing to the remorseful person. Can punishment substitute for this valuable understanding? Obviously undergoing punishment does not constitute understanding in itself. Whilst experiencing remorse *is* a painful understanding, experiencing punishment does not constitute a moral revelation.

However, we can ask whether it might lead to understanding, which might perhaps then be accompanied by or turn into an experience of remorse.

Punishment perhaps draws the offender's attention to the fact that he has broken the law and prompts him to think about the wrong that he did. This may be true in some cases but scepticism about whether this is generally the case is certainly understandable. Nietzsche (1992) voices such scepticism in his *On the Genealogy of Morals*. Whilst making the value judgement that pangs of conscience are actually not a good thing, he also presents the empirical hypothesis that punishment hinders the emergence or persistence of these pangs, saying that 'prisons and penitentiaries are not the kind of hotbed in which this species of gnawing worm is likely to flourish'. He explains why this might be:

Generally speaking, punishment makes men hard and cold; it concentrates; it sharpens the feeling of alienation; it strengthens the power of resistance. If it happens that punishment destroys the vital energy and brings about a miserable prostration and selfabasement, such a result is certainly even less pleasant than the usual effects of punishment – characterized by dry and gloomy seriousness. [Quoted in Morgan 1992: 1270]

The idea of strengthening resistance mirrors my suggestion above that one may try to detach oneself emotionally from what one is going through when being punished. He also describes what might happen if one fails to become emotionally detached, characterised by self-pity. Nietzsche concludes that experience reveals that it was precisely through punishment that the development of the feeling of guilt was most powerfully hindered. Instead, 'the person upon whom punishment subsequently descended, ... like a piece of fate, suffered no "inward pain" other than that induced by the sudden appearance of something unforeseen, a dreadful natural event, a plunging, crushing rock that one cannot fight' (ibid.).

Some more contemporary thinkers have reflected this sentiment. Teichman (1973: 344) explains that ‘some people will argue that prison is not the right place to send criminals because hardship turns a man’s thoughts away from remorse towards self-pity. Maybe, then, they believe that feelings of remorse are a better, that is to say, worse (more painful) punishment than imprisonment.’

But, here we need to distinguish between the value of the ‘pain’ involved in remorse and the understanding which it accompanies. The ‘inward pain’ to which Nietzsche refers may indeed be more painful than that inflicted in punishment but the pain *per se* is not what we value about remorse. We do not approve when an offender feels remorse simply because he must be experiencing pain. It is to this purported value of remorse that we now turn, asking whether punishment can substitute satisfactorily for it.

The offender feels bad

In Chapter Three I argued that the affective element of remorse was in some sense valuable. However, it was not valued for being painful *per se*. If this had been the case – if what we cared about was the offender simply having a bad time – then punishment could do this. Rather, the affective element of remorse was valuable in as much as it was part of the process or experience of fully recognising what one had done. Its particular emotional quality was also seen to be valuable in itself, as constitutive of ‘moral concern’ (see also Stark 2004).

That the negative affect that results from being *punished* has this revelatory capacity (or is itself part and parcel of revelation) has been shown to be unlikely. Messages of censure – which condemn the offending behaviour as something bad –

are conveyed through the 'painful' medium of punishment, but this is not the same pain as when a lucid understanding of what one has done inherently feels bad.

Hanna (2009) explains how the pain of remorse not only makes an offender's claims to understand the wrongful nature of her behaviour more credible, but also constitutes part of the moral concern which they demonstrate:

When an offender claims to believe that her act was wrong and claims to be sorry, her claims will be more credible if she seems pained by what she did. These feelings indicate sincerity. But they also indicate something else: a certain kind of concern. An offender who is pained by her wrongful act demonstrates the right sort of moral concern. Being pained in this way shows that she cares about doing what is right and about not doing what is wrong. [239]

Thus we can draw a sharp distinction between being pained out of concern for another and being pained out of concern for oneself, or even pained *per se*. Whilst the pain of remorse *is* understanding and concern, the pain of punishment involves frustration and self-pity. Since it is the fact that the pain of remorse constitutes understanding and concern which leads us to see it as valuable, this value cannot be attributed to the pain of punishment that is not so constituted. The aim to impose suffering or harm *per se* is essential to punishment, whereas it happens that remorse is comprised of certain painful beliefs and feelings. We value these feelings as they are indicators and instances of one's sincere understanding and concern.

So far we have been considering the value of the pain of remorse in comparison to the value of the pain of punishment, where we have one or the other. However, Meyer (2005) raises some very interesting questions about the potential interaction between these two forms of pain, suggesting that far from substituting for the pain of remorse, punishment might actually serve to alleviate it. She suggests that when a remorseful offender is punished, 'the physical pain seems to blunt the mental pain, to ease the biting conscience. Physical pain may seem less painful than remorse,

to replace it and externalise it, literally, to ex-piate it. Physical pain substitutes for mental agony, at least for a while' (2005: 7).

The idea here is that anything is better than suffering remorse, so if one has the discomfort of one's punishment to focus on, one might be better able to deal with one's mental state. In addition to being distracting, there may be some comfort in the pain of punishment due to the remorseful offender's sense of it being deserved, as discussed above. In getting what one deserves, one can perhaps feel a bit less bad about what one did. It is often said that the remorseful desire to be punished. Thus, and returning to Nietzsche's ideas, it could be suggested that for those who already feel remorse, punishment can actually offer some relief whereas for those who feel no remorse, punishment can serve to harden and exacerbate defiance.

However, we must be clear about the difference between the alleviation of remorse and distraction from remorse. When distracted from one's remorse it is nonetheless still there. However, perhaps there is only so much pain a person can feel and the pain of punishment has been suggested to be easier to deal with. Alleviation, on the other hand, would mean that the remorse was reduced or extinguished, not just covered up. Allowing punishment to alleviate one's remorse might suggest that the remorse did not really run so deep. Punishment does not change what one did. Perhaps punishment *can* alleviate guilt (see Chapter Two), where one feels one has broken the rules and fears punishment: if the punishment is forthcoming, there is nothing more to fear. Remorse, on the other hand, should not be too easy to alleviate, if genuine. That being said, and as I intimated above, being punished might go some way to satisfying the need to *do something*, perhaps by way of making amends. This desire to make amends is the third valuable aspect of remorse that we shall consider.

So, it seems that not only does the pain of punishment fail to substitute for the pain of remorse (as only the pain of remorse is intimately tied up with the feelings, thoughts and beliefs that we value), it might in fact distract from, or somewhat alleviate, the pain of remorse, where it was already forthcoming for the offender.

The offender's desire to make amends

The other valuable element inherent to remorse that I want to consider is the desire to make amends. In Chapter Two I explained that genuine remorse motivates the wrongdoer to apologise and try to find ways of 'making up' for his wrongdoing. However, I also emphasised that it is the tendency towards these actions that is necessary for genuine remorse, and not their successful completion. There are many factors external to the offender that could prohibit or otherwise make difficult his making amends. Moreover, a person could perform these sorts of behaviours out of a sense of dogged duty, really rather wishing he did not have to bother. Thus, we can see that, although reparative action (however motivated) can be beneficial for a variety of instrumental reasons, it is the *desire* to apologise and make amends that we find intrinsically morally valuable in remorse.

This desire to make amends is borne of the lucid understanding of one's wrongdoing and concern for the victim – both experiences being central to remorse. I have already shown that such understanding and concern are not likely to be products of punishment and certainly are not internal to it. However, whilst punishment may not be able to substitute for the remorseful offender's *desire* to make amends, it may be able to substitute for the more consequentially valued *act* of making amends. This possibility will be considered within the next question, when I ask whether

punishment can substitute for any of the utilitarian benefits that tend to stem from remorse.

In sum: both punishment and remorse do label things as ‘wrong’. However, whereas remorse facilitates and *is* a clear understanding of what one has done and its wrongness, punishment labels behaviour as bad in a way that the offender can contest. Thus, Teichman (1973: 345) explains that ‘[r]emorse itself is a form of understanding as is partly shown by the fact that you cannot feel it and at the same time ask: “Why am I feeling this?” A person can be punished without knowing why, of course, but one might say that laws and punishments are amongst the things which reflect this or that society’s understanding of good and evil’. Punishment can present some actions as bad to the offender but there is no guarantee that he will accept this evaluation of his actions. The truly remorseful cannot logically reject the evaluation implied by the experience of remorse, whereas the simply punished can reject the evaluation of society.

The only element that punishment can substitute for is that of feeling bad or ‘pained’. But this pain has no intrinsic value when it is not borne of/part of remorse. We do not just want the offender to feel bad: we value the bad feeling in remorse because it constitutes and serves to clarify the offender’s understanding of the wrong that was done and because it is constitutive of moral concern. This being said, a classical retributivist might value the offender’s pain from punishment as satisfying pure retribution. This is, however, postulating a conception of retribution as revenge, which is generally considered a very unattractive position for liberal states.³⁰

³⁰ However, consider also Gardner’s (2005) thesis (discussed in Chapter Three) that punishment (amongst other things) serves to ensure that the wrongdoer feels as subjectively ‘low’ as he is objectively ‘low’, where remorse is not forthcoming. He seemed to present this lining up of the subjective with the objective as intrinsically valuable.

We now turn to the next question in our project of determining whether punishment can in anyway substitute for remorse.

Is there *anything* punishment can achieve as a substitute for remorse? Are the (more tangible, utilitarian) consequences of remorse also achievable through punishing?

Even if punishment cannot substitute for any of the aspects of remorse that are of intrinsic value, it could be the case that remorse is also of value instrumentally, in ways that punishment *can* substitute for. What might these ways be? I suggest that the instrumental value of remorse lies in three main consequences usually associated with it: 1) the intuition that remorse results in a reduced propensity on the part of the offender to re-offend, 2) that knowledge of the offender's remorse, especially when accompanied by a sincere apology, can be of comfort to the victim, and 3) that the reparative behaviours associated with remorse go some way to making amends to society. We can now ask whether punishment can fulfil any of these positive functions of remorse.

Remorse leads to a reduced propensity to re-offend

There is an understandable intuition that an offender's remorse (and all the painful understanding associated with it) will be instrumental in dissuading him from re-offending. Because he has woken up to the reality of the consequences of his actions, he now commits himself to refraining from harmful behaviour in the future. As will be discussed in Chapter Six, the empirical evidence for a link between remorse and desistance is inconclusive. Cox's (1999:17) discussion of the evidence leads him to conclude that '[t]he disturbing fact is that there appears to be no firm evidence that

[had the remorseless offender] shown remorse, there is less likelihood of him re-offending'. Thus, although it is part of being remorseful that the individual commit himself to being a better person, we cannot unequivocally assert that this commitment is always successful.

It is much clearer that punishment does fulfil this function. Deterrence is one of the key utilitarian justifications for punishment: the world is a better place when there is less crime and the pain inflicted on the offender through his punishment is outweighed by the benefit to society accrued from effectively deterring crime. However, whereas remorse can at most serve to 'deter' the specific offender, punishment extends this function to achieve general deterrence. In punishing the offender for his offence, others are provided with evidence of what would happen where they themselves to offend.

It would therefore seem that in punishing those who break the law, a state aims to exert a more extensive deterrent effect than an individual's remorse could ever logically achieve: the whole of society's propensity to offend is reduced, rather than just the remorseful offender's. Moreover, it is questionable that remorse is strongly associated with complete desistance anyway. Indeed, one of the three retributive theories (von Hirsch 1993) that will be examined in the remainder of the thesis proposes that punishment is required as an additional prudential disincentive to offend, as normative reasons (which we imagine are readily supplied in an experience of remorse) are often not compelling enough.

It would seem, then, that rather than punishment being a substitute for any deterrent effect of remorse, it more obviously and successfully serves to deter, supplementing rather than substituting.

Remorse and apology make the victim feel better

The second possible consequence of remorse that is of value is any relief that it can bring the victim, especially if accompanied by a sincere apology. This has empirically been shown to be the case, even when the offence was serious. In fact, Ohbuchi *et al.* (1989) actually found that as harm increased, victims were more interested in hearing an apology. Part of the explanation for this was that more severe harm precipitated more aggressive feelings in the victim and the apology served to calm these feelings.

Murphy (2003) argues that victims often perceive part of the harm of the offence to be the insult or degradation that the offence communicates. Through the offence, the offender sends a message to the victim, saying that the victim is inferior and exploitable. This causes the victim distress. He may even feel that he *is* inferior and exploitable, and blame himself. However, if the offender repents, then the message is withdrawn and some of the harm lessened.

In a similar vein, Weiner's (1992) attribution theory can be used to explain how offenders' apologies enable victims to avoid feelings of hopelessness and increase their self esteem, thus reducing harm. Being the victim of crime can result in self-blame and negative feelings towards one's self. This is plausibly due to attributing the causes of the crime to internal and stable factors: victims see the cause of the crime as being related to their enduring character, not to something external. What the offender's apology and display of remorse might achieve is to shift the blame to the offender. The victim no longer feels responsible.

Whilst it is important to point out that it is the overt display of remorse, accompanied by a direct apology to the victim that is capable of having this comforting effect, it is essential to remorse that it tends towards such a display. Thus, we can legitimately attribute this desirable effect to remorse, as a possible and even

likely consequence of its being experienced. The question now is whether punishing the remorseless offender can substitute for the comfort the victim feels on receipt of knowledge of the offender's remorse or a sincere apology. I suggest that there is a sense in which punishing the offender makes the victim feel 'less bad', but that this relief is qualitatively different from that precipitated by an apology, and that it operates through different psychological mechanisms.

The desire for revenge is very basic, even to those who quickly and consistently manage to keep it in check. Getting revenge is often described as 'sweet' and this is explained by the satisfaction experienced when someone who hurt you is hurting. In this way, knowing the offender is being punished can make the victim feel less bad and perhaps even gleeful. However, there is another, less basic response that the victim can have to the offender's punishment that might also have some positive elements. The feeling that justice is being done can be of comfort to victims or relatives of victims of serious offences. This is different from a basic enjoyment of another's suffering – victims often report a feeling of 'closure' when their offender is convicted and sentenced: given what happened, things are now as they should be and the victim and/or victim's family feel they can move on. This will usually involve no glee, but often some relief.

The second of these two ways in which victims might respond to their offender's punishment is arguably the more morally preferable, although both are understandable. Regardless, it is the case that punishment can, to an extent, make the victim feel less bad. This is not through the retraction of messages of degradation or realigned attributions, but through pleasure at the offender's suffering or attaining closure following perception of justice.

So, just as remorse and punishment both make the offender feel bad but in different ways and for different reasons, remorse and punishment can both make the victim feel better, in different ways and for different reasons.

Remorse motivates efforts at reparation

The third potential benefit of remorse is the reparative efforts that it often motivates. In Chapter Two I discussed how genuine remorse involves a desire to make things up to the victim or otherwise make amends. Above, it was argued that this desire was of moral value. However, the actual act of offering reparation has instrumental value in providing the victim with compensation or society with hours of community service or whatever else the offender sees fit to commit himself to.

There are modes of punishment that can be seen to quite neatly substitute for the reparative behaviours of remorse. In England and Wales, there are three types of reparative orders that can be imposed on the offender alongside his proportionate punishment. Compensation orders can be imposed for violent offences, restitution orders on conviction of a theft offence (requiring the offender to restore to the victim the stolen property or good representing that property or an equivalent sum of money), and reparation orders (a form of community sentence available only when sentencing under 18s) (see Ashworth 2010: 361). In addition to reparative orders, community sentences have a ‘making amends’ quality to them in as far as they are beneficial for society and often require of the offender the sorts of voluntary work he might take on himself, were he to be genuinely remorseful.

Whilst it is clear that, in terms of benefiting the victim or society, these sorts of sanctions can substitute for the reparative efforts that flow from remorse, there is a further, more abstract understanding of ‘making amends’ that punishment that we can

consider. Giving a victim money or working for free in a hospital cafeteria are obviously of instrumental value: the victim has more money than he otherwise would and the hospital can save on wages or provide more efficient service. Some theorists have argued, however, that punishment of any kind (including that which does not obviously benefit a particular individual or organisation) allows the offender to make amends symbolically (Bennett 2008), or serves as an enforced penance, necessary for acceptance back into the community (Duff 2001). This highlights the more abstract value of the reparative behaviours that remorse motivates – the remorseful offender seeks atonement through them.

If the remorseless offender is not motivated himself to restore his status in his community then perhaps punishment substitutes for the work that he would have done to achieve this, had he been remorseful. Bennett (2008) argues that punishment should be understood (and is best justified) through the symbols of blame and apology, where punishment is the institutional analogue of making amends, required for the institutional analogue of forgiveness. Duff (2001) argues that punishment should be understood as a penance, necessary for reintegration into the community.

It might seem, then, that punishment can substitute very well for the penitentiary and reparative behaviours that are motivated by remorse. The next question will examine this possibility more closely.

Perhaps punishment can only significantly substitute for the penitentiary behaviours motivated by remorse: Is punishment a substitute for making amends?

Of all the possible ways in which punishment might substitute for remorse – from its phenomenology to its morally valuable elements to its more external benefits – it has

been the capability of punishment to emulate remorseful reparative behaviour that proved the strongest candidate. Punishment can provide and constitute both literal and symbolic reparation to the victim and community.

The two theorists mentioned above rest significant parts of their justifications of punishment on this idea. Although they will be discussed in greater detail in the remainder of the thesis, the ideas pertinent to the present topic are as follows.

Duff argues that criminal punishment should be understood and justified as a species of secular penance. He explains that penitential punishments ‘constitute a forceful and weighty kind of apology, which should reconcile the wrongdoer with those she has wronged’ (Duff 2001: 109). He explains that for purposes of reconciliation with the community, it does not matter if the apology that the penance represents is sincere – the offender’s fellow citizens should accept that she has completed her punishment without inquiring into her reasons for doing so: ‘She has now paid her debt’ of apology and reparation’ (ibid.: 110).

It may well seem, then, that Duff would be sympathetic to a conception of punishment as substituting for the reparative behaviours of remorse. However, to argue that he would accept this conception would be to misunderstand the subtleties of his position. To see this we need to look more closely about what he thinks the relationship between remorse (repentance, in his language) and punishment is. For him, the secular penance does not substitute for remorseful reparation; rather, he believes punishment is the *only* way to achieve full reconciliation with the community. He believes it is not possible to make amends to a great enough and public enough degree unless punished. Punishment, then, does not substitute, it constitutes. Further, Duff also believes that punishment as secular penance is the mode through which an offender can become fully repentant. In this sense, it is necessary for remorse, rather

than a possible substitute for it: ‘As fallible moral agents, we need such penances to assist and deepen repentance’ (ibid.: 108). Thus, in this sense also punishment does not *substitute for* remorse as it would be seen to be inadequate if remorse is not successfully inspired in the offender.

Later, in Chapter Six, I will argue against the necessity Duff ascribes to punishment for both remorse and adequate penance when I consider whether remorse can constitute any of the deserved punishment. For now, however, we have seen that there might still be some difficulty in asserting that punishment can substitute for remorseful reparation in this way.

However, Bennett (2008) is a theorist whose conceptualisation of punishment as the offender’s making amends is more compatible with an argument of simple substitution. He believes that the ideal scenario is where the offender (or wrongdoer) makes amends of his own volition – amends that (contra Duff) would be adequate if they were forthcoming. Punishment as amends constitutes an apology ritual, necessary because many offenders do not seek to make amends and because it must be possible for it to be successful regardless of whether it is sincere or not. One of the key differences between Bennett and Duff is that he does see punishment as a substitute for the amends that an offender would be motivated to make, were he to be truly sorry and seeking to apologise.

It could be argued that it is crucial to recognise that there is still something importantly morally different between having reparation/penance/amends imposed and offering them oneself. Hanna (2009: 254) asks whether ‘in making amends, [offenders] must aim to harm themselves or make themselves suffer? Must they make amends in order to do these things to themselves?’ Her response is that this is not the case:

No. For they could make amends, not because of any suffering or harm it may cause them but *in spite of* any such suffering or harm. And the willingness to make amends despite the harm and suffering incurred can have powerful expressive value. It can serve to show that the offender is truly sorry for what she did, that she sincerely cares about righting her wrong and so on. [ibid.]

The argument is that offenders do not make amends to feel pain but despite the pain, for other, expressive reasons. It cannot, then, be the *pain* of punishment that constitutes the amends or penance and it cannot express in the same way. Even if punishment can facilitate reconciliation through serving as a means of reparation, if it is not endured in a spirit of remorse it can at best be an adequate substitute, and not the ideal. Even Duff concedes this, as his hope is that in serving the penance the offender will come to full repentance and embrace it as his own: ‘What began as a punishment inflicted on him in order to induce repentance becomes a punishment (a fully fledged penance) that he accepts or wills for himself as an expression of that repentance’ Duff 2001: 111). It is remorse that renders the penance fully fledged.

So, it does seem that punishment can substitute for the reparative behaviours motivated by remorse, and that this is held to be the case by some retributive theorists – indeed, it is necessary for their accounts. However, it has also been pointed out that as a substitute, the amends made by way of punishment alone will be morally inferior to those attended or motivated by remorse, even if the former is adequate for reconciliation with the community. It must also be remembered that punishment has not been shown to adequately substitute for any of the emotional-cognitive elements of remorse, and that what it can substitute for is restricted to the ensuing behaviours of remorse.

By way of summary we can ask the following:

What is left out when there is punishment but no remorse?

With respect to the phenomenology of remorse it was argued that the mental inescapability which characterises remorse was not essential to being punished. One might succeed in detaching oneself from the experience of one's punishment but this is impossible in remorse, if it is genuine. Nietzsche's depiction of men becoming emotionless in prison as a method of self-preservation is not something open to the remorseful offender. It is logically inconsistent to be experiencing an intense emotion (which remorse is by definition) and to at the same time be indifferent to it. Remorse was seen to be the more thoroughly isolating: the punished can seek comfort for their suffering from the knowledge that others are also going through the same thing. The remorseful, however, should not find consolation in the knowledge that others have also done wrong - they are radically singular. Further, the sorrow and anger involved in an experience of remorse are directed at different objects when being punished, if felt at all.

In terms of the moral value of remorse, nothing could be substituted for by punishment. Punishment will not necessarily lead to the offender understanding the wrongdoing and certainly does not constitute it. The resolve to become a better person might be prompted by punishment but, without remorse, this would amount to deciding to avoid punishment in the future, and not stem from normative reasons for compliance.

There were some utilitarian benefits associated with punishment that were similar to those stemming from remorse. However, they were not exactly the same. The relief that might be given the victim when he hears of the offender's remorse and/or receives an apology cannot be provided by punishment, although feeling avenged and/or relief at justice being done can follow.

The things that punishment can substitute for were seen to be:

- 1) An aversive experience for the offender. However, (non-specific) discomfort *per se* was shown not to be what we value in remorse. Rather, remorse's *particular* emotional quality is essential for a thoroughgoing understanding of what one has done and is constitutive of moral concern, both of which are of value.
- 2) A form of reparation or penance, which is not strictly necessary in remorse but ideas of which should be present in the offender's mind as something he wants to do. It was argued that punishment, although adequate to reconcile the offender with the community, is unable to substitute for the offender's *desire* to make amends.

Having seen that, on the whole, punishment is a poor substitute for remorse, it seems naturally to pose the following, final question:

Is substituting for remorse what we are trying to do when we punish anyway?

It will be seen in the following chapter, where I set out prominent versions of (broadly) retributive theories of punishment that substituting for remorse is not the justification for punishment nor is it an aim. The theory that comes closest to linking punishment to remorse is Duff's, discussed above. However, as was explained, his theory relies on the idea that punishment *induces* repentance rather than provides another route to the valuable elements internal to remorse. Through imposing punishment we are not trying to substitute for remorse but to facilitate its emergence. This being said, the reparative behaviours motivated by remorse have been argued to be substitutable for to an adequate degree when they are imposed on the offender by way of punishment.

Again, for Duff, punishment is not substituting for remorse here but supplements the behaviours that are motivated by it in order that they can be adequately reconciliatory.

Another theory (that of von Hirsch (1993)) that will be explored posits that punishment provides an *extra* prudential disincentive to supplement normative reasons not to break the law (this is mostly thought of in general deterrent terms). The general deterrent function of punishment is something in addition to anything remorse achieves and so again punishment can be seen to supplement something that remorse begins to achieve.

However, what will become clear throughout the thesis is that for both these theories the main function of punishment is to censure the offender, communicating publicly on the subject of the wrongfulness of the offending conduct.

Other retributive theories – those concerned with correcting the unfair advantage the offender obtained through committing his offence – will be seen to be concerned not with censure nor making amends nor supplementing normative reasons we have for compliance. Such theories are even further from supporting a conception that punishment is intended to substitute for remorse.

Conclusion

Having completed our exploration of the nature and value of remorse, we can now move onto engaging fully with the institution of state punishment. The relevance of remorse to punishment can be posed as the question of whether it justifies any alteration to the severity of the sanction the offender receives. Developing a justification for this mitigating role will be the project of the remainder of the thesis. Being so, I will now turn to the task of introducing the concept of mitigation in the context of state punishment, with a focus on retributive punishment in particular.

Section Two:
Remorse and Retributive Punishment

Chapter Five

Mitigation and retributive punishment

Remorse features in the sentencing guidelines for England and Wales as a mitigating factor.³¹ However, whilst mitigation is generally associated with a reduction in sentence severity, how different mitigating factors precisely operate and what justifies their doing so can be difficult to untangle conceptually: at what stage of the sentencing decision-making process do they feature and for what reasons?

Overview

The aims of this chapter are as follows. First, I will explain why the concept of mitigation needs unpicking, suggesting that it actually has distinguishable uses which are not always coextensive. I will then go on to outline the various ways in which factors can end up affecting the severity of the sentence, pinpointing which of these can appropriately be labelled ‘mitigation’. As a result of the non-uniform operation of these factors, I will suggest that they are more carefully understood when exploring mitigation in theory as ‘sentencing factors with the potential to mitigate’, rather than ‘mitigating factors’. Next, I will explain why mitigation is a particularly retributive concern, and thus why the search for justifications for remorse-based mitigation is better focused on this tradition of penal philosophy. With this in mind, I will outline the main tenets of retributive thinking, introducing three contemporary versions of retributive (desert) theory. Finally, returning to the schematic I developed (presenting

³¹ As well as appearing in the sentencing guidelines for specific offences, remorse is identified as a relevant mitigating factor in *Overarching Principles: Seriousness* (Sentencing Guidelines Council 2004: 7), which is intended to cover all offences: ‘When the court has formed an initial assessment of the seriousness of the offence, then it should consider any offender mitigation. The issue of remorse should be taken into account at this point along with other mitigating features such as admissions to the police in interview’.

the various possible effects of sentencing factors, indicating which should be considered as instances of mitigation) I set out the possibilities for remorse. These possibilities will be explored in the remainder of the thesis, with close reference to the three desert theories introduced.

Ambiguity surrounding the concept of mitigation

Within sentencing discourse, ‘mitigation’ can describe various, distinct, penal circumstances. For any discussion of the legitimacy of a mitigating factor, it is important to be clear about what these various circumstances are, and how they differ. In everyday discourse ‘to mitigate’ roughly means to make less harsh or severe, and its more technical usage in punishment discourse is similar. Under the heading ‘mitigation’, in the *Oxford English Dictionary*, there are two entries invoking mitigation in the context of sentencing. The difference in the two definitions is indicative of the scope of the concept. One reads ‘to reduce the severity of (a punishment)’. The other reads ‘to extenuate or lessen the gravity of (an offence, a wrongdoing)’. Mitigation of the gravity (or seriousness) of an offence is not conceptually identical to mitigation of punishment, although the former often leads to the latter. A factor could make the offence less serious but a judge might, for whatever reason, refuse to reflect this in the sentence he passes. The reverse inference – reduced sentence means reduced seriousness – also cannot be made. The factors or circumstances that led to a reduction in the severity of punishment an offender received may be unrelated to the seriousness of the offence (the guilty plea discount, for example). They might, nonetheless, still be appropriately described as mitigating.

Wasik (1982: 524) argues that there is a ‘scale of excuse, running downwards from excusing conditions, through partial excuses to mitigating factors’. In what

follows I will explore this scale, enumerating all the ways facts about an offence or offender could reduce the severity of the punishment received. On the assumption that being clear about what mitigation is not might further elucidate what it is, I will start from the top of this scale, with ‘excusing conditions’ and exculpation. Once we are clear about the distinct (although sometimes overlapping) possibilities, we will be in a position to understand better when the language of mitigation is really appropriate.

In the following, I will use examples predominantly from sentencing practice in England and Wales to help illuminate the distinctions I draw. This should not prevent the discussion being of relevance to any sentencing system: the aim is to *abstract* the various mechanisms through which facts about an offence or offender can have a moderating effect on the severity of the sanction imposed, revealing which of these is correctly described as mitigation. The discussion is not intended to cover all or even most mitigating factors. Rather, the possible mechanisms for mitigation are enumerated identify those through which remorse in particular might operate to reduce an offender’s sentence.

Possibility one: exculpation

One possibility is that there are facts surrounding the purported offence that lead to a defendant not being held liable. In England and Wales, such a possibility would involve the demonstrable presence of at least one of the established criminal law defences, such as insanity, duress, necessity or mistake of law. Depending on the facts, these defences will serve to either justify or excuse the actions of the defendant. If successful the defendant will receive no punishment at all, as he will not be convicted.

In this scenario, facts or circumstances surrounding the cases result in the individual involved receiving less punishment than he would have done had the facts

or circumstances not been present – indeed, he receives no punishment at all. However, this is not an instance of mitigation. The offence was not made less serious as no offence was committed. Nor was the punishment mitigated through any other mechanism as there was no punishment to mitigate in the first place. So, here we have a scenario where, although a factor (insanity, or duress, or necessity etc.) had a traceable and significant impact on the punishment (or absence thereof) the individual received, there is no mitigation.

Possibility two: yields lesser offence

There are occasions when a fact about the offence results in a different, lesser offence having been committed, thus yielding less severe punishment for the offender. These facts can relate to culpability or harm. For example, if an offender intended to kill a person, but that person actually survived the attack, the offence the offender will be charged with would be attempted murder, not murder. A different offender, with absolutely no intention to kill, might start a fight which nonetheless results in the death of the victim. This offender would likely be charged with manslaughter, rather than murder.

Whilst facts about these cases – absence of death and lack of intention to kill, respectively – may lead to less severe punishment, they are not mitigating factors. Rather, they shift the offence into an offence category of (arguably) lesser seriousness, to which the severity of punishment must then be in proportion. Here, absence of death is not mitigating; it is a feature of the standard case of attempted murder and so does not make the offence any less serious.

There may be features of a particular attempted murder that *do* mitigate the seriousness of the offence (perhaps the offender is clinically depressed, for example –

we will come to partial defences below) but these cannot feature in the definition of the offence if they are to be mitigating. To count absence of death as a mitigating factor would be to ‘double count’ this feature of the offence.

This may seem somewhat puzzling as, in the possibilities below, it will be explained that matters relating to harm and culpability frequently *do* constitute mitigating factors. What has been explained above does not show that reduced harm or reduced culpability have no bearing on the severity of the offence, rather, they are already taken account of in the very definition of the offence, and the severity of the punishment reflects this accordingly: the standard (mid-range) cases of attempted murder and manslaughter warrant less severe punishment than do the standard cases of murder.

That these facts of the cases are incorporated in this way is not a necessary feature of sentencing systems. Although not arbitrary, one must remember that offence categories are legally constructed. That there is an offence of attempted murder, wherein absence of death is accounted for in its very definition, is a matter of convention. We can imagine a hypothetical jurisdiction, impoverished in offence categories, where there was simply an offence of extreme violence towards another person, with no separate offences depending on harm caused or offender intention. In such a jurisdiction, absence of death may well count as a mitigating factor. Whether the offender intended to kill would also have to be taken into account in mitigation (if he did not) or aggravation (if he did). In England and Wales, in contrast, this evaluation of intention is the main determinant of whether the offender is found guilty of murder or manslaughter. So, again, here, this particular factor is already taken into account by the offence.

Thus, when a fact about the offence or offender entails that a different offence category is applicable, it is not a mitigating factor. To treat it as such and allowing it to reduce the seriousness of the offence from that of the standard case would be to 'double count it'. It has already diminished the seriousness ascribed to the offence (and hence the proportionate punishment) through the mechanism of yielding conviction of a lesser offence.

Possibility three: a determinant of offence seriousness

Under the sentencing guidelines in England and Wales, judges must identify the appropriate 'starting point' for the offence before adjusting it to reflect any mitigation or aggravation in order to maintain proportionality. Thus, the first task for the judge is to make a decision about the seriousness of *standard* cases of the particular offence before them.³² In England and Wales, judges are often aided in this task by descriptions in guidelines of various levels (or categories, which are ordered in terms of seriousness) of a particular offence. They can then make an assessment of which level or category is most appropriate to the offence with which they are dealing. Having identified an appropriate offence category, the court adopts the starting point for that category before proceeding through the remaining steps of the guidelines methodically.

For example, the sentencing guidelines for the offence of 'assault occasioning actual bodily harm' divide the offence into three categories.³³ Category One covers the most serious cases of this type of assault, where there is greater harm (serious injury must be present) and higher culpability. It has a starting point of 1 year and 6

³² The Coroners and Justice Act (2009) 121: 10 requires judges to do this.

³³ The guidelines for this offence appear within the Sentencing Council's (2011) *Assault: Definitive Guidelines*.

months' custody. In order for a particular case to fall into this category, it must have at least one factor indicating greater harm and at least one factor indicating higher culpability. These category-determining factors 'comprise the principal factual elements of the offence' (2011: 12). Examples of factors indicating greater harm include: injury which is serious in the context of the offence, the victim being particularly vulnerable because of personal circumstances, and sustained or repeated assault on the same victim. Factors indicating higher culpability include (but are not limited to): a significant degree of premeditation, use of a weapon, and a leading role in the group.

Category Two covers cases where there is greater harm (serious injury must normally be present) and lower culpability *or* where there is lesser harm and higher culpability. The starting point for this category is 26 weeks' custody. The same factors as above indicate greater harm and/or higher culpability. Lesser harm is indicated where there is injury which is less serious in the context of the offence. Factors that indicate lower culpability include: a subordinate role in the group, a greater degree of provocation than normally expected, lack of premeditation and mental disorder and learning disability.

Category Three covers cases where there is lesser harm and lower culpability, indicated by the same factors as above. The starting point is a medium level community order.

Judges are provided with ordered levels or categories such as these for most offences in England and Wales to help determine the seriousness of the offence. As has been seen: when determining which category is appropriate – and therefore also the starting point for the seriousness of the offence – facts about the offence and offender are taken into account. So, if particular facts surrounding the offence position

it in a level or category of lesser seriousness than if the facts had been different or absent, does this make them mitigating factors? Here we do enter the realms where the language of mitigation starts to become appropriate, although we have not yet reached the penal circumstances that constitute the central instances of mitigation.

In the examples given above, we saw that playing a subordinate role in the group, indicating lower culpability, might move the offence from Category One to Category Two or Three (depending on the level of harm). It might be thought, therefore, that playing a subordinate role is mitigating. However, although it is true that this feature of the offence does make it less serious, it still is not strictly a mitigating factor as it constitutes a principal factual element of the offence – it does not move the seriousness (and accompanying proportionate punishment) away from the starting point, indeed, it sets it.

The aggravating and mitigating factors that are set out in the guidelines that justify an upwards or downwards adjustment *from* the starting point (determined as above) include: abuse of power and/or position of trust, gratuitous degradation of victim, attempts to conceal or dispose of evidence (all aggravating), and remorse, single blow and age and/or lack of maturity (all mitigating).³⁴

However, as will be discussed further below, the determinants of initial (starting-point) offence seriousness can sometimes be hard to distinguish from aggravating and mitigating factors as they all serve to facilitate evaluation of how serious a particular offence was. Thus, the language of mitigation is not entirely misplaced when considering factors that contribute to the determination of offence-seriousness.

³⁴ See Assault: Definitive Guidelines (2011: 13) for the full list of aggravating and mitigating factors.

Possibility four: an adjuster of offence seriousness

The fourth possibility constitutes a central instance of mitigation. Once offence seriousness has been established, facts surrounding the offender or the offence serve to adjust the seriousness of the offence, moving it away from that of the standard case.

Hartman (1984: 368) describes the logic and operation of mitigating factors as follows:

The factors presume that there are offences which are ordinary and those which are extraordinary. Strictly construed, the factors would not apply to an offence which is unremarkable. When an offence deviates from the norm (either in the character of the offender or in the manner in which the offence is committed), application of the appropriate aggravating or mitigating factors should generate a degree of individualisation in the sentence.

The more common mitigating factors include partial defences. In the above section on exculpation, I explained that there are established defences in English Law that serve to excuse or justify the offence, thus precluding the liability of the defendant. Where such factors are present but not strong enough to fully exculpate, they can be presented in mitigation, thus reducing the seriousness of the offence. Diminished responsibility and provocation are examples of partial defences: the offender is still held liable, but with a lesser degree of culpability than he would have been absent one or more partial defence. The criminal behaviour is neither justified nor excused, and the degree of harm caused is not affected, but the case is rendered overall less serious, on the grounds that the offender is less blameworthy.

In the above example of the offence of 'assault occasioning actual bodily harm', age and/or lack of maturity, and delivering only a single blow were seen to be legitimate mitigating factors. The former serves to reduce the offender's culpability, as young people are generally understood to be less responsible for their behaviour.³⁵

³⁵ Exactly why this is so is a complicated issue. See von Hirsch and Ashworth (2005: 36- 41) for a discussion of the arguments. The general ideas are that juveniles have a more limited capacity to grasp the harmful consequences of their actions and have had less time and opportunity to develop self-control.

Delivering only a single blow might be indicative of lower culpability as it suggests the offender was not wholeheartedly committed to the assault. It also suggests that the harm inflicted might (but not necessarily) be lesser than assaults that involve multiple blows. The moderating effect these factors exert on the punishment deemed proportionate constitutes a central example of mitigation in sentencing.

However, as was noted above, sometimes what would otherwise constitute a factor in mitigation is already taken account of when determining offence-seriousness, which is a matter of legal convention. In such cases, the factor would not be taken into account a second time when it came to assessing grounds for mitigation. Indeed, Ashworth (2010: 106) explains:

Many...offences cover broad areas of conduct without legal differentiation: robbery can involve anything from a push to snatch a purse to an armed hold-up of a bank, and the offence of theft has no subdivisions at all according to the value of the property or the circumstances of the offender. It follows from this that the determinants of offence-seriousness are sometimes difficult to separate from the aggravating and mitigating factors.

Even factors that constitute core mitigating factors – partial defences – are sometimes already implicated in offence-seriousness. For example, the offence of manslaughter can take the form of ‘manslaughter by reason of diminished responsibility’ and ‘manslaughter upon provocation’. Here, any mitigation that would be afforded by the partial defence is redundant as it is already taken into account. Indeed, Ashworth (ibid.: 122) describes them as ‘sentences of mitigated murder’.

Possibility five: relevant external to offence seriousness

The above section described examples of factors about the offender or offence leading to reduced punishment that constitute central examples of mitigation. When theorists or criminal justice practitioners talk about mitigating factors, the above is likely what they have in mind. However, factors affecting the seriousness of the offence do not

exhaust those that can legitimately be described as mitigating. There are some factors that provide reason to reduce an offender's sentence justified on other grounds. These might be 'quasi-retributive' or pragmatic. The concept of quasi-retributive grounds for mitigation was proposed by von Hirsch and Ashworth (2005: 174). They explain that these do not form part of the general principles of proportionality in punishment but instead relate to the wider underlying conceptions of penal censure as a response to criminal offending. Such legitimate mitigating factors might include voluntary reparation or the significant passage of time (perhaps several years) between the commission of the offence and sentencing.³⁶ For the most part, pragmatic reasons justify the discount given for a plea of guilty. Thus, the Sentencing Guidelines Council in *Reduction in Sentence for a Guilty Plea: Definitive Guideline* (2007, para. 2.2.) argues:

A reduction in sentence is appropriate because a guilty plea avoids the need for a trial (thus enabling other cases to be disposed of more expeditiously), shortens the gap between charge and sentence, saves considerable cost, and, in the case of an early plea, saves victims and witnesses from the concern about having to give evidence.

In a similar, pragmatic, vein, an offender's co-operation with police and provision of further information can serve to mitigate his sentence, on the grounds that it is beneficial to the criminal justice system to encourage offenders to do so. These sorts of factors can also reasonably be classed as mitigating despite the absence of a link to offence-seriousness. They consistently serve to reduce the severity of sentence on a principled basis.

If one wants to object and claim that only those factors relating to offence-seriousness can rightly be described as mitigating, then this is a matter of semantics.

For example, the Sentencing Guidelines Council's (2008, para 16) guideline on *Theft*

³⁶ The arguments von Hirsch and Ashworth make in relation to quasi-retributive grounds will be explored in detail in Chapters Seven and Eight.

and Burglary in a building other than a dwelling lists the impact of the sentence on the offender's dependency to drugs or alcohol as a mitigating factor. However, the factor's appearance on the list is for convenience and the SGC states that it is not a true mitigating factor but 'may properly influence the type of sentence'. By 'true', it is meant that it does not fit into the category of mitigation of the previous section – it is not related to harm or culpability so does not adjust the seriousness of the offence. It is, however, still mitigating in the sense that applies to factors that fit into the category currently under discussion: those that can be justified on other grounds. The guidelines in this instance state that it may sometimes be appropriate to impose rehabilitation, treatment or supervision requirements 'as part of a community order or a suspended sentence order in an attempt to break the cycle of addiction and offending, even if an immediate custodial sentence would otherwise be warranted' (ibid.). This justification operates outside considerations of offence-seriousness.

Further, as explained previously, mitigation refers to any instance where something is thought to justify the imposition of less severe sentence than would otherwise be appropriate. The sorts of externally justified factors under current discussion usually appear in guidelines under 'personal mitigation'. I therefore choose to position these factors as operating through a second central mechanism of mitigation.

Possibility six: motivates leniency ungrounded in principle

The final possibility is that something about the offender or the circumstances surrounding the offence motivates leniency from the court without there being any principled justification offered – neither internal nor external to offence-seriousness.

We could imagine a hypothetical situation in which the sentencing judge is an art

fanatic with a particular love of the work of the artist who he happens to be sentencing for a fairly serious offence. Although he may not articulate this to himself, we can imagine that he feels the artist to be essentially a ‘good person’ as he is capable of great creativity. He also fears that any significant spell in prison might extinguish the offender’s artistic flare, which would result in no further artwork for people to enjoy. Perhaps he puts the offender’s criminal behaviour down to his ‘artistic temperament’, an idea which he romanticises. None of these reasons would be recognised by sentencing theorists as legitimate grounds for reducing the artist’s sentence. An argument that the artist contributes to society and therefore must be kept out of prison would likely not convince many, and would certainly move away from proportionality-based arguments for mitigation. Notions of all artists being worthy people or that bad behaviour borne of ‘artistic temperament’ is something to celebrate certainly would convince none. But, if the judge nonetheless reduces the artist’s sentence from that which he would have given a non-artist, there is still a sense in which the offender’s talent mitigated the sentence. It was this fact which caused the judge to be lenient. If we return to the *OED* definitions presented towards the beginning of the chapter, we can recall that one of the definitions for the verb ‘to mitigate’ was ‘to reduce the severity of (a punishment)’. The facts about the offender, irrelevant as they may be, reduced the severity of his punishment.

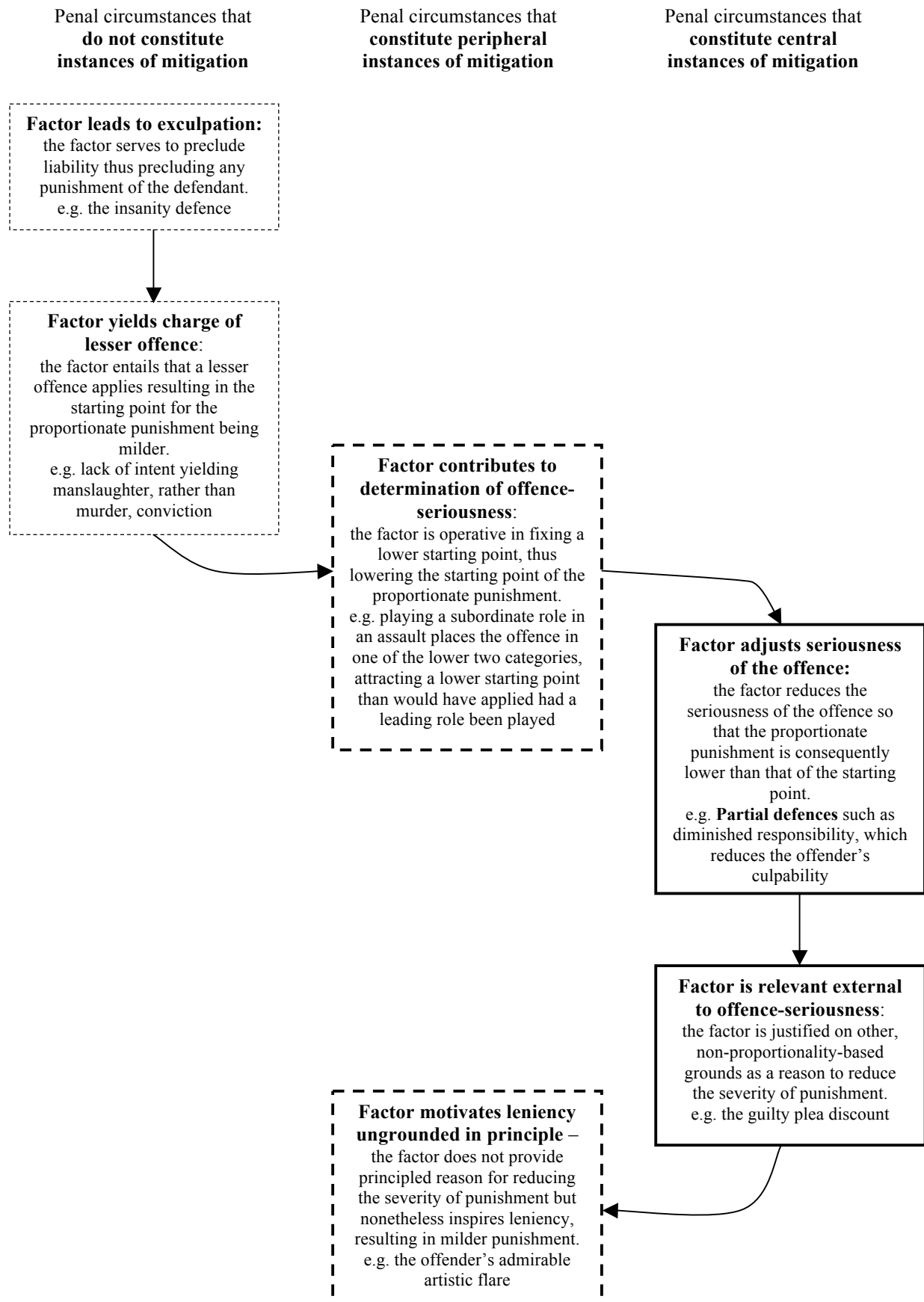
Although the language of mitigation still fits in such a scenario, this particular penal circumstance – where the reduction of the sentence is not related to principle – constitutes a peripheral (undesirable) instance of what could be labelled mitigation. Something about the offender has led to a reduction in the severity of the sentence, despite this being unjustified. ‘Artistic temperament’ could not be called a mitigating factor, as it is very unlikely that it would generally inspire leniency. Indeed, no factors

that operate without principled justification should really be labelled mitigating factors, even if they happen to (unjustifiably) motivate judicial leniency in a particular case.

Mitigation in summary

The figure below summarises all the possible ways discussed that a fact about the offender or offence might lead to a reduction in punishment. It shows which of these do not constitute mitigation, which constitute peripheral instances of mitigation and which constitute what I have argued to be the central instances of mitigation. The diagram expands on Wasik's (1982: 149) scale, introduced above, 'running downwards from excusing conditions, through partial excuses to mitigating factors'. Again, on the principle that sometimes we can see the boundaries of the concept we want to understand best when we also consider what it is not, I include the full length of Wasik's scale.

Figure One: Mechanisms through which a fact about the case may lead to a reduction in, or absence of, punishment



Sentencing factors with the potential to mitigate

As became evident in the discussion above, it is impossible to draw up an exhaustive list of mitigating factors: sometimes factors that might justifiably mitigate for one type of offence are actually defining features of another type of offence. Sometimes factors function to determine offence-seriousness and sometimes they operate to adjust the seriousness of the offence away from the determined starting point.

I therefore emphasise that it may actually be somewhat misleading to talk about mitigating factors *per se*, as there are no facts about offences or offenders that are inherently mitigating such that they will necessarily lead to reduced punishment. Even those that are easiest to distinguish (youth, for example) may not *always* mitigate once all the facts of the particular case have been weighed up by the sentencing judge. Instead there are sentencing factors which *may* be mitigating, depending on how they relate to a particular offence. I therefore suggest that it is more helpful to refer to ‘mitigating factors’ as ‘sentencing factors with the potential to mitigate’, and the theoretical literature indeed does generally speak of ‘potential mitigating factors’. Of course, in practice shorthand is needed but, when discussing mitigation in theory, language that is sensitive to the varying and contingent influence of most sentencing factors is more helpful and appropriate.

Mitigation as a particularly retributive concern

As the reader has probably gleaned from the preceding discussion, mitigation is a particularly retributive concern. Mitigation has been seen to first and foremost be a vehicle for maintaining proportionality, so that the offender does not get punished more severely than he deserves. Although mitigation also has a role to play outside

considerations of the seriousness of the offence (when justified on other grounds), it is archetypally relevant to desert. Ashworth (1994: 5) expresses this contention:

Questions of mitigation arise whenever the sentence has to be kept in proportion to the seriousness of the offence. Thus, mitigation as such is unlikely to be of great concern to rehabilitationists, except insofar as they observe a proportionality constraint. Nor is it likely to be a central issue for incapacitative theorists unless it is relevant as affecting predictions of the offender's future conduct. But in desert theory, and in Benthamite versions of deterrence theory, considerations of proportionality are particularly important, and so is mitigation. The very process of assessing the seriousness of offences involves taking account of aggravating and mitigating factors.

In emphasising the importance of mitigation to retributive theories of punishment, one is, as Ashworth argues above, also declaring it of minimal or no importance to purely consequentialist theories. Whereas retributive theories are backwards looking, focussing on the offence committed, consequential theories are forwards-looking to the effects punishment can have. Incapacitation, rehabilitation and deterrence are all means through which a safer society can be achieved through the prevention of crime. As Ashworth suggests, the sorts of factors that might justify mitigation on a desert account might be *relevant* to whether an offender is really in need of incapacitation or might be particularly suited to rehabilitative attempts or in need of less deterrence than average. However, to describe them as mitigating is stretching the concept of mitigation. Further, factors that may properly be mitigating from a retributive perspective may actually be aggravating from a consequentialist perspective.

Consider, for example, one such situation in which a factor might count as mitigating in a retributive context but aggravating if we want to try to apply the aggravation-mitigation discourse to consequentialist reasoning about punishment. It is possible that if a violent offender suffered severe physical abuse as a child, some retributivists would deem this as relevant to the assessment of his culpability, and judge him less culpable than he would have been absent the abuse in his personal

history. A consequentialist, however, would probably argue precisely the opposite: the inflammatory effect the childhood abuse has on the offender's behaviour makes him in greater need of incapacitation and specific deterrence. Thus, the factor might be seen by a consequentialist as justifying more severe punishment. Although one might well anticipate that different theories and perspectives will treat factors differently, this example highlights the scope of these differences: what can mitigate punishment within retributive theory might justify harsher punishment from a consequentialist perspective.

As can be seen, the factor that was relevant to the seriousness of the offence for the retributivist was simply a factor to be taken into account in a risk assessment by the consequentialist. The consequentialist's relative disinterest in how factors affect the seriousness of the offence makes couching such considerations in terms of mitigation and aggravation inappropriate. However, recall that for desert theorists 'the very process of assessing the seriousness of offences involves taking account of aggravating and mitigating factors' (ibid.).

Further, *Black's Law Dictionary* (1999: 236) defines a 'mitigating circumstance' as a 'fact or situation that does not justify or excuse a wrongful act or offence, but that reduces the degree of culpability and thus may reduce ... punishment'. This backwards-looking focus on culpability is essentially retributive. Interestingly, the definition does not include reducing the degree of harm caused. Although this can be a somewhat contentious issue, as sometimes lack of harm is a matter of luck rather than intention on the part of the offender, there are still occasions when attempts to minimise harm count decisively in favour of the offender. For example, an offender might return the goods that he stole, thus reducing harm in terms of loss of property. This being said, Ashworth (2010: 168) argues that many more of

the matters to be taken into account in mitigation do reflect the diminished culpability of the offender.

If mitigation is, as I have argued, essentially a retributive concern, the project of justifying remorse as a mitigating factor will involve engaging closely with retributive theories. However, so far, my description of retributive punishment has been at the most basic level. I therefore now turn to the task of more fully introducing its origin and main ideas.

Retributive punishment

At their most abstract level, retributive philosophies of punishment view punishment as either valuable *per se* or as a practice required by justice. Since justice is a good that we should promote, obligations to punish are generated where doing so serves to achieve justice. Retributive theories of punishment are backwards-looking, focusing on the offence committed. In addition to justice, desert is a key organising concept within retributivism. Desert requires that proportionality is a central consideration – if one deserves a certain quantum of punishment, then punishing more or less severely is (at least *prima facie*) unjust: the punishment must be in proportion to the seriousness of the offence (determined by the degree of harm caused or risked, and the culpability of the offender). The following summarises these claims: ‘the quest for justice is the underlying rationale of retributivism: justice is satisfied if the guilty are punished according to desert and in proportion to the gravity of the offence’ (Easton and Piper, 2008: 56).

Retributive theories come in different (and sometimes incompatible) forms. The classical view, proposed by Kant, has been modified and developed by contemporary theorists to give rise to accounts that have strong support today. I will

briefly outline the tenets of a Kant's classical retributive perspective and then go on to introduce three modern forms of desert theory.

Classical retributivism

Immanuel Kant is seen as one of the fathers of classical retributivism. He laid down the foundations for the idea that punishment should not be a means to an end but an end in itself. This is what defines all strong forms of retributivism. This notion is an outright rejection of any consequentialist notion of imposing punishment to obtain beneficial results:

Juridical punishment can 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... [2009: 210, emphasis in original]

Kant invokes what he refers to as the 'principle of equality' in his discussion of punishment. Allegiance to this principle requires that 'the pointer of the scale of justice is made to incline no more to the one side than the other' (ibid.: 211). When committing a wrongful act, the wrongdoer upsets the balance of the scale of justice. He inflicts suffering on another, and therefore renders himself deserving of suffering. So, in order to balance the scale of justice, the 'undeserved evil' the offender inflicted on the victim must be matched by the same amount inflicted on the offender: the principle of retaliation. Kant argues that only the principle of retaliation (*jus talionis*) 'can definitely assign both the quality and the quantity of a just penalty. All other standards are wavering and uncertain; and on account of other considerations involved in them, they contain no principle conformable to the sentence of pure and strict justice' (ibid.). Again, subscription to such a view entails that one renounce any

consequentialist considerations as contrary to justice and thus unjustifiable. For Kant, appropriate punishment is solely a matter of what is required to rebalance the justice scales.

Importantly, the particulars of Kant's view render an arbitrary choice of type of punishment impermissible. Kant states that the act that the person has performed 'is to be regarded as perpetrated on himself' (ibid.). This is proportionality in its strongest form. In this context, Kant employs the phrase '*Gleiches mit Gleichem*', which is usually translated as 'like for like, or measure for measure' emphasising both quantitative and qualitative matching in terms of the amount of suffering and type of punishment (Easton and Piper 2008: 58). Illustratively, the most straightforward application of this 'like for like' principle demands that murderers receive the death penalty, which Kant believes.

That Kant ultimately provides a successful retributive justification of punishment has been contested (see, for example, Byrd 1989; Murphy 1987; Scheid 1983; Tunick 1996). However, there are still those who argue that he does (for example, Johnson 2008) and it is indisputable that Kant's arguments concerning the distribution of punishment are clearly retributive.

Modern retributivism

Modern retributive theories, whilst retaining the basic retributive values (relying on concepts of justice, desert and proportionality), are often less 'thoroughly' retributive. By this, I mean that whilst retributive values are of primary import, 'quasi-retributive' or even consequentialist values also may play a role in the theory. Following Metz (2007) it may be helpful to draw a distinction between 'narrow' and 'broad' retributivism. Narrow retributivism can be defined as follows: 'The narrow sense of

‘retributivism’ is the desert or pay-back theory, the view that punishment of a certain degree is justified because offenders deserve to be harmed to a like degree for their wrongdoing’ (Metz 2007: 687). Kant’s retributivism would be narrow according to this definition. The broader notion of retributivism covers any backward-looking theory of legal punishment. In addition to the basic desert theory, the broad sense of retributivism includes two additional backward-looking accounts of the purpose of punishment:

fairness theory, the view that the government should impose a proportionate penalty because it thereby corrects the degree of unfair advantage that the criminal has taken, and censure theory (or intrinsic expressivism), the view that the government should punish in a way that fits the crime because it thereby expresses disapproval proportionate to its seriousness. [ibid.: 687-8]

I will briefly outline three retributive theories that have contemporary proponents. The first of these adopts the above-described ‘fairness theory’, the second and third adopt the ‘censure theory’, with the latter incorporating some additional, non-retributive elements. These three versions of retributivism will be considered in detail in the remainder of thesis, where I seek to develop a retributivism-compatible justification for mitigation on the grounds of the offender’s remorse.

Benefits and burdens retributivism

The benefits and burdens (or unfair advantage) justification for punishment had, for a time, been considered the most plausible theory of punishment by many philosophers. Although some have subsequently altered their thinking, the original proponents of the benefits and burdens theory include Morris (1968), Finnis (1972), and Murphy (1979). The theory is an entirely retributivist account, which seeks to justify punishment as the proper response to past crime. It seeks to characterise the institution of the criminal law as a mutually beneficial enterprise. A system of law brings

benefits to all its citizens (e.g. freedom from interference with body or property), by imposing burdens on them all (the burden of self-restraint involved in obeying the law). An offender refuses to accept the burden of obeying the law himself, whilst accepting the benefits that flow from the self-restraint of law-abiding others. He thus assumes an unfair advantage over those who obey the law. Benefits and burdens theorists justify punishment by arguing that it serves to remove the offender's unfair advantage by imposing an extra burden. Imposition of this burden is deserved by the offender due to his unsolicited taking advantage. Through his punishment, the antecedent balance of benefits and burdens that his crime disturbed is restored.

The penance perspective on retributivism

Antony Duff has developed a retributive theory of punishment at the centre of which is the concept of secular penance. His view builds on classical retributivism by understanding the criminal sanction in communicative terms. According to Duff, the sanction imposed on the offender does three things: first, it communicates disapprobation of the criminal conduct to the offender and to the wider community; second, it helps to bring the offender to a repentant understanding of his wrongdoing; third, it serves as a penance through which the offender can express his repentance, thus facilitating reconciliation with his community.

Central to Duff's theory is the tenet that offenders deserve the condemnation or censure that punishment communicates (Duff 2001: 27). Later he also suggests that offenders also deserve to 'suffer remorse' for what they have done (ibid.: 97). The criminal sanction, as secular penance, is the appropriate method for fulfilling these requirements.

Censure-focused retributivism

Andrew von Hirsch (1993) has developed a censure-focused theory of punishment. At the core of his arguments is the idea that criminal wrongdoing deserves censure. It deserves censure as 1) it takes the agent seriously as the author of his misbehaviour and 2) the state ought to take breaches of law seriously: this is owed to the citizens, and especially to the victims of crime. In order for the state to mean what it says when it declares that certain kinds of conduct are seriously wrong is to be ready to censure such conduct. Censure is retributive in nature because it is essentially backward-looking, focusing on the past wrongdoing.

The second core idea is that punishment (the criminal sanction) is not only a suitable medium through which this censure can be communicated, but also has the benefit of providing an additional prudential disincentive to offending, which is necessary as many of us find it hard to resist wrongdoing for normative reasons alone. His ‘two-pronged’ justification for punishment thus involves a consequentialist element of deterrence. However, von Hirsch makes it clear that the *primary* justification is the need to communicate censure to offenders, thereby respecting their rational moral agency and avoiding an institution of ‘beast control’ which would be engendered by purely deterrent systems.³⁷

The possibilities for remorse

Now that we have a clear understanding of mitigation and an introduction to prominent desert theories, we can consider what sorts of arguments could be made to the effect that remorse is justified as a mitigating factor. If we return to the summary

³⁷ Von Hirsch suggests that an institution utilising only the preventative function of punishment would appear to treat the offender ‘like a tiger’ (1993: 12). He therefore makes clear how, on his model, the preventative function is related to censure in a way that retains respect for the offender as a moral agent capable of understating the message of censure conveyed through the sanction (ibid: 13).

of the mechanisms through which a factor can mitigate provided by *Figure One* above, we can work out the possibilities for remorse. Causing exculpation and yielding a different (lesser) offence were seen to not be instances of mitigation. It is implausible anyway that remorse could do either of these things. It does not justify or excuse criminal behaviour and it is not integral to the definition of any offence. Nor does remorse have any role in setting the starting point for a particular offence – it is not present in the ‘standard cases’ of any offence, against which the starting points are set. Justification as an instance of this peripheral mechanism of mitigation is thus precluded.

The remaining possibilities, however, do provide us with mechanisms through which remorse could operate to reduce an offender’s sentence. The argument that remorse somehow reduces the offender’s culpability or the harm caused could be proposed. If it were to do so, remorse would serve to adjust downwards the seriousness of the offence and hence the accompanying proportionate punishment. The second possibility is that remorse is justified as a mitigating factor externally to considerations of offence seriousness. This could be on quasi-retributive or compassionate grounds, whereby the offender’s remorse is relevant to the censuring function of punishment or justifies a reduction in sentence grounded in reasons borne of a virtuous concern for the offender’s wellbeing. The final possibility is that remorse motivates unprincipled leniency in judges on occasion. It would then fail to be *justified* as a mitigating factor but might nonetheless have a mitigating effect from time to time.

If remorse can only operate via this last mechanism then the most that could be offered would be a description (and perhaps explanation) of its tendency to mitigate, rather than a justification. This would be a disappointing conclusion for

those who believe remorse should be taken into account at sentencing. The preference would be to explain why remorse operates via one of the two central mechanisms of mitigation. Both of these will be explored at length in the remainder of the thesis.

There is one further way in which remorse may serve to reduce an offender's sentence within retributive theories. If remorse were actually to constitute punishment, then the amount that the state should impose would be reduced, as the offender is already being punished to a certain degree.³⁸ If the punishment was not reduced, the offender would be over-punished, breaching proportionality requirements central to desert theories. Remorse would then be mitigating in the sense that it provides reasons why the punishment the state needs to deliver should be less than it would absent the offender's remorse. This is different from the logic involved in the central mechanisms of mitigation where the factors reduce the punishment deserved or justified. On this account, the amount of punishment remains the same, but the offender has already been subjected to some of it. This is more analogous to situations where remand time is taken off a sentence of custody: the severity of the offender's punishment has not changed; he has simply already served some of it.

With close reference to the three desert theories presented above, Chapter Six explores this possibility – that remorse constitutes or otherwise substitutes for some of the deserved punishment. Chapter Seven then explores whether remorse can reduce the seriousness of the offence, either through affecting culpability or harm. The final two chapters explore whether remorse is justified as a mitigating factor externally to offence-seriousness. This discussion serves to facilitate the development of an account that argues that it is so justified, based firmly on quasi-retributive grounds.

³⁸ Suggested but ultimately rejected by Tasioulas (2006: 309): It could be suggested that remorse 'bear[s] directly on deserved punishment, because a person who has already been punished for their wrong does not deserve to be punished for it again, and we can interpret the process of repentance he undergoes as a form of self-inflicted punishment.

Chapter Six

Does remorse constitute some of the deserved punishment?

Experiencing remorse is inherently painful. Indeed, Adam Smith goes as far as saying that ‘of all the sentiments which can enter the human breast [remorse is] the most dreadful’ (2002: 99). His description of remorse raises the question of whether living with one’s remorse might be the worst punishment of all. Whether a retributive theory of sentencing could – or should – allow that an offender’s remorse constitutes some of the punishment that he legally deserves depends on the theory’s definition of ‘punishment’ and what it takes the aims of this punishment to be. I am going to assess this possibility in relation to versions of the ‘unfair advantage’ theory, Duff’s ‘penance perspective’ and von Hirsch’s desert theory. It will emerge that remorse alone cannot constitute any of the deserved punishment in these accounts, but for different reasons in each instance.

Natural punishment

In considering whether remorse might constitute or viably substitute for some of the deserved punishment we should briefly return to the concept of natural punishment, referred to in Chapter Four. Natural punishment differs from what we might call ‘punishment proper’, which necessarily involves an agent or agents as those who inflict the punishment (see Teichman 1973). Natural punishment should also be distinguished from simple bad luck. Whereas a burglar who breaks his leg slipping from the roof of the house he is burgling might be said to have received natural punishment, the burglar who later slips and breaks his leg in a supermarket does not seem to have received natural punishment – he is just unlucky in his injury. These

examples help illustrate what is central to the concept of natural punishment: the adverse event must be a direct consequence of the illegal, immoral or imprudent behaviour. It is this intimate connection of the ‘punishment’ to the malpractice that gives rise to the related literary notion of ‘poetic justice’. Poetic justice is achieved when, often ironically, vice is ultimately punished by a twist of fate intimately related to the character's own conduct.

Whilst these concepts may seem principally metaphorical, the concept of natural punishment has been observed as operating in some jurisdictions’ sentencing practices. Notably, the Swedish law regards as mitigating factors the fact that the offender suffered injury as a result of the crime (Swedish factor 1); and the fact that the offender’s employment prospects have been damaged as a result of conviction (factor 5). Some theorists have questioned the legitimacy of the exercise of judicial compassion in cases where the offender has brought foreseeable injuries or difficulties upon himself as consequences of his committing his offence (e.g. von Hirsch and Ashworth 2005: 177). However, it should be remembered that what is at issue in this chapter (in relation to adverse consequences of an offence) is not the justification of compassion, but whether the full (unmitigated) punishment is still deserved if something adequately substitutes for some of it.³⁹

It could be thought that remorse, which can only be a direct result of wrongdoing, might constitute natural punishment and therefore justify mitigation. What will be of primary issue in this chapter, however, is not whether the label of ‘natural punishment’ is conceptually applicable to remorse, but whether remorse can adequately satisfy the definition and/or objectives of punishment theorists set out. It is the possibility that it might that begets the corresponding possibility that a reduction

³⁹ Walker (1999: 132) draws a similar justificatory distinction when he asks whether ‘retributivists are divided about the justice of regarding “natural punishment” as a mitigating factor, perhaps because it is a notion which seems to be based ... on humanity rather than desert’.

in the severity of punishment to which the remorseful offender is sentenced might therefore be justified on these grounds.

Punishment proper

Hart's conditions for punishment have become accepted by many penal philosophers as the definitive conception of punishment. Hart defines the standard or central case of punishment in terms of five elements:

- (i) It must involve pain or other consequences normally considered unpleasant.
- (ii) It must be for an offence against legal rules.
- (iii) It must be of an actual or supposed offender for his offence.
- (iv) It must be intentionally administered by human beings other than the offender.
- (v) It must be imposed and administered by an authority constituted by a legal system against which the offence is committed. [1959: 4]

However, precise working definitions used within various retributive theories are often more nuanced and also coloured by the objectives that theorists think punishment should promote. Punishment is not merely hard treatment, and close attention needs to be paid to any additional criteria theorists identify.

It should be noted that there is a difference between the possibility of remorse *constituting* punishment and of remorse being a *suitable substitute for* punishment. For the first possibility, remorse would have to satisfy the definition of punishment. For the latter it would have to be able to achieve all the objectives a theorist sets out for punishment. Either possibility could result in justifying mitigation of state punishment. If remorse were to satisfy the definition of punishment, then some mitigation of state punishment would be necessary, otherwise the offender would be over punished. If remorse were to achieve all the aims set out for punishment, this would result in there being an argument for mitigation of state punishment. The

argument would fail if theorists require, in addition, that the aims must be achieved through state punishment and that, although remorse might also achieve them, this is irrelevant to the level of state punishment required. Thus, throughout, we must be sensitive to these two ways in which remorse might count as some of the deserved punishment.

The subjective vs. objective metrics of punishment debate

A debate pertinent to, but not identical with, the one at stake in this chapter questions whether the offender's individual subjective experience of his punishment alters how severe it is and, if so, how retributive theories should deal with this. That is: should proportionality theorists adopt subjective or objective metrics of punishment? Kolber (2009) presents us with a scenario in which two offenders – Sensitive and Insensitive – having committed offences of equal gravity, are sentenced to exactly the same length of prison sentence which they serve in identical physical conditions (their cells are the same size; the guards treat them in the same manner; etc). Sensitive, however, experiences his sentence as more aversive than Insensitive, due to his fragile disposition. Kolber argues that this difference in experience means that their punishments are not identical in severity and that therefore at least one must be getting punished disproportionately. He says, 'to be retributively punished, the person punished must find the punishment aversive and the severity of the punishment is at least partly a function of how aversive he finds it' (2009: 215).

For this argument to help justify remorse as a mitigating factor the following two things would have to be true: first, remorse would have to consistently (perhaps universally) correlate with (perhaps be the cause of) offenders experiencing punishment as more aversive than the average offender. Second, Kolber's (2009)

position that subjective experience matters would have to be defended – it would have to be true that subjective experience affects, in a way relevant to retributivism, the severity of the punishment the offender endures.

However, both premise one and two can be contested. First, it is an empirical question, to which the answer is not obvious, whether remorse makes people's experiences of punishment worse. In fact, it was suggested in Chapter Four that punishment might in some cases actually be more bearable for the remorseful, as they feel it is deserved and possibly view it as a way to atone for their wrongdoing. This might not make their remorse any less painful – they are still haunted by what they have done – but may allow them to adopt an attitude of acceptance with respect to their punishment that makes the experience somewhat less gruelling. Certainly, I think there must be a multitude of different ways in which remorseful offenders experience their punishments, and that a common subjective experience profile for the remorseful offender – suggesting that they consistently find punishment more aversive – would be an unsupportable fiction.

Secondly, the argument that subjective metrics of punishment are essential to ensure proportionality is highly contested. Markel *et al.* (2011) take issue with the view that the individualisation of punishment – based on how the offender subjectively experiences it – is necessary to getting retributive justice right. They argue that 'punishment on this view is little more than a complicated pain-delivery device' (2011: 616). Coming from a communicative retributive perspective, their main point is that human understanding is informed by more than just one's experiences. The deprivation of liberty is objectively bad. The fact that we discover it is objectively bad by experiencing it as subjectively bad does not make the subjective experience *constitutive* of what is bad. Social meaning plays a role which is not

reducible to subjective experience: £500 in taxes and a £500 fine for committing an offence may ‘generate the same hedonic dip’, but the meaning is different (2011: 615). Thus, they conclude that ‘as long as it is reasonable for people in our society to think that a politically sanctioned constraint of liberty communicates condemnation, and that stiffer sanctions signal yet great condemnation, adaptation among typical offenders need not play a central or even prominent role when setting sentencing policy’ (ibid.). In contrast to theorists such as Kolber, they believe that it is the objective ‘badness’ of punishment that is most important in determining the quantum that should be inflicted.

We should not be too worried about resolving the subjective vs. objective metric of punishment debate for the purposes of this chapter. Crucial to note is that the above argument (as a way to justify an influence of remorse on punishment severity) is different from that which would be required to support the contention explored in this chapter – that remorse might adequately substitute for some of the deserved punishment. In an ‘experience matters’ argument, the main idea would be that the remorseful offender experiences his punishment as unusually harsh and therefore should be punished less than the average offender to maintain proportionality. This is not the same as the idea that remorse embodies or delivers something that approximates punishment (and, therefore delivers additional punishment) *in itself*. This latter possibility is what this chapter seeks to test.

The ‘experience matters’ argument, even if successful, would not be able to posit remorse *per se* as a mitigating factor. This is because it is experience more diffusely that is relevant, within which remorse might play a complex role. It would be offenders’ experiences of *punishment* that could at most be taken into account, not a more general measure of overall wellbeing. Although they might have an additive or

interactive effect on each other, I would argue that we can distinguish between an offender's experience of his punishment and his experience of his remorse. The question may then become whether remorse is a separate additional punishment (as opposed to a modifier of the experience of the basic punishment). It is essentially this question that this chapter asks.

So, although the themes discussed here will re-emerge in this chapter, it does not matter if we do not reach a conclusion on the subjective vs. objective metrics debate at this juncture, as it is separate to the question of whether remorse constitutes punishment. If it were to subsequently turn out that we accept the experience matters argument, this would then be a *further* consideration, not an aspect of that which is currently our focus. Whilst the conclusions reached in this chapter will suggest whether remorse should mitigate on the grounds that the remorseful offender has already been punished, the experience matters argument cannot consider remorse as mitigating *per se*: sometimes the remorseful person will be more sensitive but perhaps sometimes less. Additionally, the sentencer would have to consider all the other elements of the offender's overall disposition that might influence his experience of his punishment. Considerations of remorse would dissolve in to broader considerations of general subjective experience.

I now return to the central question of this chapter by exploring whether remorse could constitute or substitute for punishment in the first of our retributive theories: the benefits and burdens theory.

Theory one: benefits and burdens

The particulars of the 'benefits and burdens' (or 'unfair advantage') retributive theory of punishment depend on which version is adopted. However, they all share the same

basic logic: the wrongdoer's punishment is necessary to address an imbalance created by the offence. Individuals living in a just society enjoy the freedom to pursue their individual projects without unjustified interference from others. In exchange for this benefit, however, such individuals are required to assume the burden of allowing others to enjoy the same freedom: they cannot interfere with another individual's body or property unjustifiably. The offender who fails to exercise such self-restraint gains an unfair advantage over others exercising that restraint. The burden of punishment then negates this unfair advantage. The first example of this way of thinking about the justification of punishment can be found in the writings of Herbert Morris. He writes:

It is just to punish those who have violated the rules and caused the unfair distribution of benefits and burdens. A person who violates the rules has something others have – the benefits of the system – but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage. Matters are not even until this advantage is in some way erased. Another way of putting it is that he owes something to others, for he has something that does not rightfully belong to him. Justice – that is punishing such individuals – restores the equilibrium of benefits and burdens by taking from the individual what he owes, that is, exacting the debt'. [1968: 478]

So, on this theory, the aim of punishment is to redress the imbalance of benefits and burdens. If an offender's remorse is to constitute some of the deserved punishment on this account then it would have to coherently be understood as a burden or as otherwise removing some of the benefit attained by the offender. So, it could be argued that the offender's remorse is a relevant burden and therefore that a less severe state burden need be imposed in order to restore the balance. Or, the remorseful offender might be seen to have lost some of the benefit gained from the offence and therefore a less severe burden is needed to balance this reduced benefit – both possibilities will depend on what the relevant benefits and burdens are.

Therefore, in order to see if remorse can function in this way, we need to look more closely at what is meant by burden and benefit, and by unfair advantage.

The notion of balance requires that ‘items’ are most naturally of similar types, so that they can weigh against each other. Incommensurability increasingly becomes a problem as the items to be balanced exist in different domains (e.g. material; psychological; interference). Thus, we should start by considering the nature of the benefit the offender gains from his offence in order to determine what may reasonably count as a corresponding burden – one which is effective in restoring balance. Burgh (1982: 203) sketches the possibilities for what might constitute the unfair benefit. He sees four candidates:

(1) the ill-gotten gain, e.g., money in the case of robbery; (2) not bearing the burden of self-restraint, hence having a bit more freedom than others; (3) the satisfaction from committing the crime; and (4) the sphere of noninterference which results from general obedience to the particular law violated; e.g., each person benefits from property laws insofar as he is free from interference with his property.

It is obvious that, although ill-gotten gains are a benefit of sorts, they cannot be a benefit of relevance to the benefit-burden balance that unfair advantage theorists envisage. Only some crimes are acquisitive and an offence such as assault confers no material benefits on the offender – yet we obviously see assault as a serious offence. In addition, the ill-gotten gains could not count as the relevant benefit even in the case of acquisitive crimes, since we would be dissatisfied if the balance was restored simply by taking the money back from the offender (in the case of a robbery). The remaining three possibilities, however, are much more plausible. Let us examine each one in turn.

‘Not bearing the burden of self-restraint, hence having a bit more freedom than others’

The second potentially relevant benefit was, according to Burgh, the offender’s ‘not bearing the burden of self-restraint, hence having a bit more freedom than others’. There are two slightly different interpretations of this benefit, both of which involve ‘benefits’ that are perhaps better understood as the avoidance of taking on a burden. ‘Not bearing the burden of self-restraint’ can be understood in a purely psychological sense. Restraining oneself – practicing self-control – can take much effort. If a wallet has been left unattended with a large sum of money in it, it might take an individual an enormous amount of will power not to reach out and take the visible notes. The individual might have to keep repeating to himself that he should not take the money and this mental effort will be accompanied by the frustration of denying himself what he wants. The offender, on the other hand, expends none of this mental energy and feels no such frustration. He lets himself do as he pleases. On this reading, ‘having a bit more freedom than others’ might be better stated as *experiencing* a bit more freedom than others: the offender feels free, where as the resister feels inhibited.

However, there is a more objective interpretation of this possible benefit. On an alternative reading, rather than the offender *experiencing* a bit more freedom than others, the point is simply that he *exercises* a bit more freedom than others. Again, I deviate from the word ‘having’ because the non-offender does not *have* any less freedom than the offender. It is not that he cannot reach out and take the money from the wallet rather, he refrains from so doing. The offender is not more free, but he does engage in behaviour from which the majority would refrain. On this reading, the benefit is objective, resulting from engagement in a larger set of

desirable (to the offender) behaviours. The benefit is not, then, drawn simply from the subjective pleasure (or avoidance of displeasure) of not practicing self-restraint.

The benefit attained by the wrongdoer has been envisaged in a similar way by Finnis (1972). According to him,

What the criminal gains in the act of committing the crime (whatever the size and nature of the loot, if any, and indeed quite apart from the success or failure of his overall purpose) is the advantage of indulging a (wrongful) self-preference, of permitting himself an excessive freedom in choosing – this advantage...being something that his law abiding fellow-citizens have denied themselves insofar as they have chosen to conform their will (habits and choices) to the law even when they would ‘prefer’ not to [1972: 132].

‘The satisfaction from committing the crime’

The third possible benefit is the ‘satisfaction from committing the crime’. Despite being an explicitly hedonic benefit, it differs from the subjective reading of the previous benefit. Not bearing the burden of self-restraint amounts to avoiding something aversive. Gaining satisfaction from committing the crime involves an additional pleasure. We can draw a distinction between the satisfaction *of* committing the crime – the rush during the actual act, which will be necessarily temporary – and satisfaction *from* committing the crime – perhaps drawn from a more positive self-image, a sense of achievement, and the memories of the successfully executed offence, all of which may be more sustained. If the crime was an acquisitive one, one’s satisfaction might partially be derived from the additional resources. If violent, one’s aggressive impulses will have been satisfied and one might feel a sense of fulfilment.

Hart alluded to an unfair advantage-style conception of the offender being required to pay ‘the price of some satisfaction obtained by breach of law’ (1968: 47). Hart explains ‘it is pointed out that in some cases the successful completion of

a crime may be a source of gratification, and in the case of theft of actual gain, and in such cases to punish the successful criminal more severely may be one way of depriving him of these illicit satisfactions which the unsuccessful have never had' (ibid.: 131). Punishment here seems to serve the purpose of removing satisfaction.

Satisfaction as a benefit to allay also gains support from Murphy's discussions of 'retributive hatred' and the principles which it reveals. Murphy describes these feelings as 'feelings that another person's current level of well-being is undeserved or ill-gotten...and that a reduction in that well-being will simply represent his getting his just deserts' (Murphy 1988a: 89). He regards retributive hatred as 'in principle the natural, fitting, and proper response to certain instances of wrongdoing' (ibid.: 108). He argues that retributive hatred is most justified when it is believed that a wrongdoer, who has hurt someone badly through immoral treatment, is 'totally unrepentant of the wrongdoing and is in fact living a life of freedom and contentment' (ibid.: 91). Such intuitions to deprive the wrongdoer of this enhanced well-being and life of freedom and contentment provide prima facie plausibility to a balance-restoring justification of punishment where levels of satisfaction constitute some of the unfair advantage.

'The sphere of noninterference'

The fourth plausibly relevant benefit was 'the sphere of noninterference which results from general obedience to the particular law violated; e.g. each person benefits from property laws insofar as he is free from interference with his property'.

This benefit is what Morris means by the offender having 'the benefits of the system'. The offender defaults on the general requirement not to interfere with each other and thus he benefits from others' adherence. It could be argued that

everyone benefits from the system insofar as it is generally adhered to, but the offender benefits in a different way: he is a free rider, as he *only* benefits, whilst failing to uphold his part of the 'deal'.

Remorse as a relevant burden (or loss of benefit)

Having explored the sorts of things that could count as the relevant benefits that the offender gains, we can now see if remorse goes anyway to restoring the balance; to removing the unfair advantage.

Remorse as dissatisfaction

The most obvious starting point would be that remorse, as a deeply aversive psychological state, negates any satisfaction drawn from the offence. Any sense of achievement or enjoyment of the products or effects of the offence will be balanced by the emotional pain of remorse. It would actually be more correct to say that remorse indicates a loss of the benefit of satisfaction, rather than that it acts as a counterbalance. Genuine remorse would result in the offender retaining no satisfaction at all from the offence, in his memory, self-image or sense of fulfilment. These satisfactions would instead be replaced by horror at the memory of the offence and a greatly depleted self-image. The offender might also experience horror at his own satisfaction or pleasure that he originally gained from the offence. This might be particularly true for sexual offenders.

However, even if remorse does quash any satisfaction from the offence, is a psychological burden really what the unfair advantage theorists envisage as serving to satisfactorily restore balance? There is no denying that many of the sanctions imposed upon offenders are psychologically burdensome. Moreover, the idea that we want to

stop an offender feeling good about his or her offence is appealing – the smugness or gloating of an offender would repulse us. But offsetting satisfaction does not seem to be at the heart of what balance restoration means to unfair advantage theorists. If we consider what the implications for a benefits and burdens theory based on quashing ill-gotten satisfaction, we can begin to understand why such a model would not appeal.

The overriding point is that any model of punishment concerned only with the subjective would not only be impossible to implement, but would be inadequate even if possible.⁴⁰ Such a model would not be practicable as the complex ‘satisfaction’ would be made up of an indeterminate set of indices, which would be very difficult and intrusive to measure, even if identifiable. More importantly, setting punishment against satisfaction would not be desirable. Levels of satisfaction would vary drastically from person to person. The satisfaction one offender might get from, for example, an assault could be much greater than that of another offender who committed exactly the same offence.

For example, suppose offender A socialises in a subculture which values and praises aggressive, dominant behaviour. As a result of his offence, offender A grows in confidence, is proud of his behaviour, and has a much higher opinion of himself. The offence leaves him greatly satisfied. Compare this with offender B who gains some satisfaction from hurting someone he did not like but, due to being surrounded by disapproval, gains none of the ‘status-satisfaction’ that offender A does.

Moreover, the satisfaction any offender acquires will likely be influenced day to day by myriad factors – from the offender’s mood, to what people say to him about

⁴⁰ We should note that this would be much more extreme position than the ‘subjective experience matters’ argument discussed above. Whilst Kolber (2009) wanted to argue for the relevance of the offender’s experience of his punishment to the estimation of its severity, he conceived of offence seriousness in a conventional way. What the current idea involves, however, is that the seriousness of the offence is *also* a function of the offender’s subjective experience – how much satisfaction he feels as a result of it. Offences would be as serious as they are satisfying.

his offence, to what he learns others have gotten away with in comparison. It is likely that all ‘types’ of satisfaction will generally decrease over time. Varying levels of an offender’s satisfaction would potentially result in a different level of severity of punishment being deserved every day.

The same problem would appear when trying to work out how much dissatisfaction the sanction should mete out.⁴¹ Although all retributive theories of punishment should be sensitive to the reality that a given severity of punishment will ‘seem’ more or less severe to different offenders, a benefits and burdens theory that seeks to balance satisfaction with dissatisfaction would have to be completely engaged with these discrepancies. It would have to work out how this particular person would be dissatisfied to a degree equivalent to the satisfaction he acquired from his offence. Further problems would arise when we consider that a few offenders might actually gain satisfaction from some aspects of serving their sentences. Being in prison can be seen as a status symbol amongst a small minority. For others, despite forfeiting their liberty, quality of life might actually be marginally preferable in prison and so not entirely dissatisfying on some level. Of course, this would only be true for a very few offenders, but is something an entirely subjective benefits-burdens model would have to contend with.

The unfair advantage cannot just be constituted by how happy, content or fulfilled one feels following one’s offence. Indeed, Finnis says that punishment serves to ‘negate, cancel out, the advantage the offender gained in the crime – the advantage *not necessarily of loot or psychological satisfaction*, but of having pursued one’s own purposes even when the law required that one refrain from doing so’ (1999: 102,

⁴¹ Kolber (2005) would argue, however, that although criminal justice institutions might refuse to take subjective experience into account on the grounds that it is too difficult to make assess, they nonetheless need to accept that the sentences they impose will often be disproportionate, given differences in subjective experience.

emphasis added). Through ‘not bearing the burden of self-restraint’ one pursues one’s own, forbidden purposes. We must now consider how remorse might constitute a relevant burden if we focus on ‘bearing the burden of self-restraint’ as the relevant burden.

Remorse as re-imposing self-restraint

It could be argued that the remorseful offender re-imposes restraint on himself. If an offender were to be genuinely remorseful, then he would (at least strongly intend to) not commit the same offence again, if he were to find himself in a similar situation. He would now restrain himself. However, a new enthusiasm for resisting a particular temptation does not seem to counterbalance the benefit of previously indulging one’s ‘(wrongful) self-preference’. It does not seem to constitute a burden over and above that which the majority of people take on in obeying the law when they might be tempted not to. However, if we consider what Finnis says about the important aspects of punishment, we gain a clearer understanding of what could constitute the relevant burden, and whether remorse can go any way to being an appropriate substitute.

Finnis says that punishment is to be defined:

not, formally speaking, in terms of the infliction of pain (nor as incarceration), but rather in terms of the subjection of will (normally, but not necessarily, effected through the denial of benefits and advantages of social living: compulsory employment on some useful work which the criminal would not of himself have chosen to do would satisfy the definition). [1992: 133]

So, according to Finnis, to offset the offender’s indulgence of his wrongful self-preference, his will needs to be subjected. He states explicitly that this ‘punishment’ does not necessarily have to be incarceration, but that in most cases it is achieved through denying the offender the ‘benefits and advantages of social living’. If this is what is required to offset the offender’s unfair advantage, might remorse go some way

to effecting such denial? Remorse certainly should have a profound effect on the offender's capacity for enjoying even the most simple of social pleasures. With the reality of what one has done firmly in the forefront of one's mind, one likely feels incapable of engaging in all the activities one once would have, and would not gain much pleasure from them even if one did engage. Duff portrays remorse as categorically rendering the offender bereft of the benefits and advantages of social living. Of remorse, he argues that:

it will prevent me from enjoying what I would otherwise enjoy. If I truly recognise and repent what I have done, I will not be able to enjoy, for instance, my usual social pleasures. I do not decide to deprive myself of them as a punishment; but I cannot enjoy them, any more than I could enjoy them whilst mourning the death of a friend. I cannot enjoy them because my attention and my concern are dominated by the wrong I have done. [1988: 164]

This goes beyond the function of remorse to inspire renewed commitment to self-restraint, which I argued did not constitute a burden. The burden here is that the offender is deprived enjoyment of life. Importantly, as Duff notes, the remorseful offender does not impose this deprivation on himself; rather, he is incapable of finding pleasure in the things he used to enjoy. As with state punishment, this deprivation is not a choice – something Finnis seems to require in order to achieve 'subjection of the will'. Perhaps, then, remorse does impose a relevant burden.

However, it is not clear that what Finnis requires for subjection of the will is simply that the offender is rendered unable to enjoy life. A closer reading reveals that the issue is more about the offender's freedom. The benefits and advantages of social living are not so much to do with hedonic pleasure, but rather are the freedoms that we have as citizens in a free society. Thus, Finnis explains that the 'disadvantage' the offender must endure is of 'having his wayward will restricted in its freedom by being subjected to the representative "will of society" (the "will" which he disregarded in

disregarding the law' (1992: 133). The offender's will is not subjected by remorse, but is merely modified or stifled by its experience. Something external is required for subjection of the will. The offender, having permitted himself an 'excessive freedom in choosing' through acting according to his tastes (ibid.: 132) is to have his freedom curtailed – his choices restricted – by society to restore balance.

Is there any way, then, that remorse can be relevant within this scheme of benefits and burdens? Within his writing generally on the relevance of 'private burdens' to fair-play theories, Bayern (2009) argues that it can. Bayern considers the offender who 'no longer believes that the excess freedom associated with a prior crime is worth having' (2009: 40) – something integral to a remorseful response. Bayern asks what the point is – within fair-play theories – of imposing punishment 'if we can confidently determine that the offender is no longer the sort of person who considers the excess freedom from her crime a benefit (or, to put this differently, considers it a *freedom* in the first place' (ibid.).

The argument seems to be that what needs to be offset is the wrongdoer's positive experience of exercising freedoms that are usually renounced. The remorseful offender no longer enjoys a sense of freedom in connection with his offence, as his remorse reframes the past events as rebellious wrongdoing. Since the offender no longer values or associates his offence as appropriation of freedom, there is no benefit, and hence nothing to offset.

I find the above argument unconvincing. Its primary fault is found in the conception of 'freedom' adopted. We can expand on the distinction I drew between *experiencing* a bit more freedom than others and *exercising* a bit more freedom than others. The former is subjective, whilst the latter objective. Bayern nods towards this distinction when he wonders if his argument would be convincing to someone

adopting a conception of freedom that was particularly 'abstract'. However he does not develop this line of enquiry. Bayern's argument requires a subjective conception of freedom, which must boil down to the offender's maintenance of positive attitudes regarding the offence, particularly involving feelings of liberation and so on. This, however, is then no different to the benefit of drawing satisfaction from the crime, and fails to be of central relevance for the same reasons.

'Freedom', then, must be understood objectively: the offender appropriated excess freedom through engaging in prohibited activity, regardless of whether he felt liberated or satisfied by it. Punishment, within Finnis' theory must come from the state as only this can be representative of the 'will of society'. The offender's punishment must be a societal matter because the benefits and burdens at issue arise from a system constructed by, and imposed upon, society. Remorse is not a burden which represents the will of society and therefore cannot be a burden relevant to unfair advantage theories.

Theory two: Duff and the 'penance perspective'

I will now consider whether remorse might constitute, or be a substitute for, deserved punishment in relation to Duff's penance perspective. For Duff, the imposition of penances is the constitutive aim of the criminal sanction: it is on these grounds that he justifies its existence. Since remorse is intrinsically tied to what Duff envisages the ideal penance being and achieving, it seems likely that an offender's remorse will have some relevance to the punishment he receives. I am going to argue that the main tenets of Duff's view result in the possibility of a remorsefully-imposed penance constituting some of the deserved punishment. However, I shall also highlight why

Duff would disagree with me, challenging his likely opposition and discussing vacillation in his work.

State punishment and self-punishment

In a 1988 article, we find Duff arguing that an account of criminal punishment should be founded on the idea of self-punishment. Crucial to this argument is his idea of prototypical penance as punishment that is undertaken voluntarily. Duff begins his argument with the strong claim that ‘criminal punishment...should ideally aim to become self-punishment’ (1988: 159). Given the primary importance Duff places on self-punishment, and its purported supremacy over state punishment, might there be room here to portray remorse as constituting part of this self-punishment? It will be argued that this would be possible within Duff’s account but that there would be two necessary conditions which may be difficult (although not logically impossible) to meet. It will emerge that for remorse to constitute some of the deserved punishment in the case of a particular individual, he must 1) be capable of experiencing/achieving complete repentance independently and 2) adequately express his remorse through a self-imposed penance. Duff would see both these conditions as realistically unattainable by any existing individual.

Duff’s argument seems to be that state punishment is a less-than-ideal substitute for self-punishment. Despite state punishment not being the ideal, Duff claims that it is justified nonetheless. His argument is that in the majority of cases, where offenders do not impose punishment on themselves, there is the hope that the offender will eventually come to want to have punished himself. This would render it more like self-punishment (since the offender retrospectively accepts the state punishment as what he would have wished for himself if he had felt/thought then as

he does now). State punishment can meet the relevant aims since it, according to Duff, both serves to facilitate repentance and to express this repentance to the community. It thus functions like a penance and ‘in accepting it as a penance, [the offender] accepts it as a punishment which he *should* impose on himself’ (ibid.: 163).

However, Duff’s position is that the state-penitential model is derivative of more authentic self-imposed penance. If state punishment is to be justifiably derived from the self-punishment model, we must start with identifying the value of self-punishment. Duff argues that ‘what is essential to [state] punishment is an outward or manifest suffering which gives symbolic expression to the pain of remorse. This meaning is *most easily seen* in the case of self-punishment; and to justify criminal punishment we must then show how imposed punishments can come to have this meaning...’ (ibid.: 160, italics added). Here we see one of the key features of value of self-punishment – that it expresses the offender’s remorse. The vandal who desists and proceeds voluntarily to clear a whole neighbourhood of gratuitous graffiti would likely be seen to understand and to have been moved by the damage he has caused. This prompted him to express his contempt for his behaviour through giving up all his free time in penitential reparative action.

Eliciting repentance

The purpose of state punishment, for Duff, is not only as a vehicle for expressing repentance, either concurrently or retrospectively. In addition, ‘punishment aims to persuade the criminal to accept the justified condemnation of her past crime, and the understanding of the nature and implications of that crime, which it expresses’ (ibid.: 162). So, state punishment is also a means of eliciting repentance, through focussing the offender’s mind on his offence and the harm he has caused. The hard treatment

‘aims to make the offender hear the message it expresses – to force his attention onto his crime’ (ibid.: 162).

The relationship between hard treatment and self-punishment seems less obvious for the question of focussing the mind. It would seem that the offender’s mind must already be focussed repentantly on his offence if he is to impose self-punishment, which can then function to express his already-repentant understanding. However, Duff suggests that self-punishment is required in order for the offender to reach a *fully* repentant understanding of what he has done. Thus, Duff describes penances (self-imposed or otherwise) as ‘various kinds of hair shirt which, being essentially uncomfortable, provide an ever-present reminder of my wrongdoing’ (ibid.: 165). This is what makes the ‘separate’ or ‘extra’ pain associated with penance necessary. It needs to involve some ‘imposition, or deprivation, or burden, which is painful quite apart from its penitential meaning’ (ibid.: 163).

Given Duff’s project – to show how the values of self-punishment transfer to state punishment, rendering it justified, we might ask whether self-punishment might constitute part of the punishment an offender deserves, especially if self-punishment is actually the ideal. We must be aware from the outset, however, that repentant self-punishment is much more than just the experience of an aversive psychological state. It is remorse *plus* engagement in the penitential behaviour it often motivates. It is this ‘active remorse’ that plausibly could constitute some of the deserved punishment, within Duff’s theory. We can now explore how it might do so.

Let us look first at punishment’s function of facilitating a fully repentant understanding of the wrongdoing. It might be argued by Duff that full repentance cannot be achieved passively. Indeed, he is cynical about the success individuals can have with experiencing the emotions and cognitions which would characterise the

optimal repentant response: ‘repentance is achieved through time and only with difficulty; and the difficulty flows from our unwillingness to confront our wrongdoings’ (ibid.: 165). He imagines (almost) perfect beings that recognise and fully repent their wrongs as soon as they are done. These beings, he says, would not need penances or punishments as distinct from the pain which belongs with repentance. In contrast, we mere mortals, it would seem, require some kind of ‘hair shirt’. But the question then becomes whether a self-imposed hair shirt can suffice. Duff might argue that a self-imposed penance can be effective, but that it still has to be prison – only this is sufficiently mind-focussing. If this were the case, then Duff’s vision of the ideal of self-punishment would require that every novice repentant offender would request to go to prison in order to develop his or her repentance (and to express it).

However, this does not seem to be Duff’s position. Duff gives the example of the offender who, having committed acquisitive offences, embarks on a period of material deprivation: ‘this would express and reinforce my condemnation of the excessive concern for material goods which my action manifested’ (ibid.: 165). One could argue that self-imposed penances are actually better suited to focussing one’s mind especially if, like Duff’s example and the graffiti example above, they involve an offence-related element. Unlike the pain of being in prison (of which one would of course be acutely aware) the pain of these sorts of self-imposed penance forces the offender to think about the exact nature and consequences of the wrong.

Further, if one chooses to impose something on oneself then one is somehow motivated to do so, and in these cases it is because of what one has done. State punishment, on the other hand, even if not ignorable, does not have this psychological link built in. Indeed, Duff sees the burden of proof on this matter to rest with the

justification of state punishment. He argues that in order to justify state punishments as penances we would ‘need to show how, *just as a self-imposed penance can assist repentance*, so punishment which is imposed by others can help bring a wrongdoer to repentance (and can thus become a penance)’ (ibid.: 166).

It would therefore seem that on the question of an offender becoming fully repentant, self-imposed penance (motivated by at least inchoate repentance) can suffice. According to Duff, an offender cannot become fully repentant without this assistance. Although this is an empirical matter, if we accept his assumptions then remorse must be accompanied by self-punishment in order to achieve full repentance. However, there seems to be nothing in Duff’s argument to suggest that repentant self-punishment cannot match or even exceed state punishment’s effectiveness in this regard.

Expressing repentance

What of the expressive function of punishment? Again, Duff imagines (almost) perfect beings for whom no symbolic penance would be required. He says that ‘they would know of each other that their repentance was genuine and their apologies fully sincere’ (ibid.: 166). However, beings such as ourselves need to provide an outward and public expression of our painful remorse in order to assure others of the sincerity of our repentance. That we are incapable of this sort of perfection is more clearly the case than the assumption that we are incapable of achieving full repentance without a mind-focussing punishment. We are not mind-readers and, unless we know someone really well, we require a lot of behavioural evidence in order to make confident psychological attributions. Even more prohibitively, we would not even know of this behavioural evidence unless we had seen it first hand or it had been made publicly

available. So, it would seem that Duff is right to require *some* sort of penance if the wrongdoer's repentance is to be expressed to the community. The question, as with the discussion of focussing the offender's mind, is whether self-punishment suffices.

It could be argued that a self-imposed penance is not public enough – formal state punishment is the only sort of penance of which the community will be aware. Tasioulas makes this argument, saying that a penance undergone at the offender's own discretion does not 'bear a public meaning accessible to the victim and the wider community' (2007: 505). Legal punishment, in contrast, is 'something inflicted on the offender irrespective of his choice, and it provides him with a *public* vehicle for manifesting his repentance' (ibid.). The issue of whether voluntariness *per se* makes the penance more or less suitable aside, the question really is whether a self-imposed penance can adequately express remorse. It would seem that this will not invite a categorical answer. Obviously, a hair shirt concealed by one's garments will express nothing, as it would not be seen.

Not all penances are so private, though. Imagine a woman who, after significant tax evasion, wishes not only to pay what she owes, but to serve society, instead of take from it. She commits herself to working for free, serving soup in a homeless shelter four nights a week. Moreover, when people enquire why she decided to work there she openly tells them about her wrongdoing and that she wants to make it up to her community. Such a self-penance would be very public and it might become widely known about, at least within her local community. If this was not considered public enough, then it could easily be noted at sentencing, allowing anyone who is interested in the case to know of her self-imposed penance. The general public primarily know about prison sentences through the news media and, if self-imposed penances were documented alongside the state-imposed penance, they

would be just as ‘public’. Indeed, many more people would actually see the tax evader serving soup than would see her in prison.

Moreover, it strikes me that there is more to penitential expression than its ‘loudness’ (or, public character). There is also its quality. Even if one were to argue that a state punishment is still more public due to its formality, there is still the important point that the quality of a self-imposed penance is better than that of a state punishment. If Duff’s concern is really about expressive efficacy (as a function of *what* reaches *whom*), then the quality (the *what*) of the self-imposed penance cannot be ignored. Such penances express an offender’s repentant understanding with more certainty than state punishment. The community does not know if the prison ‘penance’ is accepted as such or not, it can only know that it has been imposed.

An objector might argue that a self-imposed penance can never express as much as a state penance because it is nowhere near as burdensome. This might be true in some cases: a violent assault, even if responded to repentantly, might require years in prison in order for repentance to be expressed sufficiently – there is no number of hours in a soup kitchen that could equate. In addition, even if there were, a judge could not know that the offender would continue to impose this work upon herself for years to come. Moreover, some offenders have very little time to engage in penance prior to being sentenced.

I do not think these objections threaten my argument. The question is not whether remorse – or, rather, a repentantly-imposed penance – eliminates the need for state punishment, merely whether it can constitute part of it. Even if the assailant cannot express his repentance in its entirety, he might express some of it, and convincingly so, to a very wide audience, if his penance were to be noted at sentencing.

Self-punishment can constitute some of the deserved punishment

It would therefore seem that self-punishment, in its mind-focussing and repentance-expressing capacity, can constitute – *is* – some of the deserved punishment, on Duff’s account. Indeed, for Duff, ‘self-punishment’ is a viable concept, implying a rejection of the third of Hart’s criteria: ‘this is deliberately imposed by an agent authorized by the system of rules against which the presumed offence is committed who is acting within his or her official capacity.’ Duff, in fact, specifies something different, that punishment is imposed ‘by someone with (supposedly) the authority to do so’ (Duff 2001: xv). Stating that punishment should be imposed by someone with the authority to do so is subtly different from claiming that it must be *an* authority that imposes the punishment. Having ‘the authority to do something’ means that one is entitled to do that thing. Being ‘an authority’ means that one is recognised as having a particular, formal, standing with relation to another person or group of people. It would make sense that we have the authority to impose punishment on ourselves.

What Duff’s criterion does is suggest that, *either* punishment is not punishment if it is imposed by someone without the authority to do so (perhaps it is torture or simply cruel treatment in such a case), *or* that punishment is only justified when imposed by someone who has the authority to do so. The unauthorised are then still punishing, but unjustifiably. Either way, the important point is that one can punish oneself – one is entitled. However, as Duff emphasises, the discussion of the definition of punishment must ‘rapidly become a normative discussion of how punishment can be justified, if it is to produce a useful account of what we *should* mean by “punishment”’ (ibid.: xiv, emphasis). Whether Duff wants to formalise his definition of punishment or not, it is clear that within his theory a remorsefully-imposed self-penance constitutes punishment, not merely achieves its aims.

That the offender's self-punishment should have weight at sentencing is consistent with another of Duff's arguments. He says that '...if the law's demands and the criminal trial are to show a due respect for the citizen's autonomy, they must be understood not merely in terms of what is *done to* or *imposed on*, the citizen, but in terms of a *dialogue* between state and citizen' (ibid.: p 160). If Duff really wants this *dialogue* to be a reality, then the state should not simply write the script for the offender. If his sanction is his penance – his way of expressing repentance – then this expression is chosen for him, and he may not even know what he is expressing. Instead, the dialogue should be open to listening to the offender's own expression. Duff's model would be patronizing if it were to refuse to accept that an individual can come to repent by himself and go some way to expressing this repentance (through imposing self-punishment) as a result.

Indeed, this is not really a surprising outcome for a theory of punishment that claims to model itself on – to aspire to be – self-punishment. At the outset, Duff tells us that self-punishment should be the model for, and justification of, state punishment. This being the case, self-punishment *must* be effective at the things that Duff requires of state punishment – what would be the point of modelling criminal punishment on impossibility? Duff ushers in the paradigm of self-punishment to justify state punishment. He explains the value of self-imposed penances because he wants to show penance as something worthwhile, and something to emulate in criminal justice. It would be bizarre if he were then to deny that self-punishment was capable of doing the things that he needs it to do in order for him to use it to justify the institution of punishment.

Duff sees the redundancy of state punishment as a possibility reserved only for those beings 'more perfect than ourselves', not even for those who are repentantly

imposing penances on themselves. If he insists on his claim that full repentance is impossible without it, then this conclusion is inevitable. However, I would argue that this is an overly cynical outlook on the human condition. In addition, the argument that only state punishment is ‘public’ enough was met with the challenge that penances undertaken voluntarily can be announced by a judge just as loudly as a formal sanction, and that the expressive quality of the former might actually be clearer. So, if self-punishment can fulfil the objectives Duff sets out for punishment – and I have argued that it can, perhaps even rendering it ‘punishment’ – at least some mitigation seems justified.

It should be reiterated that this argument requires that the offender has done a lot. It is not enough for him to feel the psychological state of remorse. He must also have embarked on an expressive, burdensome penance that can be communicated to, and understood by, the community as such. It is this penitential expression that can form part of the dialogue between state and citizen. This notion of acknowledging the offender’s response to his offence – his input into the communicative enterprise – will become even more important in chapter seven. There, however, it will be argued that the offender’s moral response should result in mitigation even if no burdensome penance has been self-imposed.

Duff’s vacillation

This most explicit of Duff’s work on self-punishment is both consistent and at odds with some of his remarks elsewhere. Indeed, a tension seems to run throughout all his work. On the one hand, Duff repeatedly cites his concept of penance as the constitutive justification for punishment. He also proposes that sentencers work with ‘negative proportionality’, which sets upper and lower limits on sentences but gives

‘sentencers room to attend to the concrete particularities of the crime, without worrying about rendering it commensurable with all other crimes in terms of its seriousness...’ (2001: 139). According to such a scheme, and if I am right to be optimistic about the efficacy of self-imposed penances to express repentance and the possibility reaching full repentance without state intervention, then remorseful self-punishment should be a key player in reducing sentences to their lower limits. (For a similar point, see von Hirsch and Ashworth 2005: 102).

On the other hand, elsewhere, Duff expresses explicit commitment to public communicative aims, despite these being of no consequence to an offender’s engagement with penance – the constitutive justification. Indeed, in relation to already repentant offenders, Duff argues that there should be no mitigation of sentence because ‘to impose a lighter sentence on a repentant offender is thus to imply that repentance renders the crime less serious. But this is not normally true’ (ibid.: 120). First, Duff seems here to deviate from his negative proportionality proposal, suggesting that there is only one appropriate level of punishment suitable for any crime of a particular degree of seriousness. In fact, he cites von Hirsch in relation to this point about maintaining proportionality. Later, however, Duff says that von Hirsch’s adherence to positive proportionality is wrong because it is too ‘strict’ a position. He says:

Proportionality, as a relationship between the seriousness of a crime and the severity of its punishment, is only one dimension of the proper relationship between crime and punishment on a communicative conception of punishment as a penance. A communicatively appropriate punishment should communicate not just a degree of censure proportionate to the seriousness of the crime, but a more substantive understanding of the nature and implications of the crime as a wrong. [ibid.: 142].

So, Duff’s work seems to be at variance on two levels. First, he vacillates between his stated constitutive justification for punishment, penance, and adherence to ancillary

aims of public valuation of conduct. Second he seems unclear as to the strictness that should be afforded to the proportionality requirement. Moreover, the levels intersect: whilst a purely penance perspective would significantly dilute proportionality requirements (see von Hirsch and Ashworth 2005: 102), commitment to the public communicative functions of punishment may serve to reinforce these requirements. However, Duff seems to be undecided on this point too, suggesting that communicative punishment has communicative functions that transcend proportionate censure.

Ambiguity aside, what has been shown is that if penance is taken as the constitutive justification for punishment, and given that self-punishment constitutes the ideal penance, then remorseful self-penance does constitute some of the deserved punishment.

The communicative aims, that Duff introduced inconsistently, are central to the final theory we will consider: von Hirsch and Ashworth's work on censure and sanction.

Theory three: Von Hirsch and Ashworth: censure and sanction

For von Hirsch, punishment is constituted by both censure and sanction. He writes: 'punishing someone consists of visiting a deprivation (hard treatment) on him, because he has committed a wrong, in a manner that expresses disapprobation of the person for his conduct' (1993: 9). This emphasis on the censure-communicating facet of punishment suggests the need for a further criterion to be added to Hart's definition of punishment. Indeed, Joel Feinberg (1970) points out that Hart's definition fails to distinguish punishment from mere penalties such as overdue notices, late fees and so on. Punishment, when conceived as von Hirsch conceives it, includes an *expressive*

element. Thus, a further clause must be added to Hart's definition if we want to use it in relation to von Hirsch's theory. Plausibly: (vi) it conveys an appropriate degree of censure.

Now that we understand what punishment entails within this sentencing perspective, we can consider whether remorse might either constitute some degree of punishment, or at least be able to serve the functions von Hirsch and Ashworth require of punishment.

Remorse as a substitute for sanction

One potential way of demonstrating a degree of redundancy of punishment when faced with an offender's remorse emerges from the way in which the sanction (hard treatment element of punishment) is justified. Whereas some method of delivering censure is necessary to communicative retributive theories, von Hirsch argues that the imposition of sanction is justified contingently. The sanction provides an additional penal disincentive which helps agents adhere to the moral norms to which they may already be committed. Thus, a prudential reason for desisting is provided to supplement any normative ones that already function for the offender (see von Hirsch and Ashworth (2005: 23)). Justifying the hard treatment element contingently means that retributive justice could theoretically still be achieved without it. Indeed, von Hirsch and Ashworth argue that, within their theoretical framework, punishment (or, at least, the harsh sanction (as dissociated from any symbolic censure) loses its justification if it can be achieved by other means:

If the institution of the criminal sanction were incapable of carrying out these [expressive and preventative] public functions (or if those functions could be performed without having to resort to the intrusive features of punishment) then that institution would lose its *raison d'être*. [von Hirsch and Ashworth 2005: 100]

The question can then be raised: can remorse carry out any or some of the expressive and preventative functions, rendering the criminal sanction to any degree redundant?

Remorse as a deterrent

It could be argued that remorse might serve as some sort of deterrent, preventing offending. The intensely aversive nature of the experience of remorse, as described by Adam Smith (2002; see beginning of this chapter), is likely to be something that an offender would wish to avoid. Knowing now how wrongdoing makes him feel, the offender may use this as an additional prudential disincentive. He commits himself to avoiding engagement with the wrongdoing that elicits in him this terrible sentiment.

Moreover, there is something else important about this self-originating deterrent: experiencing remorse is much more certain for the offender than punishment prescribed by the court. It is an empirical fact that people are deterred more by the certainty of punishment than its severity (Bentham 1970). If the offender has reacted remorsefully on this occasion, she may believe it likely that she would react this way again, regardless of whether she was caught the next time.

If remorse did turn out to have this deterrent effect in reality, then there would be at least an argument for reducing the additional prudential disincentive, as the remorseful offender would be providing one himself. However, the existence of a deterrent effect of remorse is an empirical one and the evidence to date does not provide strong support for this effect. Cox's (1999: 17) discussion of the evidence leads him to conclude that '[t]he disturbing fact is that there appears to be no firm evidence that [had the remorseless offender] shown remorse, there is less likelihood of him re-offending'.

However, the intuition still remains that the fear of ‘feeling bad’ if one engages in some particular activity does serve to inhibit the desire to re-offend. The empirical reality is likely to be that some types of offences are susceptible to being partially deterred by remorse whereas others are not. This may account in part for the inconclusive research outcomes to-date. An offender might re-offend despite intense remorse and self-loathing if, for example, the offence is highly emotional or motivated by addiction. Jones (1956: 123) suggests that ‘deterrent methods are of less value in reducing the incidence of those crimes in which strong passions or deep psychological problems are involved’.

There is a paradox, though, in the suggestion that fear of remorse can serve as a deterrent. If remorse were to be genuine, then the offender should resist offending because he believes he should, not because he is scared that he will feel bad if he offends. Fear of feeling bad is more plausible where guilt is the emotional response to the wrongdoing. As discussed in Chapter Two, the concept of guilt developed by some theorists, which I called ‘corrupt guilt’, involves, principally, awareness of breaking rules and the fear of the consequences for oneself when one transgresses. It would be compatible with this sort of response to wrongdoing that an offender might resist offending for fear of experiencing it again. Following genuine remorse, however, the offender would be more likely to refrain due to wanting to and believing that he should, not because of a fear of aversive emotions. Remorse would provide normative reasons to refrain, not prudential ones, and von Hirsch argues that additional prudential disincentives are helpful even to those who have internalised the normative reasons to refrain.

There is, then, no strong argument that remorse could be a viable substitute for the prudential disincentive that the criminal sanction supplies. Furthermore, even if

remorse did have some deterrent effect, that would not make it eligible to fill the role of the sanction. This becomes clear when we return again to the justification for the criminal sanction. It is not just that it serves as a deterrent. The sanction is chosen as the means of conveying censure (as opposed to, say, purely symbolic means) because it has this additional function. But it is, first and foremost, a mode of expression: ‘the preventative function operates...only *within* the framework of a censoring institution’ (von Hirsch and Ashworth 2005: 24, emphasis in original). The sanction, then, does some of the expressive work – might remorse go any way to fulfilling *this* function?

Remorse as delivering censure

It seems natural to think of remorse as constituting self-blame. We talk of people emotionally ‘kicking themselves’ and being angry with oneself is a familiar concept. It is even possible to imagine individuals who actually reprimand themselves in their heads, saying the same sorts of things that someone in an official capacity might say to them. Perhaps some might even be harder on themselves than others would be. Might such self-blame legitimately be perceived as censure?

The barrier to such a possibility is that censure, particularly for von Hirsch and Ashworth, is characterised by its authoritative nature. In fact, they perceive this to be the crucial difference between blaming and censoring. On this view, the most the remorseful offender might be able to do is blame himself, but he does not have the requisite standing to himself to convey censure. Censure requires that the person or institution delivering it stands in authority to the wrongdoer. Thus, even if remorse did provide some prudential disincentive, it could not express censure in any meaningful way.

However, as implied by the assertion that ‘a condemnatory response to culpably injurious conduct might be expressed *either* in a purely symbolic mode, *or* through the visitation of hard treatment’ (2005: 24, emphasis added), hard treatment is only one option for expressing censure. Although delivering deserved censure is always going to be desirable within a retributive perspective, the criminal sanction does not *have* to be the way of achieving this. There is a theoretical possibility, then, if remorse were a powerful deterrent, the symbolic mode of expression would become adequate. However, I have already noted above that von Hirsch and Ashworth would not accept this, due to their pessimism surrounding the ability of average people to refrain from offending motivated purely by normative reasons. This is what leads them to argue that a supplementary prudential disincentive be provided by the threat or experience of hard treatment. I shall expand on this further.

Remorse as reattachment to normative reasons to refrain

Von Hirsch and Ashworth would probably argue that although remorse demonstrates attachment to normative reasons for desistance, this does not make prudential ones redundant. Their theory proposes that the criminal sanction is an *additional* disincentive. They already recognise that many people are motivated by normative reasons to resist offending. The assumption is that will power and intention alone are not strong enough preventative forces. In some cases it may take wrongdoing and subsequent remorse to wake an individual up to these normative reasons: the reasons become real for him. The point is that von Hirsch and Ashworth see an additional prudential disincentive as necessary even for people who are already attached to their moral convictions. This is what prevents their prudential disincentive from being ‘beast control’: it aids people in adhering to norms to which they already know they

should be committed. Indeed, whilst general deterrence might be powerful enough for most, an offender's remorse might show him to be in *particular need* of specific deterrence.

So, given their theory, von Hirsch and Ashworth would have to accept that if no additional disincentive were needed, and there was an alternative, equally effective way of communicating, then the harsh sanction would be redundant. They endorse the use of hard treatment as the means of communication since it also performs the secondary, preventative, function. Since remorse cannot perform either the preventative or expressive function, mitigation cannot be justified on these grounds.

Conclusion

This chapter has considered whether remorse might constitute some of the deserved punishment, or otherwise provide a viable substitute. In relation to the benefits and burdens theory, remorse could do neither. Despite punishment not necessarily requiring incarceration on this account, the appropriate burden required objective restriction of freedom, which necessarily had to be imposed by an external agent.

The case for remorse constituting some of the deserved punishment was strongest in relation to Duff's work. I argued that remorse-motivated penance could actually *be* punishment within his account, especially as penance is his constitutive justification. The reasons why Duff would disagree were found in his pessimism about the expressive efficacy of self-imposed penance and about the capacity of human beings to reach full repentance without state intervention. I challenged both of these. He also disagrees on the grounds that communicative aims require that there is no deviation from proportional censure. This was shown to be an inconsistent line of thought throughout his work. It was noted that the very most that could be argued here

is that remorsefully-imposed *penance* could constitute punishment. Remorse alone, or half-hearted attempts at expressing one's remorse, cannot suffice as they do not adequately express one's repentance to the community.

In relation to von Hirsch and Ashworth's perspective, it was demonstrated that remorse cannot constitute punishment as, for them, punishment requires that it be imposed by an authority. Further, remorse could not go any way to achieving the aims of punishment since it provides no prudential disincentive additional to normative reasons. Indeed, any disincentive that remorse provides is predominantly in the form of these normative reasons. Moreover, even if remorse did have a deterrent effect, it could not adequately fulfil the expressive function of punishment, as censure is necessarily authoritarian in character.

Having found no strong justification for mitigation of the grounds of remorse, based on its constituting or substituting for deserved punishment, I shall now move onto an alternative possibility. The following chapter will explore whether remorse might reduce the seriousness of the crime (the culpability of the offender and/or the harm caused) and thus justify a reduced punishment. For this argument, it would not matter that remorse does not constitute any of the deserved punishment, since it instead would mitigate what is actually deserved.

Chapter Seven

Can remorse reduce the seriousness of the offence?

Within retributive theories of punishment, the severity of the punishment the offender receives must be proportional to the seriousness of his offence. Thus, if the seriousness of the offence is reduced, the punishment must also be reduced. The seriousness of the offence is a function of the harm caused or risked by the offender, and the extent to which he was culpable (roughly, responsible) for it. If an offender's experience of remorse serves to reduce either the harm of the offence, his culpability for it, or both, then he should receive less punishment than would otherwise be deserved. This chapter will consider both these possibilities, attending to harm in the first section and culpability in the second. It will ultimately be argued that remorse does not have a relevant bearing on either and so does not reduce the seriousness of the offence.

Section One: Harm

The degree of harm an offence causes or risks is one of two elements that determine the seriousness of the offence. The greater the harm caused or risked, *ceteris paribus*, the more serious the offence. It is uncontroversial to state that common assault is less serious than murder, or that petty theft is less serious than burglary with ransacking. In both examples, the difference is constituted by the degree of harmfulness of the respective offences.

In order to come to a conclusion about whether an offender's remorse can reduce the harm of the offence, we will have to explore what is meant by the harm of the offence, and assess whether remorse has any impact on this dimension. If it is

plausible that it does, we must then ask whether this effect is relevant for sentencing purposes. With close reference to the most comprehensive discussion of how the harm of the offence should be conceptualised and assessed (von Hirsch and Jareborg 1991), I will construct the strongest argument for the thesis that remorse does reduce the harm of the offence. I will then level objections that I argue render the argument ultimately unpersuasive.

However, before doing so, I will bring to the reader's attention the not uncontroversial nature of focusing only on harm as (in addition to culpability) the determinant of offence-seriousness.

Harm as the relevant factor

The controversy arises within a more general debate about whether an action's being harmful is necessary or sufficient for criminalisation. The answer to this question has consequences for the way in which we should assess the seriousness of an offence – on the features relevant to its gravity. The debate is broadly between those who adhere to some formulation of the 'harm principle', and those who see this conception of what justifies criminalisation as deficient. The latter group would, for example, supplement or replace considerations of harm with those of wrongs or violations of deontological rights. The motivation for disagreement frequently stems from proposed hypothetical scenarios that involve a wrong that causes no harm, where the intuition is that the wrong should be criminalised (see below).

Dissatisfaction with a focus solely on harm as the non-mental element of offence-seriousness is also hinted at in the sentencing literature. Even those who conceptualise offence-seriousness as being determined by the harm of the offence and the culpability of the offender acknowledge that harm is not always a rich enough

concept to capture what we should take into account when determining the seriousness of an offence. The notion of many offences as wrongs – as well as or instead of – harms is considered by some to be important to recognise. Summing up this sentiment, Ashworth writes:

...the violation of a protected interest is one key component of offence-seriousness, often expressed as harm or harmfulness but also including the concept of a wrong, since it is not merely the physical or psychological consequences but also the *nature of the wrong* done to the victim that is relevant in assessing seriousness. [2010: 109, emphasis added]

Let us conceive of a wrong with no harm. Imagine a stranger (there is no possibility of implied consent) takes and returns your car without your knowledge, even filling it up with petrol and fitting new brake pads. In this scenario, it is hard to see what *harm* is done to you. Some theorists seem to be able to find ‘harm’ in any hypothetical example of harmless wrongdoing. However, in doing so, the notion of harm becomes so stretched that it seems to incorporate destruction of rights – the result being that all wrongdoing is also harmful (see Stewart (2010) on this debate).

Von Hirsch and Jareborg’s (1991) account of how harm should be conceptualised for the purposes of sentencing is one example of this extensive use of the term ‘harm’. In the above example of the stranger taking your car, they would explain that there was harm done in the form of violation to your interest domain of privacy/autonomy. Others, in comparison, might understand this scenario in terms of violations of rights – something which they would distinguish from harm and the frustration of interests.

Von Hirsch and Jareborg provide the most comprehensive treatment of harm in relation to sentencing. It is their account on which I will focus in order to ask whether there is any plausible way in which remorse might reduce the harm of an offence. Even if there is disagreement about what should be labelled harm and what

wrong, their inclusive interpretation of harm will at least not leave anything relevant out, even if their conceptual approach could be contested.

Von Hirsch and Jareborg's 'living standard' analysis of harm

Different offences harm victims in different ways – they involve intrusions into various personal resources or interests that people have. Von Hirsch and Jareborg aim to find and articulate a guiding idea which can be used to focus on and assess the various relevant interests involved. They suggest that the most natural organising concept has to do with the quality of a person's life: 'The most important interests are those central to personal well-being; and, accordingly, the most grievous harms are those which drastically diminish one's standard of well-being' (von Hirsch and Jareborg 1991: 7). This guiding concept they term 'living standard' and they intend it to include not only material support and amenity, but other non-economic capabilities that affect the quality of a person's life.

Von Hirsch and Jareborg suggest how harms can be gauged according to the extent to which they affect someone's living standard by constructing a grading scheme. Their scheme consists of four living standard levels: level one is categorised as 'subsistence', level two as 'minimal well-being', level three as 'adequate well-being' and level four as 'enhanced well-being' (ibid.: 17).⁴² Offences can comprise intrusions into interests required to maintain any of these levels. The most serious offences intrude into interests required for subsistence (level one). Examples would be murder or mayhem. Intrusions that only marginally affect the living standard, and do not prevent the victim from maintaining an

⁴² For a description of each living standard level, see von Hirsch and Jareborg (1991: 17). Detailed knowledge of every concept employed by von Hirsch and Jareborg will not be necessary to follow my argument. Where detail is required, it will be elaborated in the discussion.

adequate life would be ranked at level 4 (enhanced well-being). An example would be theft of a small amount of money.

Even within these examples, we can see that the intrusions the offences make are of different sorts. The former comprise fatal or significant damage to physical well-being, whereas the latter is a very minor dent to financial well-being. Von Hirsch and Jareborg suggest four generic-interest dimensions that can be intruded upon, that they believe the state has reason to protect. These are: ‘physical integrity’, ‘material support and amenity’, ‘freedom from humiliation’ and ‘privacy/autonomy’ (ibid.: 19).⁴³

Different offences will intrude upon one or more of these interests, and to different degrees. For example, assault affects the victim’s physical integrity and his freedom from humiliation, whereas burglary mainly affects the victim’s privacy and his material possessions (ibid.: 19).

How might remorse repair a victim’s living standard, or prevent its decline?

Only one of the generic interest dimensions seems to have any potential for being affected by the victim’s awareness of the offender’s remorse. Remorse could do nothing to repair a victim’s physical integrity – bodily wounds will heal no more quickly. Nor could remorse enhance the material support and amenity the victim enjoys – barring acts of reparation, which go beyond an offender expressing his remorse, the property the victim owns remains unchanged. Privacy and autonomy are also not affected: the offender’s remorse does not revoke any intrusion into the victim’s personal space; nor does it render any of the criminal activity of the victim’s choosing.

⁴³ For a description of each interest dimension, see von Hirsch and Jareborg (1991: 20). Again, detailed knowledge of every concept employed von Hirsch and Jareborg will not be necessary to follow my argument. Where detail is required, it will be elaborated in the discussion.

However, if a victim is made known of the offender's remorse, he may no longer feel humiliated. Von Hirsch and Jareborg explain that freedom from humiliation – or 'degrading treatment' – 'refers to those injuries to self-respect that derive from others' mistreatment' (1991: 20). The perception that victimising criminal activity sends a message to the victim that may result in his feeling humiliated or losing self-respect occurs elsewhere in writings on retribution.

Murphy (2003) argues that victims often perceive part of the harm of the offence to be the insult or degradation that the offence symbolically communicates. Through the offence, the offender sends a message to the victim, saying that the victim is inferior and exploitable. This causes the victim distress. He may even feel that he *is* inferior and exploitable, and blame himself. However, if the offender repents, then the message is withdrawn and some of the harm lessened. Hampton (1988) espouses a similar conception of the symbolic communication, suggesting that the 'insult' conveyed by the offender can prompt the victim to question his own value, conceiving that the victimising conduct was somehow justified by his lack of worth.

So, if such perceptions do indeed play out in the experiences of victims, it is possible that remorse might reverse, or undermine, or counterbalance the message which the criminal behaviour communicated. Knowing that the offender is remorseful could serve to reaffirm a victim's sense of worth – the victim thinks of the offender: 'he is (justifiably) feeling bad because he wronged *me*...he sees me as valuable enough to dwell on the harm he caused me and regret it'. Thus, the victim may no longer feel humiliated as he ceases feeling worthless.

The living standard time-frame

If we grant that remorse may have the effect of reducing or eliminating the humiliation an offender feels, allowing him to regain self-respect, we still need to consider whether this can be relevant for sentencing purposes. One source of uncertainty relates to the temporal perspective of von Hirsch and Jareborg's living standard analysis. They consider what sort of temporal perspective should be adopted in order to appropriately judge the impact of an offence on a victim's well-being, concluding that a mid-term perspective is suitable: 'The appropriate perspective is a middle-term one – something approximating "How has your year been?"' (ibid.: 22). This is appropriate given that living standard judgements are intended to relate to the global quality of a person's *life*.

So, when assessing the intrusion into the physical integrity of an assault victim (for example), we need to think longer-term than whether he was bleeding and in pain on the night of the offence. Perhaps – and I speculate – pain and/or bruising would have to persist for at least a month or so in order to qualify as significantly affecting one's year physically. Or, if the pain inflicted was so severe, even though it may not persist much beyond the assault, it may by far constitute the most excruciating physical experience of the victim's year. In this vein, von Hirsch and Jareborg consider that 'a physical assault, and its immediate trauma, may soon be over; yet if the experience was painful or humiliating enough, it may still loom large in the evaluation of, say, the quality of a whole year's experience' (ibid.: 22).

Although the experience must be detrimental to the victim's living standard from the perspective of the proposed timeframe in order to be significant, it must be remembered that this time frame remains mid-term. If, in four years time, the victim

cannot even remember what pain or injury he sustained from the assault, this does not detract from the significance it had within the year that immediately followed.

So, as suggested, it is plausible that, were a victim made aware of the offender's remorse, humiliation that may have affected the victim's well-being for many months might be alleviated. The above remarks from von Hirsch and Jareborg attest to their belief that it could be the humiliation precipitated by an assault *alone* ('painful *or* humiliating enough') that diminishes the victim's living standard, from the mid-term perspective. The harm that can be forestalled, therefore, is not insignificant. The victim may actually (perhaps not virtuously) feel a sense of power in the face of the offender's remorse – it is now the offender who experiences humiliation. Thus, from a mid-term perspective, the effect on the victim's living standard may be altered significantly. The argument then takes a final step: if knowledge of remorse reduces the harm done to the 'freedom from humiliation' interest dimension, it thus serves to mitigate the seriousness of the offence. If the offence is less serious, less severe punishment is justified.

Challenge One – Distinguishing between the time-frame from which living standard judgements are made and the time-frame of the offence

However, we need to disentangle the two distinct temporal dimensions involved in judgements of harm. On the one hand, there is the actual duration of the offence; on the other, there is the timeframe from which we should view the importance of the offence and its effects for the living standard of the victim. It is not the case that the seriousness of the offence continues to be established thorough out the year. Whilst it might be true that the importance of the offence and its effects might not become completely clear for some time, the offence has nevertheless already occurred. It is

not somehow still being committed. So, what we need to be clearer about is the timeframe of the *offence* – that is, how the offence is to be delineated as a complete action or actions – and the importance of the offence to the victim – in terms of continued reduction of his living standard – from the perspective of a given timeframe.

We can begin to make this clearer if we think about theft of a particular sum of money. Imagine that an offender stole £100 from a house. He climbed in through an open window, leaving no damage whatsoever. The following day, the offender returns and gives the victim the money back. In this scenario, we might wonder what harm remains: the victim has no less money than he started with. Perhaps knowing that someone had been into his house might still leave him feeling humiliated, and his privacy has still been violated. This being said, the total harm is conceivably less than experiencing humiliation and violation *and* being £100 poorer.

Whether we see this instance of ostensible harm reduction as relevant for estimating seriousness turns on how we delineate the ‘harm of the offence’. It could be argued that the burglar of £100 burgled £100 at the time of the offence and that *this* constitutes the degree of harm. A broader construal might allow that the harm of the offence consisted in the taking and returning of £100.

However, such an extended perspective on the duration of the offence becomes unrealistic once the criminal activity is discontinued. For example, as long as the burglar is on the victim’s premises, he is arguably still committing the offence, but as soon as he leaves and pursues alternative activities he is no longer committing the offence and thus its harm has been established.

If we were not to delineate the harm of the offence in this way, it could potentially be indefinitely indeterminate, and this is neither theoretically rational nor practically desirable. For example, the harm of a serious assault may not be fully

determined until the doctors and surgeons treating the victim have completed the course of treatment, which may be a matter of weeks – should we be waiting to find out how bad the victim ends up once stabilised before assessing harm? Moreover, different doctors and surgeons could have different levels of success, making the ‘final’ degree of harm somewhat a matter of luck. In a way, the level of harm might be indeterminable until the victim’s death: the £100 that did not seem such a big deal at the time might be seen as lifesaving when, as an elderly man, the victim has no money to heat his house in a freakishly cold winter month. Some harms might even conceivably emerge or persist after the victim’s death: one’s honour or memory can be blemished indefinitely.

So, in light of the indeterminacy of harm over time, we need to draw a line.⁴⁴ The same sort of limiting argument suggested for the harms of burglary can be applied to remorse-begotten reductions of humiliation and low self-worth. Although the offender’s remorse might go some way to assuaging the victim’s continued feelings of humiliation and worthlessness, the offence nonetheless constructed conditions such that the victim would otherwise have experienced humiliation. Had there been no known remorseful response from the offender, the victim’s humiliation may have loomed large in the evaluation of the quality of his year. Moreover, if the humiliation at the time of the offence was particularly severe, then it may still be significant from the mid-term perspective, despite not persisting. Since the offender’s

⁴⁴ Restorative justice perspectives would be much less concerned with determining the precise level of harm of the offence and instead focus on how a remorseful apology might alleviate the victim’s ongoing harm. From a retributive perspective, however, we need to make an assessment about the seriousness of the offence so that we know how much censure/punishment is deserved.

remorse is not within the duration of the offence⁴⁵, it cannot eliminate the *potential* that was created by the offence for a reduction to the living standard of the victim.

However, there still remains the intuition that the unfulfilled potential for humiliation is not as serious as an experience of humiliation that dampens a whole year. The argument might then be modified to propose that remorse-abated harm from humiliation could be considered at sentencing in a similar way to aborted attempts or offences of endangerment and risk-creation (e.g. dangerous driving, unsafe working conditions). We will consider these possibilities next.

Challenge Two – Can remorseful harm-reduction plausibly be considered in a similar vein to aborted attempts or risked harms?

If awareness of the offender's remorse prevents significant humiliation, then perhaps such circumstances should be seen as akin to aborted attempts or offences where there is greater harm risked (the offender was reckless) than actually transpired.

However, there are problems with drawing comparisons with attempts which render this line of thought unpersuasive. The concept of attempt would be misplaced, as it refers to the offence itself, which *was* completed, rather than to any of the secondary effects of the offence. The offender still assaulted the victim, it was not aborted and it did not fail. The subsequent alleviation of some of the effects of the offence does not change its status. All elements of the offence are present, so the offence is complete. To suggest otherwise would again be to confuse the duration of the offence with the importance of the experience from the overall mid-term perspective.

⁴⁵ But, see Duff (2001: 120-1) on 'immediate repentance'. This concept will be discussed in the following section in relation to culpability.

The concepts of endangerment and risk-creation may seem to have more purchase. Von Hirsch and Jareborg explain how their model makes room for discounts for threatened or risked harms. They say that ‘many crimes only create a threat or risk to a given interest. Their harm rating should depend not only on the importance of the interest but the degree to which it was risked’ (ibid.: 30). In our scenario, the offender creates the risk of substantial humiliation but then prevents it from materialising by remorsefully revoking the messages of degradation. This could be seen as him having risked or threatened humiliation. However, there are still temporal issues to resolve. It would have to be decided within what time frame a domain of harm could still be affected by the offender such that it was only risked or threatened. The difficulty in placing temporal limits on a post-offence downgrade from actual to risked harm again highlights the distance of the remorseful response from the elements of the offence.

It does not seem plausible that such allowances could be permitted in other domains of harm. If an offender returns the stolen goods this constitutes reparation, not risked threat to the interest dimension of material support. It would, if anything, provide basis for an argument for mitigation outside the assessment of offence seriousness. However, humiliation does seem the most fluid and flexible domain of harm. An offender might not fully experience humiliation until a few days after the offence. This would not prevent its status a harm of the offence.

If these misgivings have not yet succeeded in removing the idea from the table, there are further objections that have more force.

Challenge Three – Humiliation is usually not the only harm and frequently is the less serious

Even if we entertain the idea that the remorseful prevention of significant humiliation can be conceptualised as a threatened harm there are still barriers to its resulting in mitigation. Humiliation will rarely be the only harm caused and it will often be less serious than the violations to other interest dimensions.⁴⁶ Indeed, in their examples, von Hirsch and Jareborg refer to humiliation as a ‘secondary harm’ (ibid.: 31). These secondary harms they operationalise as ‘exacerbating features’ in their model (ibid.). This would mean that risked (rather than actual) humiliation would at most serve to preclude exacerbation of the seriousness of the offence, rather than mitigate it.

Challenge Four – What is really meant by ‘humiliation’ within the living standard account?

An interlocutor might still remain adamant that my objections have not entirely defeated the argument that an offender’s remorse reduces the harm that should be accorded to the offence. Within von Hirsch and Jareborg’s account, the concept of threatened harms *is* operant, and it is coherent that known remorse may conceptually reduce the violation of freedom from humiliation to a threatened or risked harm. However, there are still further problems that I will argue render this irrelevant to sentencing.

I am going to argue that von Hirsch and Jareborg employ a concept of humiliation that does not actually lend itself easily to the arguments I have discussed. Recall the descriptions (noted above) employed by Murphy and

⁴⁶ Rape could be an exception to this.

Hampton. They explain how the commission of an offence sends out an insult or message of degradation to the victim. It was suggested that this leads the victim to feel humiliated and without self-worth. Knowledge of the offender's remorse, it was suggested, can retract the message and enable the offender to regain his self-esteem.

However, that von Hirsch and Jareborg have in mind such abstract exchanges of messages – serving to precipitate and (perhaps) alleviate humiliation – is questionable. The first clue to this is that they seem only to see violent offences as damaging the interest of freedom from humiliation.⁴⁷ They state that this interest is 'affected by a variety of criminal acts, from physical assault to verbal harassment' (1991: 20). Moreover, when they apply their living standard analysis to hypothetical cases, it is only in relation to violent offences that they cite the interest of freedom from humiliation as having been intruded upon.⁴⁸ This is in contrast to the ideas espoused by Murphy and Hampton, who would consider messages of degradation as also being sent out by the offender who, for example, ransacks a person's house: 'You do not matter enough to me to prevent me from doing what I like with your possessions'.

Looking more closely at some of von Hirsch and Jareborg's examples we can begin to understand why their concept of humiliation is closely tied to violent offences, and in so doing better understand their concept. They say that what makes a beating 'deeply humiliating' is the 'being put at someone else's mercy. The person beaten is literally abased – knocked down, abused – and the beater establishes direct *physical* dominion over him' (ibid.: 25, emphasis added). This

⁴⁷ Although, at odds with what features in their examples, they do say that 'the idea of humiliation and the idea of loss of privacy are closely related: if one's privacy is intruded upon, one is almost necessarily somewhat humiliated' (1991: 32).

⁴⁸ The cases where freedom from humiliation is deemed to be intruded upon are: 'assault and battery', 'petty assault', 'forcible rape', and 'date rape'.

focus on the physical relationship of the victim and the offender is different to, although not incompatible with, the offender assuming an ostensibly-elevated status of worth in relation to the victim.

In von Hirsch and Jareborg's above example, we can imagine the offender sneering down at the frightened, bleeding victim, smirking at his situation. This contrasts with the more subtle message that Murphy and Hampton envisage. The burglar may have no thoughts about the victim, but his behaviour suggests that he does not see the victim as being of significance. This is the insulting aspect of the message – 'I don't have to think about you, I can just do as I please' – and not necessarily a triumphant display of brute power.

Attention to a second example serves to further reveal the precise concept of humiliation von Hirsch and Jareborg have in mind. In relation to 'petty assault' they say that 'having one's face slapped *is* humiliating. But one is not being made helpless, as in the case of the beating. One can extricate oneself with dignity – remonstrate, move away, call the authorities, etc' (ibid.: 25).⁴⁹ Here we can see again that von Hirsch and Jareborg have in mind a more public-related conception of humiliation as a physical power struggle. One may indeed remain dignified, but this does not change the symbolic communication from the offender. The communication would remain the same whether the offender responded with resentment or a more dignified indignation.

Of course, I agree that the humiliation and loss of self-respect is much greater in the case of assault and battery than it is in the case of petty assault, but the point remains that von Hirsch and Jareborg seem to have a more objective and public conception of what constitutes humiliation resulting from victimising

⁴⁹ The observation that – in the face of 'insult' – one can choose to protest with dignity instead of fighting back mirrors Hampton's distinction between 'indignation' and 'resentment' (see Hampton 1988: 35-87).

behaviour. It seems they are primarily concerned with physical power relations and the offender's exercise of this power. Indeed, they rephrase violation of the interest of freedom from humiliation as 'degrading treatment'.

This emphasis on the behaviour of the offender, rather than the feelings of the victim mirrors the distinction between being humiliated and feeling humiliated. It is not irrational to describe an interaction between two people as humiliating even in full knowledge that no humiliation is felt.

Perhaps it is not surprising that they want to utilise this more objective conception of humiliation given their additional comments on 'psychological harm'. Of course, it is not that they do not consider that feeling humiliated will follow from being treated in a degrading manner – it is precisely this that makes degrading treatment psychologically harmful. It is simply that they focus more on the behaviour that precipitates these feelings as the violation, not the actual feelings that may or may not result.

Von Hirsch and Jareborg provide an explanation for why they do not want to include 'psychological harm' as a legitimate interest for the state to protect. Some people may fear irrationally, for example, and it is undesirable for the harm of the offence to depend so much on the capricious subjective experiences of victims. However, they acknowledge that many aversive psychological states *justifiably* flow from various forms of criminal victimisation. They propose that these rational modes of distress constitute part of the intrusion into the interest dimension involved. Thus: 'when examining conduct that affects self-respect...the sense of humiliation is the injury that reduces the living standard' (ibid.: 23).

Conclusion

It has been argued that the arguments for remorse reducing the harm of the offence are unpersuasive. Conceptualising remorse-prevented humiliation as downgrading the harm from actual to threatened or risked seemed *prima facie* possible. However, it was rendered ultimately unconvincing. The difficulty in placing temporal limits on a post-offence downgrade from actual to risked harm highlighted the distance of the remorseful response from the elements of the offence. Moreover, the concept of humiliation that von Hirsch and Jareborg employ was shown to be less amenable to the effects of remorse, given its objective nature. They do see the concept of ‘risked’ humiliation (ibid.: 31) as plausible. However, it would likely be the *nature of the degrading treatment* – completed within the duration of the offence – that would determine the ‘risked’ compared with ‘actual’ status.

Even if they were to acknowledge a harm reducing effect of remorse, it could only lead to:

- 1) an absence of exacerbation (not quite the same as mitigation) for
- 2) violent crimes with
- 3) a clear victim who
- 4) knows about the remorse and
- 5) ceases to feel the humiliation that he would otherwise have felt.

Incorporating this into sentencing practice does not seem theoretically attractive. It would limit the ‘mitigating’ role of remorse to the few cases that fit the description above.

Having argued that remorse cannot reduce the harm of the offence in a way that is relevant to sentencing, I will now turn to the other of the two components of offence seriousness: culpability.

Section Two: Culpability

Culpability in retributive sentencing refers to the mental elements of the offence that, in combination with the harm caused or risked, establishes the seriousness of the offence. These mental elements include factors of intent, motive, and circumstance that determine the extent to which the offender should be held accountable for his act (see von Hirsch (1986: 64-5)). The concept of culpability employed by retributive theorists is by definition temporally static: once the offence has been completed, the offender cannot retrospectively change his state of mind at the time of the offence. This dependence on viewing crimes as categorically ‘complete’ in this way may seem somewhat ‘formalistic’ – a crime is ‘complete when we say it is’ (Hoerber 1986: 377). However, for the purposes of assessing crime seriousness ‘the punishment a defendant deserves is, to put it somewhat metaphorically, fully congealed at the time of the crime’ (Garvey 1996: 1030). It would therefore seem that an offender remains just as culpable as she was, regardless of whether she has subsequently experienced genuine remorse.

Overview

After considering a possible exception to the temporally static concept of culpability, I will broaden out the discussion to consider questions of personal identity, and concepts of blameworthiness related to, yet distinct from, culpability. Taking my lead from arguments found within von Hirsch and Ashworth’s theory, I will first consider whether the remorseful person is in fact a different person from the one who committed the offence. Having argued that remorse does not effect such a metaphysical transformation, I then explore the possibility that an offender’s remorse might render him less blameworthy. Two versions of this thesis will be examined: that

remorse reduces the offender's responsibility for his wrongdoing, and that remorse makes him a morally better person. The former possibility will be shown to be incoherent, whilst the latter will be seen to be coherent *per se*, but inappropriate for the purposes of retributive sentencing. This will lead me to consider deserved blame – as distinct from any notion of blameworthiness – as a more promising line of enquiry in the following chapter.

'Immediate repentance'

Anthony Duff makes a concession which begins to stretch, temporally, the notion of culpability. He argues that if an offender is 'immediately repentant' then this can 'cast a different light on her crime by showing it to have been a momentary aberration' (2001: 120). He claims that this different light reveals that she was not wholeheartedly committed to her crime, but that the crime instead was 'an aberration for which she already condemns herself' (ibid.). Duff therefore thinks that immediate repentance alters the offender's culpability – the offence is less serious as a consequence.

Such an argument does not allow remorse *per se* to retrospectively affect culpability. Rather, the immediate remorse leads to the inference that the state of mind at the time of the crime must have been such that we would consider it less blameworthy than, for example, a cold, hard, commitment to harm. However, this inference, although probably valid in many cases, does not describe something that is true by necessity. Although psychologically improbable, there is nothing logically inconsistent in the idea that someone could be committed to his crime in the most reprehensible way, and then immediately feel repentant.

I would suspect that if Duff were presented with such a scenario, and knew it to be such a scenario, he would not want the immediate repentance to be seen as

mitigating culpability. Rather, in his argument, he is giving the benefit of the doubt to the immediately repentant offender, considering it very likely that a more apprehensive and vacillating state of mind was present at the time of the offence. Perhaps ‘concurrent repentance’ might be closer to what Duff actually wants to make relevant to assessments of culpability.

His comment that immediate repentance shows the offence to be a momentary aberration has parallels with von Hirsch’s progressive loss of mitigation theory which explains the increase in sentence severity for persistent offenders. Von Hirsch argues that a first offender should receive a penal discount, partly as a sympathetic recognition of human fallibility – of ‘the all-too-human-weakness that can lead to a first lapse’ (von Hirsch 1991: 55). The additional justification he offers is that the first time offender has not yet had the chance to respond to the censure that punishment attempts to communicate to her. It is likely that immediate repentance can only be seen as showing the offence to be a temporary aberration in the case of first time offenders. Duff could therefore not argue on the grounds of the temporary aberration, confirmed by immediate repentance, for anything other than first time offences.

Despite seeming to make room for some temporal elasticity of the concept of culpability in cases of immediate repentance, Duff explicitly rules out a more significant mitigating role for remorse. This is due to his ‘penance perspective’ of punishment. As discussed in Chapter Six, he argues that repentance is not something that ‘can be achieved and completed in a moment’ (2001: 119). Repentance, he suggests, is deepened and strengthened appropriately by the penitential punishment. He may be right that repentance is not a momentary process, and it was argued in Chapter Three that a genuinely remorseful response will linger. This being said, I

argued in Chapter Six that is not obvious that the deep understanding that Duff hopes for is impossible to achieve without punishment.

However, he has a further argument which, given the structure of his penance perspective, more decisively rules out mitigation on the grounds of remorse within his account. On Duff's account, punishment serves the aims of reconciliation with those the offender has wronged (both the victim and the community as a whole). That reconciliation is achieved through penal hard treatment, which constitutes a forceful and public apology (see Duff 2001: 199). The whole sentence serves the dual purpose of inducing *and* expressing repentance. Thus, the offender's repentance would not adequately be expressed if he received a mitigated sentence.⁵⁰ Although Duff expressly denies that remorse should have any mitigating role (except in the case of 'immediate repentance'), we do not need to conclude that no justification can be found. We will see later that the justification I develop - inspired by Duff's commitment to communicative censure, von Hirsch and Ashworth's 'quasi-retributive' grounds for mitigation and attention to a distinction between blaming and blameworthiness - has significant implications for even the penance perspective.

Von Hirsch also employs a temporally-fixed understanding of culpability within his work (see, for example, 1993: 29-30). However, from the outset there appears to be more scope for incorporating a mitigating role for remorse into his account based on concepts closely related to culpability. I will argue that arguments he makes (in collaboration with Andrew Ashworth) to support the mitigation of the offender who is sentenced after a time delay would require him to consider similar arguments in relation to remorseful offenders. The argument that I will ultimately develop (in Chapters Eight and Nine) is not only *compatible* with his account, but is

⁵⁰ However, I argued in Chapter Six that self-imposed penance, borne of remorse, would in some cases adequately substitute for some of the state-imposed penance.

supported by aspects of his time delay argument and by the broader retributive values and conception of censure to which he adheres.

Quasi-retributive grounds for mitigation

In their 2005 book *Proportionate sentencing: Exploring the Principles*, von Hirsch and Ashworth engage with the possibility of there being quasi-retributive grounds that, despite not manifesting agreement with the proportionality requirement (strictly conceived) should nonetheless affect sentencing. Von Hirsch and Ashworth state that these grounds ‘relate to the wider underlying conceptions of penal censure as a response to criminal offending’ (2005: 174). Their discussion of these grounds arises from a discussion of ‘equity mitigation’ within Sweden’s sentencing provisions. In the Swedish statute, equity factors do not affect the offence’s penal value, but are nevertheless to be treated as reasons for reducing the severity of the sentence. Quasi-retributive grounds, they argue, are one possible source of justification for equity mitigation. Such justification arises from the nature and purpose of censure, which is at the heart of their desert theory. They argue that censure for an offence provides the offender with reason and opportunity to reflect on his wrongdoing and on why he committed it. What quasi-retributive grounds might do is ‘address special situations which relate to this reflective process’ (ibid.).

I wish to begin from their arguments for the sentence mitigation of an offender who is being sentenced a long time after the commission of his offence. It is from their arguments here that I think we can attempt to develop a quasi-retributive notion of blameworthiness (as distinct from culpability) and begin to consider how blame and censure feature generally in von Hirsch and Ashworth’s retributive theory.

The ‘time delay’ argument

The first statement von Hirsch and Ashworth make about time delay mitigation is the following: ‘when the offence for which [the offender] is being sentenced was committed several years earlier, the process of eliciting a reflective response may become more problematic after such a long delay’ (ibid.: 174). The second argument is this: ‘with the lapse of time, the possibility increases that the actor may have changed significantly – so that his long-past act does not reflect badly on the person he now is’. These two arguments are not the same. The former focuses on the offender’s capacity for reflection (regardless of whether the act reflects badly on the actor now). The latter focuses on what we might call blameworthiness. The actor now is not as blameworthy for his actions as he was when he committed them. The reason why this might be so, they suggest, is that he has changed significantly. More succinctly, whilst the former argument considers the function of censure, the latter considers its legitimacy.

Time, change and personal identity

Let us consider the second argument first – the suggestion that the offender has changed so much that his offence no longer reflects badly on the person he now is. What might be meant by this? The language von Hirsch and Ashworth use invite considerations of personal identity. The strongest version of such an argument would be that the person now being censured is not actually the same person as the person who committed the offence. If this were the case, then to punish him would be as illegitimate as punishing a random person picked from the street for this offence. The offence was committed by a person who no longer exists.

I doubt that von Hirsch and Ashworth have such a strong thesis in mind – they still consider censure and the associated sanction appropriate, but to a lesser degree

than if the offender had been addressed by the sentencing judge soon after commission of the offence. What they could propose, however, is that the offender sentenced after a long time delay is *somewhat* discontinuous with the person who committed the offence, without being an entirely new metaphysical entity.

In line with the thinking of philosophers such as Derek Parfit (1986), some punishment might then still be deserved because the offender is related (psychologically) strongly enough to the person who committed the offence, even if he is not 100 percent identical with him as a consequence of his changing significantly over time. Identity, on such an argument, would not be an ‘all-or-nothing’ concept, but would permit of degrees in a way relevant to punishment. Importantly, the offender being a ‘different person’ (or not fully identical), on this view is not simply that he is a different person in the sense that he seems to think about things and feel towards things in a new way. Rather, the person standing before the court does not share one-to-one the identity of the person who committed the offence. This person has not fully survived.

If von Hirsch and Ashworth were to propose such an argument for the sentence mitigation of offenders facing the justice system years after their offence then, I argue, the same reasoning they use could support the mitigation of the remorseful offender’s sentence. I assume here that by ‘changed significantly’ they mean that the offender might now be described as something like law-abiding, responsible, empathetic etc. – an offence would surely not cease to reflect badly on a person unless he had become somehow ‘morally better’. The argument then is that if these sorts of changes (which von Hirsch and Ashworth assume are possible with the passage of time) result in the offence no longer reflecting so badly on the offender, then why would an offender’s remorse not function the same way?

It is very likely that the sorts of ‘significant changes’ von Hirsch and Ashworth envisage are likely to correspond with the changes that remorse represents and facilitates. The change in attitudes that remorse involves are likely to constitute much of the change they are looking for. Unless they consider the passage of time to be necessary for significant changes or think that time should pass so we can be sure about these, then it seems to me that the remorseful offender may have changed significantly so that his offence no longer reflects quite so badly on the person he now is. Arguably, we can be *more* certain about the changes to a genuinely remorseful person than those that might have occurred to an offender who is simply distanced in time from his offence.

So, if their argument should, therefore, also apply to the remorseful offender, we need to see if it really does have purchase. In order to do so, we will need to explore in more detail the idea that remorse might lead to discontinuity of identity.

Discontinuity of identity

For Parfit (1986: 217), what is of importance when we think about the survival of persons is Relation R (and not identity *per se*). Relation R is determined by how *psychologically connected* one is to one’s purported past self, and the *psychological continuity* that endures (ibid.: 206). Psychological connectedness exists when psychological connections hold directly between a person at one time and at another time. The sorts of connections Parfit has in mind are between specific experiences and present memories of them, between intentions and their being acted on, between beliefs, desires and personality traits at one time and another. Psychological continuity, the other relation of importance, refers to the holding of overlapping

chains of *strong* connectedness. It is this relation, in addition to psychological connectedness, that allows persons to continue to exist over long periods of time.

How might this perspective on persons and the way in which they endure help to frame the argument that remorse makes the offender discontinuous with the person before the court? Reminiscent of von Hirsch and Ashworth, in the context of retributive punishment, Parfit proposes two scenarios which might lead to the psychological connections between the offender and the present person in court being insufficiently psychologically related to support desert-based punishment. He suggests that a) a person convicted many years after commission of the offence and b) a person who has undergone ‘some great discontinuity’, such as a religious conversion, would both be insufficiently connected to the respective past offenders to justify punishment (ibid.: 325). The beliefs, desires and personality traits of the remorseful person might be so discontinuous with those of the offender that they are inadequately connected and therefore punishing the remorseful person is not justified.

Before we consider this possibility further, let us think about an alternative example, discussed by Duff (1986: 15-22), where it could be argued that a lack of continuity renders punishment unjustified. Duff suggests that when an offender (who was sane at the time of the offence) later becomes insane, there might not be the continuity of identity needed to justify punishing him. This sort of great discontinuity, where the psychological functioning of the insane person is radically different from that of the person who committed the offence, does seem a plausible example of the sort of changes that could render punishment no longer justified. However, in addition to the discontinuity in identity, we might also note that the censuring function of punishment, central to the theories of Duff and von Hirsch and Ashworth would have little purchase on a person with severely impaired powers of comprehension. It may

be that this is a more straightforward way of justifying the sparing of the offender who becomes insane.

So, how might remorse affect 'relation R', and with what consequences? It could be argued that a powerful, genuine experience of remorse constitutes a 'great discontinuity' not entirely dissimilar from a religious conversion. Or, more specifically, the things about a religious conversion that we might suppose Parfit deems sufficient to establish great discontinuity are likely to be the sorts of things we would offer as evidence of discontinuity between the offender and the remorseful person facing sentencing: a new perspective on oneself and the world; new beliefs about wrongdoing; re-evaluation of one's behaviour; the modification of one's personality traits or at least a commitment to try to modify those which are less desirable, and so on. These changes, it might be thought, are sufficient to render the remorseful offender different enough from his previous offending self to constitute discontinuity between these persons. On these grounds, punishment, or at least full punishment, is no longer justified.

There are some initial comments I would like to make about Parfit's arguments before I evaluate the effect of remorse on identity in particular. The first is that Parfit's two scenarios – sentencing after a significant time delay and sentencing after undergoing 'some great discontinuity' – involve neutral or positive changes in the offender. In fact, Parfit's (1986: 326) own time delay example involves a drunken assailant becoming a (ninety-year-old) Nobel Peace Prize laureate. Likewise, a religious conversion is assumed to be a positive thing inasmuch as it demonstrates a new commitment to moral values, peaceful behaviour and so on. Could it be that it is the positive aspects of his examples of change which make us more inclined to agree with his conclusion that less (or no) punishment becomes deserved?

Imagine, for example, that an offender undergoes a great discontinuity, which involves dramatic changes to his beliefs, his desires and his personality traits. These changes, however, are very much for the worse. Twenty-five year old Bob is being sentenced for a not-very-serious common assault that occurred six months previously, and had been perceived at the time to be out of character. During those subsequent six months, however, Bob became acquainted with, and influenced by, a particularly malicious group of individuals. His whole attitude towards the world and other people has changed. He now often makes his family members cry and delights in the power he has to do this. He no longer does anything to help them, having developed the belief that people should be running around after him. He has left his job, and his earlier friends lament that it is as if the Bob they knew no longer exists, claiming that they want nothing to do with this 'new' Bob.

Such a significant change in Bob's character, values and behaviour suggests that he has undergone 'some great discontinuity' and if we return to Parfit's criteria for psychological connectedness, such as steady beliefs, desires and personality traits, he does indeed seem to be distantly related to the man he was when he committed the offence. Would such a case of great discontinuity render punishment unjustified? It does seem that he is not the same person who committed the offence but I suspect that our intuitions about the consequences this should have for punishment are not the same as those for the offender who became a nun or won a peace prize. In Bob's case, it would be the fact that he became a terribly nasty person that would lead to a reduction in punishment, which does not, I think, seem palatable. If he has somehow been brainwashed by this group, then there may be additional things to take into consideration, but the fact would remain that on Parfit's view, he is importantly discontinuous with the man who committed the offence. If we are to allow that

positive changes should have a bearing on personal identity and punishment, then negative changes that follow the same structure should too.

In a similar vein, we can imagine a religious conversion that constitutes great discontinuity, but that would likely inspire discomfort if it were to lead to reduced punishment. Imagine a burglar who undergoes a genuine religious conversion before sentencing. However, instead of the religious conversion leading the offender to dedicate himself to helping the poor and to practicing love and forgiveness, he instead becomes an extreme Christian fundamentalist who is intensely homophobic, Islamophobic and publicly celebrates the death of cancer sufferers as ‘God’s will’.⁵¹ That this religious conversion could result in a situation where there is insufficient psychological connectedness to support desert would strike many as abhorrent. Yet, such an example of discontinuity follows Parfit’s logic. I suggest, therefore, that it is the positive overtones of the changes he provides in his examples that make his ideas somewhat more persuasive than they might otherwise be.

John Kleinig (1984: 46) makes a similar point when he argues that unwillingness to punish a person for a much earlier offence is felt not because the person who committed the offence is no longer present, but more persuasively because a change in character is a morally relevant factor. This is similar to the argument that I will explore below: that the remorseful offender is less blameworthy on the grounds that he is now morally better.

Even if we were to limit the effect that psychological discontinuity could have on punishment to situations where the offender is somehow ‘better’ (or, at least, not ‘worse’) than he was, I am going to argue that remorse cannot be seen to effect *any* change in personal identity. With arguments similar to those of Radden (1995), who

⁵¹ For examples of such religious conversions, see Louis Theroux’s documentary on Westboro Baptist Church, *Louis Theroux: America’s Most Hated Family In Crisis*, BBC Two. Information available at: <http://www.bbc.co.uk/news/magazine-12919646>, accessed on 31st July, 2011.

has also explored the implications of remorse for personal identity, I suggest that the very nature of remorse presupposes an enduring self.

Remorse presupposes an enduring self

As discussed in Chapter Two, remorse felt over someone else's wrongdoing would be unjustified. If I, remorsefully standing before the judge, am not identical to the person who committed the offence, then why do I (or should I) experience remorse? As Radden (ibid.: 71) suggests, 'one's remorse and contrition are felt over one's own actions – not those of others and not over states of affairs that have befallen one'. Moreover, remorse possesses moral significance precisely because it is the same person. To suggest otherwise would be to deprive our moral attitudes of their 'essential ground' (Madell 1981: 125-6).

Joel Feinberg (1983) elucidates the intelligibility and moral significance of remorse when he suggests that references to offenders being 'different people', inspiring reluctance to punish, must be taken metaphorically. In reality, he argues, the same person persists in such cases; indeed, 'descriptions of the gentle sensitive person in Death Row presuppose for their intelligibility an *identity* with the savage beast who earlier committed a murder, and a continuity of development of the same self' (1983: 480, emphasis in original). It is only by the offender being the same person that his remorse makes sense and has moral significance.

When we think about what a denial of this would mean for remorse, we are left with a very undesirable picture, lacking the rich moral value that was discussed in chapters Two and Three. One possibility would be that the offender feels remorse rationally, but then at some point the remorse is strong and enduring enough that he becomes discontinuous with his offending self. The remorse then becomes illogical, if it continues, as he is now a new person. This would mean that *wrongdoers* never

really felt remorse (perhaps just the beginnings?) as it is someone different who experiences remorse for the things that they did. It would also mean that it is illogical for people to *anticipate* feeling remorse for their wrongdoing, as it will be someone else's remorse (see Perry 1978: 5-6).

Further, the question would arise whether, on *ceasing* to feel the (now) irrational remorse, the new person might somehow go back to being the previous person. Without the identity-changing state of mind and set of beliefs that the remorse consisted in, does he perhaps resume his old identity? Or, has he been irreversibly cut off from the offender, thus enduring as a new person? These scenarios would either make rational remorse very short-lived – on experiencing it one quickly becomes discontinuous with one's offending self – or would make it wax and wane: it tracks the offending identity which disappears with remorse but re-emerges when remorse fades as it becomes irrational. Neither picture fits the conception of remorse developed earlier in the thesis and neither looks morally desirable.

But, one might argue, on Parfit's view, remorse might not lead to a complete disjunction of identity, and just enough connectedness might remain that remorse is still rational, whilst there also being enough discontinuity to warrant reduced punishment. This would also have some peculiar consequences. If remorse leaves the offender only slightly connected to person being sentenced then this person being sentenced would rationally have to feel remorse to a lesser degree than if he was 100 percent identical with the offender: it was only 'sort of' him that committed the offence. But, if remorse (as Parfit might suggest) is transformative of identity then this new instance of moderated remorse would distance him from his (new) 'discontinuous self', who would then have to feel even less remorse, *ad infinitum*.

So, it has been argued that remorse requires identity to make sense and to be morally valuable. The valuable elements drawn out in chapters Two and Three are valuable precisely because it is the *offender* who regrets, understands what *he* did wrong, apologises, and so on. However, it should be pointed out that just because remorse has been seen not to transform identity, this does not provide a total refutation of Parfit-style reductionism. Whilst it does not make sense to understand the remorseful offender as discontinuous in ways that have consequences for desert, there might be other situations in which matters of personal survival can be illuminated by thinking about the extent to which someone is psychologically connected and continuous with his or her past self. It was suggested that the offender who becomes insane might be a convincing example of the offender not ‘surviving’, due to great psychological discontinuity. The justification for not punishing this person, however, is probably more conclusively found in the conditions necessary for censure, such as the offender being able to understand.

We cannot, therefore, use a personal-identity-based argument to explain why an offender’s having changed significantly through his remorse justifies a reduction in punishment. However, even if remorse does not alter identity, it could be argued that this (identical) person is somehow less *blameworthy* as a consequence of his remorse. It is to this argument that we now turn.

Time, change and blameworthiness

When considering von Hirsch and Ashworth’s time-delay argument – and, by extension, the remorseful offender – we can posit a weaker, and more plausible thesis to justify mitigation. The thesis is that, despite the offender being metaphysically the same person who committed the offence, he is no longer as blameworthy for it. Such

a possibility requires that we invoke a concept which applies in a way different to culpability, which was discussed above. An offender's culpability is set at the time of the offence: the offender either intended to commit the offence or did not, he either foresaw the likely consequences of his actions or did not – these facts cannot be changed retrospectively. The alternative concept of blameworthiness, in contrast, would allow temporal flexibility, allowing a person to become more or less blameworthy over time. The utilisation of such a concept would be necessary if von Hirsch and Ashworth want to argue that past acts can reflect more or less badly on a person over time – the inflexible concept of culpability is not enough to make viable their argument.

We need next to explore the proposed concept of blameworthiness. What might be meant by the assertion that an offence no longer reflects badly (or as badly as it might otherwise) on a person? I suggest that there are two possible arguments. The first possibility is that the offender is no longer responsible for his offence. The second is that the assessment of blameworthiness is made about the moral quality of the *person*, and so the morally improved person is less blameworthy, despite still being responsible for the offence. The two options, then, for an argument of the sort that von Hirsch and Ashworth make, are that blameworthiness is located either in responsibility or in character (constituted by attitudes etc.). I will attend to the possibility of remorse reducing the offender's responsibility for his offence first.

Blameworthiness as attributing responsibility to the offender

In relation to the remorseful offender, I will first consider the possibility that he is less blameworthy for his offence because he is no longer as responsible for it. Meir Dan-Cohen (2007) argues for such a possibility in his paper *Revising the Past: On the*

Metaphysics of Repentance, Forgiveness and Pardon. I will pay attention during my discussion of his argument to the potential differences between repentance and remorse, and to the particular understanding of repentance held by Dan-Cohen. I will argue that difficulties – resulting from his concepts of responsibility and repentance – result in Dan-Cohen’s argument failing to offer support for the possibility that the remorseful offender is no longer responsible for his offence. It will emerge that no version of this argument could prove successful.

Dan-Cohen’s argument attempts to characterise repentance as one of three ‘revisionary practices’ (the others being forgiveness and official state pardon) which redraw the boundaries of the offender so that his wrongdoing is left outside, eliminating his responsibility for it. The consequence, Dan-Cohen argues, is that negative reactive attitudes towards the offender for his wrong are no longer appropriate. Indeed, they become entirely misplaced. Thus, if an offender redraws his boundaries in this way, then he would cease to be blameworthy for his offence. In effect, he is changed, so that his wrongdoing does not reflect badly on his newly-bounded self.

Dan Cohen uses a political analogy to illustrate how a person’s boundaries can change so as to exclude the wrongful act. As a consequence, the person remains the same person (there is no change in identity in any strong sense) but ceases to be responsible for his wrongdoing. He gives an example to illustrate his analogy, as follows:

Imagine that the state of Arcadia has near its border a pollutant that causes environmental damage to the neighbouring states. As a matter of course, Arcadia bears responsibility for this pollutant: it is required to take measures to reduce the damage, to compensate the affected states, etc. It is equally obvious that this responsibility would expire if, say by treaty or by war, Arcadia’s border were redrawn so as to exclude the offensive site. [2007: 122]

The wrongdoer's repentance is meant to constitute a redrawing of his boundaries in an analogous way, leaving him no longer responsible for his wrongdoing and rendering negative reactions towards him for his wrongdoing unsupportable. Dan-Cohen explains further how this is possible in the case of a person, drawing on the notions of 'self-constitution' and 'social construction'. These ideas pertain to the proposition that human beings create themselves: individuals can be the authors of their own identities, and social practices – discursive and otherwise – also serve to shape them. Such symbolic interactionism takes place in the medium of meaning. Dan-Cohen (*ibid.*: 128) elaborates this: 'the self is the product of the web of meanings we spin around various objects and events, most importantly the human body and its career'. So, although we cannot change facts of the past, what can and does change is the significance we attach to past events and our attitudes toward them.

Dan-Cohen (*ibid.*: 130, n. 28) states that his account implies that 'the revisionary practices are inconsistent with a continued insistence on punishment, so that they must either occur subsequent to punishment or as a substitute for it.' Thus, the pre-sentence repentant offender eliminates the justification for his punishment, through his relinquished responsibility: he is no longer blameworthy for his wrongdoing and thus punishment would actually be unjust.

I am going to argue that there are significant difficulties with the political analogy that Dan-Cohen uses, making the conclusions he draws from it problematic. Following this, and for additional reasons, I am going to argue that the very notion that the remorseful offender can lessen his blameworthiness through shaking off responsibility is untenable. This will have the result of raising the question of what Dan-Cohen really means by repentance, and whether its practice would be entirely positive.

The problems with the political analogy arise from an understanding of responsibility that does not have application when we come to consider the wrongdoer. This can be seen when we consider a couple of features of Dan-Cohen's description of the state of Arcadia's altered responsibilities.

First, he mentions a couple of times the reallocation of responsibility which would occur in the political scenario. He suggests that 'the political scenario makes vivid the constructive and indispensable role played with regard to the reallocation of responsibility by the process or action by which the boundary is changed' (ibid.: 122). Later: 'the answer to the question what a boundary change is and to the question how it releases Arcadia from responsibility is one and the same: a border change ultimately just *is* reallocation of responsibility concerning the pollutant (ibid.:125, emphasis in original).'

Such a political scenario may well seem plausible but problems with drawing an analogy to personal boundary change become apparent when we think about the important notion of 'reallocation'. In the personal scenario there is no one (cf. nowhere) to whom we can reallocate the wrongdoing. The very idea that we might is ridiculous. Thus, the need for *someone* (a state) to be responsible for the pollutant in the political scenario is solved by the fact that the pollutant will now sit on the other side of the border, in a different state. In the personal scenario, however, if boundaries are redrawn in the way Dan-Cohen suggests, the result is that no one is responsible for the wrongdoing. No one can be held to account. Perhaps this disanalogy is not fatal to the argument. Indeed, we could conceive of a 'no man's land' into which the pollutant becomes situated. Perhaps then no state is responsible.

However, there is a more significant disanalogy between Arcadia and the repentant offender concerning the basis for responsibility. The pollutant just happens

to be within Arcadia's borders: it does not make sense to make any attributions of fault with respect to the pollutant. The wrongdoing of the offender, however, did not just happen to fall within his boundaries. To suggest that it did would be to deny his moral agency and his freedom to make choices, even if they are bad ones. In his case, it does make sense to make ascriptions of fault. Perhaps we might schematically draw a distinction between X generating responsibilities and being responsible for X. The former more accurately describes the Arcadian scenario: the pollutant generates responsibilities for Arcadia, but Arcadia is not responsible for the pollutant in any culpable way. The offender, on the other hand, *is* responsible for his wrongdoing in this sense. It may also generate responsibilities (such as making up for the harm caused) but responsibility in the personal scenario runs deeper than in the political scenario.

If we imagine a situation where Arcadia might be said to be responsible in the latter, deeper of the two senses, we can see that the political reallocation of responsibility scenario no longer has intuitive appeal. Imagine Arcadia hides landmines in the fields close to its border in an aggressive attempt to stop non-Arcadians from gaining access to its territory. Whilst the minefields are still active, a border change results in them being outside of Arcadia. In this situation, it seems less easy to accept that the neighbouring state now bears full responsibility for these mines. It would sound unacceptable for Arcadian authorities to say 'sorry, but your state is now responsible for these mines as they are within its border'. The difference here is that the mines are easily attributable to Arcadian action, whereas the pollutant was probably either naturally occurring or a result of complex environmental impacts and factors which are difficult to trace. Possibly, we would want to say that the minefields generate responsibilities for the authorities of the neighbouring state – they should

have acutely in mind the protection of their citizens – but we still would not want to say they are responsible *for* the minefields. Arcadia is still responsible *for* them.

However, Dan-Cohen is careful to spell out exactly the relationship he sees holding between for what an entity is responsible, and the entity's boundaries. He says that it is not the case that Arcadia is responsible for the pollutant because it is on Arcadia's territory:

Rather than it being the case that a state is responsible for X because X is within its boundaries, it is the other way round: X is said to be within the state's boundary and counts as part of the state insofar as and in the sense that that state bears responsibility for it. [ibid.: 125]

The pollutant is thus understood to be in Arcadia because Arcadia is responsible for it. If we work this way from responsibility to territory then the intuition would be that the minefields would always be within Arcadia's borders, politically, regardless of what happens geographically to any boundary line. In fact, it would seem that in cases of deep responsibility, it might never be possible to redraw the political boundaries. This then confirms the intuition in the personal case, that responsibility in this sense cannot just be reallocated. Further, if we were to accept Dan-Cohen's responsibility-boundary logic for persons nonetheless, we would be left with the result that anything could be within a person's boundary just by virtue of his taking responsibility for it.

Perhaps this does make sense for responsibility in the weaker sense, for example, if a person in a shared house takes responsibility for the vacuuming, this entails obligations and 'being the person who does the vacuuming' comes to constitute part of the person's identity (albeit a very small part). However, we would not want this to be the case for deeper responsibility. Dan-Cohen (ibid.: 132) says that 'reactive attitudes such as guilt and resentment play a decisive role in defining a person's responsibility and that by doing so they participate in the construction of a self'. Although this may be true in the sense that we are to some extent shaped by the

way we see ourselves, we would not want ascriptions of deep responsibility to depend entirely on whether we or others thought we were responsible. If we were committed to entirely subjective determination, pathological feelings of guilt and responsibility would generate obligations and justify punishment.

Dan-Cohen sees as a potential objection to his thesis the claim that revisions in the boundaries of the self, although conceivable, never occur. The evidence the objector could offer, he suggests, is that if any offender is asked point blank whether he had done the wrongful act, he would have to affirm that he had, despite having repented. This is true, yet Dan-Cohen (*ibid.*: 128-9) claims that ‘it is inappropriate for the truly repentant to dwell on the past misdeed and the hypothetical interrogator would be deemed nasty and cruel or obtuse for bringing up the nasty event.’

The norms to which Dan-Cohen alludes – those that prevent us from feeling comfortable bringing up the past event with the repentant offender – do in fact seem to operate. However, I would argue that what generates them is not the absence of the offender’s responsibility, but rather the extent to which it becomes right to blame the repentant offender. (This argument will become key in the development of my positive thesis in the following chapter.) If we reinterpret the norms in this light, the problems that arise for Dan-Cohen disappear and, I will argue, that a more realistic picture emerges. Further, the understanding of repentance that Dan-Cohen must be relying on is brought to the fore in these comments, and we will be able to contrast this with our conception of remorse.

Why is it inappropriate for the person to ask the offender about her responsibility for the incident if she has repented? Dan-Cohen argues that it is because multiple versions of the self can co-exist, depending on which revisionary practises have occurred and who is observing the individual. Thus: ‘if one insists, cruelly or

obtusely, on unearthing the older version and on resorting to it, one is not strictly speaking mistaken, merely cruel or obtuse' (ibid.: 130). However, the language Dan-Cohen uses here might be masking what really is inappropriate about asking the repented offender now about his responsibility then. Is referring to the responsibility cruel *per se*? I think it is not.

Imagine the mother of a young woman who mugged a passer-by gently starting to talk to her daughter about the event months after. Inquiring into why she had done it, the mother asks about her motivations and feelings at the time, perhaps asking 'darling, but how could *you* have done something like that?' In this scenario the mother is still very much invoking her daughter's responsibility for her offence but, out of love, tries to understand her daughter. Such a scenario does not strike me as cruel or obtuse. The daughter, even having repented, may actually welcome the chance to explore what happened, and why, with her understanding mother. So, reverting to the previous version of the person and inviting discussion about her is not cruel *per se*.

What I think makes Dan-Cohen's examples cruel, and what constitutes the difference between these and the curious mother, is the level of blame implied in and by the questioning. When the offender has repented, continuing to impart strong messages of blame to the offender does seem to be cruel. It is this intuition in particular that I will focus on and expand later.

Finally, Dan-Cohen's comments concerning the repentant offender's need to resist dwelling on the past highlight the unusual understanding of repentance on which his argument must rely. Dan-Cohen (ibid.: 129) suggests that 'a serious wrongdoing invariably casts a long shadow over the offender's life in the form of lasting negative attitudes and other consequences. When, due to the operation of the

revisionary practices, the shadow disappears, we ought to conclude that its source in the wrongdoer's life has been removed.' But as discussed in Chapter Two, lasting negative attitudes and lingering shadows are constitutive of remorse. We would be likely to find an offender's purported remorse suspect if it did not involve such attitudes. As Duff (1988: 164) says of repentance: 'it will prevent me from enjoying what I would otherwise enjoy. If I truly recognise and repent what I have done, I will not be able to enjoy, for instance, my usual social pleasures.' The shadow remains.

Moreover, remorse requires that one *accept* responsibility for what one has done. However, as discussed in Chapter Three, remorse certainly does not require that one condones one's behaviour. In fact, it was argued, part of the discomfort of remorse is the 'thin identification' with the deed: knowing that one did it (that one was responsible) yet not recognising it as part of whom one is. Duff sums up this idea (1988: 165): 'repentance is anyway not something which is done in a moment. It requires a proper understanding of what I have done, which both owns the wrong as mine and disowns it as something I condemn – a determination to improve myself, and to make such apology and restitution as I can'.

Duff's comment here, that repentance is not something that is done in a moment, is important. Dan-Cohen says nothing about what he actually means by repentance. It cannot mean something similar to what we are calling remorse, as he conceives repentance as assuaging an offender's guilt. Guilt here does mean guilt feelings, as he is referring to guilt as a reactive attitude. So, on Dan-Cohen's account: through repentance the offender removes his responsibility base by ceasing to understand himself as being responsible and, consequently, he ceases to feel guilty. It is hard to understand what repentance is on this reading, other than that it must be a process, since there is a before and after. But everything Dan-Cohen says of

repentance could be attributed to something like self-deception or neutralisation (see Matza 1990). If I convince myself that I was not responsible then I seem to have repented. Is this really something we would want to see as morally valuable? Remorse, in contrast, requires an ongoing painful recognition of responsibility.

It would seem, then, that the first possibility – that a remorseful offender might become less blameworthy due to changes in what he is responsible for – cannot be upheld. The deep sort of responsibility, as opposed to the weaker (X generating responsibility) sort, is difficult to shake off. It is true in a trivial sense that how we see ourselves is how we see ourselves, so that my ‘version’ of myself would cease to involve responsibility for X if I deceived myself into believing that I was not responsible for X. But Dan-Cohen’s analogies were shown to lend no support to this having any serious consequences for punishment. Further, I argued that a remorseful response of value necessarily involves continued acceptance of responsibility and is logically inconsistent with assuaging itself. I argued that continued insistence on bringing up the past wrongs of the remorseful offender can be inappropriate not because he is no longer responsible, but because it can be a mode of blaming which strikes us as excessive.

So, von Hirsch and Ashworth could not make use of the argument that the ‘significant changes’ in a person constitute a change in his boundaries: they could not argue that the offence does not reflect badly on the person he now is due to elimination of responsibility. This leads us to consider the second way in which we might want to claim that an offender is less blameworthy: as a person.

Blameworthiness as moral assessment of the offender

The second way we can interpret von Hirsch and Ashworth's argument that over time, the offender can change significantly, so that his offence no longer reflects badly on him, is that there is a sense in which the offence no longer shows the offender to be the bad person (or as bad a person as) he was when he committed the crime. The crime has lost its reflecting capacity because the 'picture' of the offence-committing offender and the picture of the changed offender are not the same. Reflection implies corresponding images.

The word 'badly' – reflecting badly – implies some assessment of the offender and thus lends itself to an interpretation of blameworthiness that has to do with the offender's attitudes and intentions. The leading desert theorists, however, only talk of blameworthiness as interchangeable with culpability, the concept from which we want it distinguished. How then might an offender become less blameworthy?

As noted, von Hirsch and Ashworth's argument suggests that there might be an important disparity between what the offender was like at the time of the offence, and what he is like after a significant passage of time. Having ruled out the possibility of the offender ceasing to be responsible for the offence, we instead must look to what he is like as a person: the attitudes he holds and so on. The argument might be that at the time of the crime, the offender was blameworthy for the complete lack of concern he demonstrated towards other human beings and the messages of derision that his action sent out. Over time, and (or) as a result of remorse, this lack of concern is replaced with compassion, and the derisory messages retracted. The offender might now be seen to be less blameworthy.

Such a characterisation of moral wrong was proposed by Jeffery Murphy in his writings on forgiveness. He explains that the main significance of repentance is that it negates an inference from the past wrong – that the wrongdoer still disparages

the victim. Murphy argues that forgiveness becomes appropriate since the repentant wrongdoer is not ‘*now* conveying the message that he holds me in contempt’ (1988b: 26).

However, to argue that such an understanding of reduced blameworthiness should justify mitigating the offender’s sentence does not seem quite right. Despite this ‘moral assessment’ notion of blameworthiness seeming a coherent idea in itself – we *can* judge a person’s present attitudes to be morally good or bad – it does not give us what we need to fully support the argument based on von Hirsch and Ashworth’s ideas. The sanction would focus too much on the offender’s present attitudes, which are not connected to the past crime. All that would be relevant is summed up by the following: ‘Is he holding anyone in contempt? No. In that case he is not presently blameworthy’.

This way of assessing blameworthiness results in the possibility that the offender’s blameworthiness could change daily or even hourly. One day he could be experiencing compassion towards all he encounters and the next he may be full of hate and malice. Present attitudes are far too capricious a phenomenon to constitute the means by which deserved punishment is determined. They may also be totally unrelated to the offence. In addition, there is still an uncomfortable intuition that a person who becomes less blameworthy should therefore feel less bad about what he or she did. Further, such an account of reduced blameworthiness does not explicitly acknowledge the moral value of remorse that was argued for in chapter three. What has become known as ‘sentencing on character’ is almost universally opposed by retributive theorists. To mitigate on the grounds of reduced blameworthiness – as a function of the offender’s present attitudes – would be sentencing on character in its most extreme form.

Conclusion

It has therefore been shown that neither conception of blameworthiness can support the mitigation of a remorseful offender's sentence. In the following chapter I am going to argue that instead of talking in terms of an offender's blameworthiness, we should talk about the extent to which an offender should be blamed. I will draw out this distinction and use it to make my positive argument for why a reduction in the extent to which we should blame an offender justifies mitigating his sentence. Von Hirsch and Ashworth's second argument for time-delay-based mitigation – concerning the offender's capacity to reflect on censure – will serve to support my argument.

Chapter Eight

Developing the positive thesis: remorse and censure in a dialogical context

Overview

The argument I will put forward to justify mitigating an offender's sentence based on his remorse will comprise five main stages. First, drawing on the work of Scanlon and Bennett, I will argue at some length that we can draw a distinction between a person's blameworthiness (as variously conceived in the previous chapter), and the extent to which we should blame or censure him. Second, I will note that the concept of censure is central to the principal communicative retributive theories. Third, I will argue that within these accounts, a dialogical model of censure is superior to a model in which there is no room for input from, and response to, the offender. The principal accounts will be shown to have reason to concede this model. Fourth, I will argue that communication of remorse is the most valuable, relevant input that the offender can make to the dialogue: his 'comment' on the wrongfulness of his conduct. Fifth, I will argue that the dialogical model requires that the censure visited on the offender needs to respond to his communication of remorse. Mitigation embodies this response. Von Hirsch and Ashworth's arguments relating to quasi-retributive grounds for mitigation will be shown to reinforce my arguments. Sixth, if the censure the offender deserves is mitigated, then, I demonstrate, the punishment that communicates this censure is correspondingly mitigated. Finally, I will defend my account against some potential objections. The consequence of this chapter will be to put the burden on potential objectors to show why one-way communication is the superior model for sentencing offenders, or why remorse is irrelevant even in a dialogical context.

The extent to which we should blame the offender

Scanlon (2008) explores the idea that the concept of blameworthiness can be distinguished from that of blame. For him, to claim that an agent is *blameworthy* for an action is to claim that this action indicates something about that agent's attitudes towards others that impairs his relations with them (ibid.: 6). This is similar to the second notion of blameworthiness explored in the previous chapter. To *blame* someone, on the other hand, is to go on to hold attitudes towards him that differ, in ways that reflect this impairment, from the attitudes required by the relationship one would otherwise have with this person (ibid.).

Bennett (2008), whose retributive theory of punishment draws extensively on the practice and symbols of blame and apology, also recognises a similar distinction, saying 'We can distinguish...the conditions that make a person *blameworthy* from the conditions that make it appropriate *overtly* to blame them, or to *express* our blame' (2008: 167).

We might, however, further distinguish all possible blame-related states of affairs: 1) a person can be blameworthy, in a sense that is comparable with the legal notion of culpability or, as Scanlon suggests, as a result of acting in such a way as to reveal relationship-impairing attitudes; 2) a person can assess another person as being blameworthy; 3) a person might go on to hold an attitude of blame towards a person in virtue of his blameworthiness; 4) a person might behave towards another person in a way informed by his attitude of blame; 5) a person might remonstrate with another person about his blameworthiness; 6) a person might remonstrate with another person about his blameworthiness from a position of relevant authority; 7) a person might

cease to do any or all of these things, perhaps resulting in/constituting the forgiving of that for which another person is blameworthy.⁵²

Conditions that justify one or some of these states of affairs do not necessarily justify them all. Further, a person's instantiation of one or some of these states of affairs certainly does not mean he or she is instantiating them all. For example, it might be justifiable for a person to hold an attitude of blame towards a stranger that he hears has cheated on her husband. It is debatable, however, whether he would be justified in behaving towards her (perhaps by glaring at her angrily and refusing to engage in conversation) in a way informed by his attitude of blame. It seems even less likely that he would be justified in remonstrating with her about her infidelity. That blameworthiness and all 'steps' of blaming are distinct is most conclusively shown by the fact that a person can cease to blame – perhaps to forgive – without the subject of the prior blame ceasing to be blameworthy.

It might seem that such a conceptualisation of the phenomenology of blame is only applicable to personal relationships, and thus cannot be transferred to a state-citizen model. However, this is not the case. Central to all relationships, explains Scanlon (2008), are intentions and expectations about how the parties will act towards one another. He elaborates how this can extend to moral relationships, holding between all human beings, and in what such a relationship might consist. He explains that, unlike personal friendship, 'in the case of morality, the relevant conditions do not concern the parties' existing attitudes toward one another but only certain general facts about them, namely that they are beings of a certain kind that are capable of understanding and responding to reasons' (ibid.: 139). And, for Scanlon, morality requires that human beings hold certain attitudes toward one another simply in virtue

⁵² Presenting these steps as a sequence is somewhat artificial. Not all of them will be passed through and none are necessary for the one that follows. However, there is a logic to the order I present them in, which I trust will make sense to the reader.

of the fact that they stand in the relation of ‘fellow rational beings’. Thus, when an individual does not manifest the mutual concern we should have for other human beings, the moral relationship is impaired, and blame becomes appropriate.

Bennett (2008) explicitly argues for the relevance of blame (and apology, to which I will return below) to state punishment. For him, private individuals who are engaged in intrinsically valuable joint enterprises (from friendships, to teaching, to simply being neighbours) must respond with blame to individuals’ failures to meet the basic standards expected of ‘qualified members’⁵³ of the enterprise. An extension of this is that the state must respond with condemnation to any qualified members of the political community who fail to meet the basic expectations of being a citizen. He uses a ‘right to be punished’ strategy to argue for this conclusion for both the sphere of private individuals and the state’s interactions with its citizens. The main idea is that failing to respond to those who act in ways that demonstrate disrespect for the legitimate norms and expectations of an intrinsically valuable relationship would fail to treat them as fully rational, qualified members who should be held to the demands of the relationship. Failing to condemn and temporarily withdraw goodwill is, Bennett argues, incompatible with maintaining the relationship and denies the right to be punished.

It should be noted that not all of the ‘steps’ of blaming can properly be transferred to the state. The state, for example, does not hold attitudes of blame towards its citizens (step 3) in a way that fully meets the phenomenology on an individual level. Whilst it can make assessments of blameworthiness (step 2) (based on the criminal law and sentencing legislation), it does not make sense to ascribe

⁵³ Bennett contrasts ‘qualified members’ with ‘apprentices’. Whereas a qualified practitioner can be expected to ‘recognise and comply with the responsibilities of her place in the practice autonomously, without external supervision’, apprentices (for example, within the collective project of political community, children) require help and direction to learn (2008: 95).

thoughts and feelings to the state. The mode of remonstrating with, or censuring, the offender can *only* be authoritative (step 6). Whereas, a teacher can censure a student formally from her position of authority, it is conceivable that she could censure the student for the same thing informally, person-to-person, in a way that would be qualitatively different. In sentencing offenders, the state always operates in a formal capacity. It could be argued that it would be inappropriate to apply the language of forgiveness (step 7) to the state – it is not the state’s place, the argument might go, to forgive the burglar, but the victim’s.⁵⁴

Bennett (2008), for one, would not agree with this. He argues that the reacceptance of the offender into the enterprise (of political community) ‘looks a bit as if he has been granted a sort of forgiveness: the offence is not quite forgotten, but put into the past: it no longer conditions the offender’s relations with the authority and his official status in the group’ (2008: 172). This is one of the main reasons that Bennett argues that the symbolism of apology as well as blame should be incorporated into the criminal sanction – that the state should impose proportional amends rather than simply hard treatment. The amends, he argues would represent what the offender would be motivated to make himself, were he to be truly sorry. This is necessary because ‘the offender’s reacceptance looks like the kind of forgiveness that results from a person’s having made a sincere and adequate apology’ (ibid.: 172). Consequently, in order to earn the institutional version of forgiveness, the offender has to be made to do something that is the institutional version of apology.

Thinking about the varying appropriateness of blame and the effects of apology invites the possibility of repairing the moral relationship (*à la* Scanlon) or facilitating one’s reacceptance into the political community (*à la* Bennett). When one

⁵⁴ However, if we were to adopt a political model in which all offences are conceived as being offences against the state, forgiveness (or, perhaps, official pardon) may sit more naturally as something the state can appropriately grant.

communicates one's genuine remorse, the offender's concern is shown to be reinstated and recommitment to the demands of the collective enterprise demonstrated.

This possibility of an offender going at least some way to repairing his moral relationship with the community is found in Duff's work. He suggests that

The crime created a moral breach between the criminal and her fellow citizens or her community... That breach can be repaired, thus reconciling her with those from whom her crime threatened to separate her, only if she herself is prepared to heal it by repenting and forswearing her crime, which necessarily also involves a resolve and attempt so to reform herself that she avoids such wrongdoing in the future... [Duff 1996: 48-9]

This is similar to Bennett's (2008) requirement of an offender performing the ritual of apology (through serving his amends-inspired sentence) in order for the offender to have his full status returned to him and cease to be condemned.

However, whereas Bennett envisages that the ritual of apology follows state condemnation (or perhaps is concurrent with state condemnation – the condemnation is meted out by requiring the offender to make proportionate amends), I am going to suggest that remorse can affect the appropriateness of blaming (or, at least, the severity that is appropriate) at *prior* steps. This is particularly the case, I will argue, when censure is approached in dialogical terms, which I will try to show the state has compelling reason to do.

Affecting the severity of the blame deserved

If it were to be the case that blameworthiness and deserved blame were distinct, it would be possible that two offenders could be equally blameworthy for their respective offences, but that they might not deserve blame of equal severity. I am going to argue that an offender's remorse functions to reduce the severity of the blame

that he deserves, and conclude that this translates into the requirement that the state censure him less harshly.

There are two possible bases for such an argument – convention and principle. An argument from convention would assert that cultural norms control when communication of blame is tempered and that these same norms should influence what we can allow to mitigate an offender’s sentence. Such an argument would draw on such norms as discussed in relation to Dan-Cohen’s (2007) work. If it seems cruel to insist on the same intensity of blame for the already remorseful wrongdoer – for example, by persisting to reiterate the wrongdoer’s responsibility for his wrongdoing – then this norm should inform the censure we visit upon offenders through punishment.⁵⁵

Manson (2011) proposes an argument from convention, pointing out that retributive theorists already concede that some aspects of sentencing depend on what has become conventional. Citing von Hirsch and Ashworth – ‘when judged in absolute rather than comparative terms, the censure expressed through penal deprivation is in part *a matter of convention*’ (quoted *ibid.*: 48) – Manson argues that amongst the justifications for mitigating factors is the justification of ‘legitimate sympathy’.⁵⁶ If the factor tends to evoke sympathy and is legitimated by serving sentencing aims (rehabilitation and reparation), he argues, then it should mitigate the offender’s sentence. Manson asserts that the ‘truly remorseful offender is worthy of sympathy’ (*ibid.*).

I will argue in the following chapter against the supposition that it is principally (or even at all) sympathy that remorse evokes. However, for now I want to

⁵⁵ I the following chapter, I compare in detail the account I develop with one put forward by Tudor (2008) that begins its justification of remorse-based mitigation with an appeal to these same social norms.

⁵⁶ The others justifications being: 1) drawn from a ‘principled approach’ and 2) ‘systemic rationale’.

argue that remorse can more forcefully be included as a mitigating factor justified by a principled approach. To do so, I will show that communicative retributive theories have reasons internal to their constitutive arguments that oblige them to accept remorse as a mitigating factor.

Censure as central to communicative retributive theories

The idea that the offender must be censured for his or her wrongdoing is central to the principal communicative retributive theories. For Duff, ‘punishment communicates the condemnation or censure that offenders deserve’ (2001: 27). Offenders have done wrong and therefore deserve to be confronted with disapprobation. Von Hirsch (1993) also considers it obvious that the penal sanction conveys censure or blame.⁵⁷ Drawing on the work of P.F. Strawson, he describes the capacity to respond to wrongdoing with reprobation or censure as ‘simply part of a morality that holds people accountable for their conduct’ (p.9). People judge wrongdoers adversely because their conduct was reprehensible. For von Hirsch, ‘censure consists of the expression of that judgement, plus its accompanying sentiment of disapproval’ (ibid.). He also specifies some of the ‘positive moral functions of blaming’, citing its capacity to address the victim, its capacity to address the offender, its provision of the opportunity for the offender to respond, and its ability to express a normative message to third parties, appealing to their sense of the conduct’s wrongfulness as a reason for desistance.

However, despite these ‘positive moral functions’, it should be emphasised that for both von Hirsch and Duff censure is *deserved*. This is what makes their theories retributive at their core. For censure to be deserved and not just ‘a good idea’ (i.e. useful to achieve something else), censure must be a ‘Good’ in itself. If censure were

⁵⁷ Von Hirsch and Ashworth explain that censure is the authoritative analogue of blame (2005: 19). This authoritative character is the only distinguishing feature that they mention.

simply a way of trying to manipulate how people behave, it would not be retributive, as it would be simply a way of achieving crime control. Instead, censure is delivered because it is the appropriate response to wrongdoing, even if it goes on to have other benefits, such as persuading the offender to repent. But, some positive consequences *are* part-and-parcel of what censure involves. Duff argues that it is internal to censure - part of what it is - that it seeks a response from the offender, persuading him to see his behaviour for the wrong that it is:

...if a citizen does commit such a wrong, the law should aim to bring him to recognise and to repent that wrongdoing: not just because that is a method of persuading him not to repeat it, but because that is *owed* both to him and his victim...To take wrongs seriously as wrongs involves responding to them with criticism and censure; and the aim internal to censure is that of persuading the wrongdoer to recognise and repent his wrongdoing. [Duff 2001: 81-2, emphasis added]

So, whilst censure may have positive consequences in the effects it can have on the offender's attitude towards his wrongdoing and future behaviour, it is delivered because it is owed. It takes wrongdoing seriously.

Importantly, both Duff and von Hirsch see the criminal *sanction* as part of the communicative enterprise. Punishment *communicates* censure; it does not simply follow it. Duff takes the strongest view, that a communicative conception of punishment provides its complete justification. That is, hard treatment is necessary because it can serve the communicative aims of punishment more adequately than can mere convictions or symbolic punishments. (The communicative aims being those of 'transparent persuasion' (Duff 2001: 101), whereby the offender is brought to recognise the wrongs he has done and that he should take steps to avoid such wrongs in the future, and the offender's communication of his remorse through penance.) So, for Duff, punishment is censure and penance, and neither will be adequately communicated without it.

Von Hirsch takes a weaker view, arguing that the harshness of the sanctions cannot be justified on communicative grounds alone. He does see punishment as serving a communicative function saying that, 'through the censure expressed by...sanctions, the law registers disapprobation and blame' (1993: 12). However, he does not see this function as sustaining the severity of the criminal sanction. He argues that the severity is justified by its capacity to provide an additional prudential disincentive which supplements the normative reasons conveyed by penal censure (p.13). He makes it clear that the two functions (and justifications) of punishment are not independent, rather, 'the censure and the hard treatment are intertwined in the way punishment is structured' (p.14). He also states that in his justification of punishment, the blaming function has primacy, and it is carried out by visiting criminal sanctions upon offenders as they can also play a supplementary role as a disincentive.

So, crucially for the argument I am going to make, blame or censure is conveyed by punishment. For von Hirsch, punishment also serves to deter, but it primarily communicates. If it were the case that censure was independent to punishment – that censure could be adequately achieved simply through conviction or the judge verbally imparting blame on the offender – then differences in the level of censure appropriate for a particular offender would not be intrinsically linked to the severity of the punishment he should receive. Since my argument will focus on the effect of remorse on deserved censure, it must be communicated by the sanction in order for there to be necessary consequences for the severity of this sanction. That this is the case - that censure severity is communicated by sanction severity – is, however, held to be true even for von Hirsch's 'two-pronged' justification of punishment. He explains that by altering the harshness of the onerous consequences of breaking the

law – consequences which both constitute hard treatment and express reprobation – the degree of censure conveyed will be altered (p.14).

Having established that censure is the central element of the communicative function of punishment, I am now going to argue that a dialogical model of this communicative enterprise is superior to one in which censure is simply an isolated message, inattentive to the content of the offender's prelude or response. Moreover, I will also argue that principles to which the two theories adhere make adoption of this model appropriate.

Communication as dialogue

As explored above, both the retributive theories developed by Duff and von Hirsch are communicative. Duff proposes a 'communicative conception of punishment as communicating to offenders the censure that their crimes deserve' (2001: 79). He emphasises that punishment should be communicative, rather than merely expressive, arguing that:

communication involves, as expression need not, a *reciprocal* and *rational* engagement...it aims to engage [the person with whom we are trying to communicate] as an active participant in the process who will receive and respond to the communication, and it appeals to the other's reason and understanding – the response it seeks is one that is mediated by the other's rational grasp of its content. Communication thus addresses the other as a rational agent, whereas expression need not. [2001: 79-80, emphasis in original]

Von Hirsch's conception of punishment is less explicitly dialogical, yet still seems to imply that a response from the offender is both expected and is an appropriate object of evaluation. In collaboration with Ashworth, he writes:

'...punishment, we believe, should be conceptualized as an expression of censure. Penal censure has important moral functions that are not reducible to crime prevention. A response to criminal wrongdoing that conveys disapprobation gives the individual the opportunity to respond in ways that are

typically those of an agent capable of moral deliberation: to recognise the wrongfulness of the action; to feel remorse; to express regret; to make efforts to desist in future – or else to try to give reasons why the conduct was not actually wrong. [2005: 92]

Following essentially the same remarks elsewhere, von Hirsch writes: ‘a reaction of indifference would, if the censure is justified, itself be grounds for criticizing him’ (von Hirsch 1993: 10). So, von Hirsch’s conception of the communicative function of punishment involves an expectation that the offender respond (in a way that ‘answers’ the censure delivered) and that his answer in turn is something that we can appropriately assess as being morally better or worse.

This emerging picture of an exchange; of a conversation that shapes the appropriate censure suggests a model of sentencing as a species of dialogue; as something that is two-way, in which both participants seek to better understand the subject matter. However, we need at this juncture to explore what really can be meant by ‘dialogue’ in this context, and why it is better than a non-interactive model.

The classical model of dialogue would be something like the Socratic Method: an intellectual attack and counter-attack designed to uncover truth or to discover new ways of thinking about things. This cannot be what is meant by dialogue here: a sentencing hearing is not a drawn out iterative exchange where truths slowly dawn on the offender. It is not a critical reasoning exercise.

When I suggest that censure should be conceived in dialogical terms, I am not imagining a literal conversation that occurs between the offender and the judge. Although the offender may physically say something to communicate his remorse, this will not be in the context of a drawn out exchange. Moreover, the sentencing dialogue I have in mind is, in some ways, *more* inclusive than the classical model. I want to incorporate more diffuse modes of communication. Censure through sanction is already extending the meaning of communication – it is not literal speech. In

addition, I suggest that the offender's remorse can be communicated in ways other than speech – perhaps through a letter, through a display of emotion, through significant life changes, through efforts at reparation, but also through speech. Moreover, this communication may occur before the offender physically meets the sentencing judge and will most likely not be directed at him (at least, not primarily). I want all these instances of communication to be potential elements of the censure dialogue.

One might ask why dialogue is the term I have used, if I am using it in such a non-standard way. I would reply that dialogue involves the essential elements that I want to use in my ideal model of censure: it is responsive – what is communicated by participant A is influenced by the prior communication of participant B, with a view to his subsequent response; it involves a shared topic – dialogue is incompatible with talking at cross purposes; it necessarily involves attention to one's interlocutor – one is not involved in dialogue if one ignores the other participant's input. Mere expression or even simple communication does not have these characteristics. One can communicate something without expecting, nor attending to a response.

So, whilst dialogue incorporates the characteristics I will argue should be transferred to the practice of censure, it is not to be thought of as a sequential spoken exchange that occurs in the court room. Instead, we can conceive of a more suitable notion of dialogue, which fits more comfortably with the sentencing context. Johannesen (1971: 374) suggests that 'dialogue seems to represent more of a communication attitude, principle, or orientation than a specific method, technique, or format. One may speak of a spirit of dialogue in the human communication process'.

This focus on the attitudes involved in dialogue makes it a more plausible candidate as a model of sentencing. Such an understanding of dialogue is found in the

works of Buber (1970, 1985), who suggested that dialogue essentially involves an effort to recognise the value of the other, and in so doing to see him as an end and not merely as a means to achieving a desired goal. This mirrors von Hirsch's (1993) requirement to respect the offender's moral agency and avoid mere 'beast control'. This alternative conception of dialogue is also implicit in Dewey's (1927) work on the democratic community. Here, dialogue is not an iterative process of truth discovery. Rather, an attitude of attention to others generates shared understanding, which can then inform collective action.

Dialogue as an attitude of communication is a very plausible model for sentencing. Understood this way, it does not matter that there is not a clear sequence of speech acts emerging from and alternating between the judge and the offender. The offender's input into the dialogue might have been first made when he expressed his remorse to the victim and tried to apologise. It need not even be entirely verbal: perhaps the offender's remorse is partly expressed, for example, through his seeking help from anger management programs. As I argued in Chapter Five, any self-imposed penance expresses something. Similarly Buber argues that the dialogue can be either spoken or silent, as long as the dialogic attitude is maintained in a communication situation (1970: 39-40, 1985: 3-4).

Core argument

My argument so far, then, is that by engaging dialogically with the offender on the subject of his offence, the quality of the censure will be improved. The offender is better able to understand the wrongful nature of his offence: that it is indeed something in light of which the remorse he feels is appropriate. A modified response to his remorse could communicate this. If censure seeks a response, and this response

is already forthcoming, to nonetheless continue to seek this response as if it were absent devalues the censure. If censure is not responsive then it universally assumes zero moral transformation and speaks to the remorseful offender as if he is yet to be convinced of the wrongfulness of his conduct. However, censure that communicates modified disapprobation, as a response to the offender's remorse, has greater moral force. This moral force stems from taking into account the remorseful offender's attitude with respect to his offence, addressing him appropriately to his understanding.

Recall Duff's insistence that communication should be reciprocal, appealing to the other's reason and understanding. Censure that ignored the offender's remorseful communication would fail to be reciprocal. It would also fail to appeal to the offender's reason and understanding because it does not address him in his present of understanding but deafly seeks the response it should have already heard. Von Hirsch also emphasises that a moral response is sought from the offender. If this ideal moral 'response' has already been communicated prior to censure, it might be thought that the censure now appropriate would be qualitatively different.

To illustrate this, let us imagine it were a *literal conversation*. The remorseless offender is sentenced. The judge says (very approximately): 'We have to explain to you that what you did was a really bad thing. You must see that. You hurt a lot of people and showed utter disregard for anyone but yourself.' This seems fair and appropriate. Imagine the next scenario. The remorseful offender comes to the judge: 'I'm afraid I've done something really bad. I feel utterly terrible and can't stop thinking about all the people I hurt. I hate myself for it and want to do something to make up for what I did'. The judge then addresses the offender using exactly the same words: 'We have to explain to you that what you did was a really bad thing. You must see that. You hurt a lot of people and showed utter disregard for anyone but yourself.'

This does not seem quite right. Directing the offender to see the wrongfulness of what he did seems inappropriate. It is as if the judge had not heard the offender. Instead, perhaps he might say, 'I'm glad that you can see that what you did was terrible, because it really was – you are right that you hurt a lot of people and trying to make amends would indeed be appropriate'.

In this response, the judge is still condemning the offence and bestowing blame on the offender, but in a way that is responsive to the offender's remorse. He addresses him as someone who already understands the wrongfulness of what he did, whilst also confirming that the offender has 'got this right'. This communication has greater moral force as it confronts the offender in his current level of understanding. This hypothetical exchange is over-literal but serves to illustrate my argument. I have developed a model of dialogue that incorporates non-verbal as well as verbal expressions of remorse that may be made before as well as during the sentencing hearing. Punishment as communication of censure extends the model of dialogue in a similar way.

If the change of response seems right and appropriate in the literal example, then it would seem that we should try to mirror it in censure communicated through punishment. Censure is still censure – anticipating a moral response – even if it is communicated through punishment. If we accept that the change in the quality of the censure is appropriate in the literal exchange with the remorseful offender above then, I argue, this has consequences for sentencing. Since censure communicated through a sanction seeks to confront the offender in an analogous way, we should accept that the dialogical model I have presented is superior to unresponsive, unidirectional penal communication.

Support from communicative retributive theories - rational agency

As well as making the argument independently for the superiority of a dialogical model of sentencing, I now wish to argue that other values which the principal retributive theories explicitly hold lend support to the acceptance of a dialogical model of censure. Von Hirsch and Ashworth's retributive theory emphasises throughout the need to address the offender as a rational moral agent. That is, someone who is 'capable of moral deliberation' (ibid.: 17). A communicative enterprise that ignores the offender's moral response (where present) would seem to be less respectful of rational moral agency than it could be. Of course, such respect for rational moral agency would preclude attempting to *elicit* a remorseful response – as von Hirsch and Ashworth argue against (ibid.: 18) – but should give the offender an *opportunity* to make such a response.

That communication is at its optimum when it engages with the offender, instead of proceeding indifferently, is also a theme that recurs in Duff's work. He too characterises communication as seeking a response and addressing the other as a rational agent:

Communication...is essentially a two-way rational activity. We communicate with another, who figures not simply as a passive recipient, but as a participant with us in this activity; our communication appeals to her understanding, not simply to her unreasoned feelings, and seeks a rational response from her. Communication, that is, addresses the other as a rational agent... [Duff 1996: 33]

Such engagement most naturally suggests a dialogical attitude: multiple participants engage in the activity of censuring the offender for her wrong, including the offender herself. Moreover, when seeking a response from a rational agent, one consequently puts oneself in a position to respond to the agent's response.

So far I have argued that a dialogical model of sentencing should be the preferred model within communicative retributive theories that focus on the censure an offender deserves. Engaging with the offender enhances the quality of the censure and best serves the aim of addressing the offender as a rational moral agent who is expected to respond. I have also shown how the principal retributive theories are not only compatible with a dialogical model, but seem to require it.

Some preliminary concerns

However, there are two general objections that could be raised to a dialogical model of censure: non-participation and inability to understand. It could be argued that dialogue requires two participants and, if the offender refuses to participate, a dialogical attitude is therefore inappropriate. However, I would argue that it is important to give the offender the opportunity to contribute if he so chooses and so a predisposition towards dialogue should always be maintained. If dialogue, as I have conceived it in the sentencing context is an attitude of attention, this does not become redundant if the offender offers no contribution.

A more difficult objection to answer is the potential argument that since some offenders may not be able to understand the meaning of the censure, or are unable to engage in dialogue (broadly construed) effectively, it is not fair to adopt a model that potentially favours (in terms of outcome) those that can. However, if, as I have argued, a dialogical attitude is generally superior to an unreceptive one, due to the superior quality of the communication achieved by the former, it seems right to commit ourselves to this model across-the-board. It would be incorrect to reduce the quality of the censure for all offenders – to cease to engage with them as rational

agents – on the basis that some may not be able to grasp entirely the content of the communication.

Remorse as the offender's ideal, relevant contribution to the dialogue

Of course, not everything the offender might communicate is relevant. The communication is about the offence (not, for example, whether the offender has done good work for charity in the past), and the offender's response to the offence can form part of this communication.

If we conceive of the sentencing hearing as dialogical, we can consider what the constituents of the dialogue would be. The subject matter is the offence for which the offender is being sentenced. The judge communicates disapprobation to the offender – censure that is embodied in the sanction that is imposed and thus continues to be communicated. Both Duff and von Hirsch expect that the offender will respond to this censure, but he might respond in many ways. He might, for instance, respond with indifference, he might respond defiantly, he might respond remorsefully, or he might offer a complete non sequitur. Importantly, any of these 'responses' might be offered by the offender prior to, or during, sentencing – as 'responses' to the anticipated censure or simply as reactions to the wrongdoing itself.

Some responses may not be relevant to the dialogue. As noted, Duff (2001: 79-80) explains that the response sought is one that 'is mediated by the other's rational grasp of its content'. Perhaps an offender might respond with a list of good things he had done in the past. However, this would show that he had not really grasped the meaning of the censure. He is being censured for his offence – his offence is the subject matter of the dialogue. To ignore the subject of the wrongfulness of his actions and instead try to draw attention to something non-offence-related would not

be an appropriate input. The appropriate response of the interlocutor would then be to reaffirm the censure.

Perhaps another offender might respond indifferently to the message conveyed in censure. In such a situation, it is possible that the offender had fully grasped the content of the censure but simply did not care. Hypothetically: 'I understand that it was wrong but that does not bother me'. This response *is* focused on the subject matter at hand. It reveals the offender's attitude both to his wrongdoing and to the censure he is receiving. However, as von Hirsch suggests (1993: 10), such a response would be grounds for further criticism. A remorseful response, on the other hand, is both relevant – it focuses on the offence – and desirable: it is the appropriate moral response following the commission of the offence. It demonstrates understanding and moral concern.

Further support from quasi-retributive grounds and von Hirsch and Ashworth's time-delay argument

At this juncture, we can now consider the second argument von Hirsch and Ashworth advance to explain why the person who is sentenced a long time after the offence should receive mitigation. It serves to emphasise their focus on the reflection-facilitating function of censure, and the occasions for mitigation they think this can create. Conceiving of censure as reflection-facilitating shows a commitment to censure being not merely a 'statement', but as a communication that is sensitive to the responses it can provide opportunity for.

Recall that von Hirsch and Ashworth express commitment to quasi-retributive grounds that 'may address special situations which relate to this reflective process' (2005: 174). Although the censure communicated by the sentence is not tailored to

elicit remorse, they emphasise that ‘the rationale underlying the censuring response has something to do with the capacity of an offender, as a moral agent, to engage in certain kinds of moral reflection’ (ibid.: 175).

Their time-delay argument – considered in the previous chapter – shows that they accept that there are ways in which the offender can change, particularly in reflective capacity, that are relevant to the censure he receives. I will argue that if this holds for the time-delayed offender, then a parallel argument can be made for the remorseful offender.

Von Hirsch and Ashworth (2005) say that ‘the process of eliciting a reflective response may become problematic after such a delay’ (ibid.: 174). The assumption made here is that an offender may no longer easily be able to think very deeply or feel very strongly about an event that was long ago. I am not sure this is an obvious assumption: sometimes the passage of time can give us a clearer perspective on things. However, what it does show is that they deem the function of eliciting a reflective response important for censure, so much so that incapacity of the function might result in the censure being mitigated.

It is inherent to remorse that it constitutes a reflective response and one that is deemed the optimum: ‘when the offender is thus censured, a moral response on his part would be deemed appropriate – for example, an expression of concern, an acknowledgement of wrongdoing, or an effort at better self restraint’ (ibid.: 18). It would seem peculiar if the remorseful offender – the offender *most* engaged in the reflective process – did not receive mitigation whereas the offender who evaded detection for a long time did receive mitigation, on the grounds that he may not be able to reflect effectively. Indeed, it could be argued that an *increase* of censure might

be necessary in order to jog his memory and reiterate the reality and nature of his past conduct.

Von Hirsch and Ashworth (2005: 177) also recommend mitigation for the offender who has offered reparation voluntarily, for reasons that are ‘related to the values underlying the desert rationale’. They identify this value as giving the defendant the opportunity to provide recognition of how he has wrongfully injured the victim. Again, they relate this to the reflective process that censure gives the offender reason to engage in. They argue that when an offender offers reparation voluntarily he not only has reflected on and recognised the wrongfulness of his actions himself, but has taken action to convey that recognition to the person wronged, in a manner that is closely linked with the wrongful conduct itself.

The remorseful offender, as we have conceived him also has reflected on and recognised the wrongfulness of his actions, and should be disposed to make attempts to apologise and repair. However, von Hirsch and Ashworth argue that the case for mitigation may become ‘considerably weaker where the offender merely feels or expresses remorse but takes no such action’ (ibid.: note h). I will argue that their position over emphasises reparation and underemphasises remorse.

It is in remorse and not reparation that reflection on actions and recognition of their wrongfulness occurs. It may be that voluntary reparation makes us more *convinced* of an offender’s remorse, but it is remorse itself that is bound up with the reflective process censure is intended to facilitate. Indeed, voluntary reparation could occur without remorse – perhaps the offender simply wants the victim to ‘get off his back’. However, it was also argued in Chapter Two that remorse that did not motivate any attempts to apologise and repair the harm caused would to some degree be suspect.

It may be the case, then, that voluntary reparation makes us more certain that an offender's remorse is genuine, but we must be clear that it is the remorse that is tied to the reflective process, and not the reparation *per se*. Reparation may have other instrumental benefits for the victim, but is not intimately linked to the quasi-retributive values relating to the censuring response. So, I argue, it is remorse that should justify the mitigation, even if voluntary reparation is a fairly robust marker.

In sum: if, as von Hirsch and Ashworth argue, the offender's reflective process is important to the rationale underlying the censuring response, then an offender who independently and readily engages with it should be censured in this context. Adopting a dialogical model of censure achieves this, making room for input from the offender on the subject of his wrongdoing. This of course does not mean that censure is not appropriate in such cases where self-blame is already occurring. The censure needs to be reiterated and confirmed, but the offender's mind does not need to be drawn to it and the quality of the censure required in response is altered.

Mitigation of punishment is necessitated by the change in the meaning of the censure

As explained above, communicative retributive theories – including the two being discussed – hold that the sanction, or hard treatment, does the work of communicating the censure. Therefore, if the censure deserved is less harsh, then the punishment required to communicate this censure is mitigated. Modulation of the tone of delivery of the sentence by the judge would not be sufficient as the *sanction* communicates the censure and the meaning of this censure must change in the face of the offender's remorseful input into the dialogue. Remorse therefore mitigates the punishment deserved.

It might be argued that allowing such mitigation on the grounds of remorse might send out the message to society that the conduct was not as serious as the unmitigated sentence would have suggested. However, in line with treating offenders foremost as moral agents and not as means to achieve deterrence, von Hirsch and Ashworth say that the offender ‘is being confronted with disapproval in virtue of the wrongfulness of his conduct, and *not solely in order to produce preventative or other societal benefits that such censure might achieve*’ (2005: 17, emphasis added).

Further, since judges should explain when passing sentence how they have reached their particular decision, all mitigating (and aggravating) factors should be explained. In theory, the public can then be made aware that the conduct was not less serious but that the reduction in censure constitutes a response to the offender’s communication of remorse. This also sends the message that remorse is the appropriate response to moral wrongdoing. To argue that the public usually only hear about the sanction and so would not understand how the sentence had been reached is true, but to suggest that this precludes responding dialogically to the offender’s remorse would also lead to the conclusion that all mitigating and aggravating factors functioning outside offence seriousness be eliminated.

An argument against the plausibility of thinking of punishment in dialogic terms

In a recent paper, Brownlee (2011) raises some serious objections to dialogical conceptions of punishment. Since these objections could be levelled against my account, I must show the reader why it nonetheless remains plausible. I will outline her concerns. These relate to conceiving of censure and sanction in dialogical terms generally, and also to applying such a dialogical conception to remorseful offenders’ interactions with the state in particular. I will then argue that, whilst her concerns raise

important points about the appropriateness of dialogue as the correct model for criminal punishment, they do not apply with equal force to *all* communicative theories. In consequence, the dialogical model I have argued for remains an appealing modification of von Hirsch and Ashworth's theory. Further, I will show that within her arguments there is a very important difference between assessing whether current practice *does* satisfy the conditions for dialogue and asking whether practice *could* satisfy these conditions. Drawing the conclusion that dialogue is not the appropriate model because what we have does not currently match it is a mistake; one to which Brownlee seems susceptible. The account I have developed shows both how sentencing could become more dialogical, particularly with respect to remorse, and that this would be a good thing.

Brownlee uses Duff's communicative theory of punishment to test the appropriateness of applying the language and logic of dialogue to criminal conviction and punishment. According to Duff, 'the [criminal] trial seeks to engage the defendant in a rational dialogue about the justice of the charge which she faces, and to persuade her – if that charge is proved against her – to accept and make her own the condemnation which her conviction expresses' (Duff 1986: 233). Brownlee emphasises that in Duff's view, punishment not only communicates both condemnation and a desire for repentance and reformation by the offender, but also gives the offender an opportunity to communicate her repentance by accepting the punishment, apologising, and making reparation where possible.

Conditions for dialogue

In order to assess whether accounts of punishment as a communicative practice can be conceived of as analogous to interpersonal moral dialogue, Brownlee outlines five

conditions for dialogue. Broadly, ‘a *dialogue* is a sustained, purposive conversation or verbal exchange of thought carried out by two or more persons. A *moral* dialogue is such an exchange that has as its subject either a moral issue or an issue that has moral implications. Often a moral dialogue will involve or address some moral disagreement’ (2011: 57, emphases in original).

The first condition Brownlee identifies is *reciprocity between the parties*. She suggests that, for a dialogue to occur, each party must be an active participant in the interaction in that each must intentionally play the dual roles of communicator and receiver. Additionally, ‘a reciprocal exchange is marked by mutual recognition of the conventional communicative claim-rights, duties, and privileges that each has as an active participant in this interaction’ (ibid.: 58).

The second condition is *sustained and extensive interaction*: a dialogue is a more sustained and extensive interaction than other kinds of broadly reciprocal exchanges such as a simple call and response or an exchange of threats, or a wordless meeting of minds (ibid.).

The third condition requires *purposive conversation marked by mutual recognition of each party’s rights and duties*. Brownlee describes dialogue as ‘reason-giving, argument-based, progress-oriented interaction’ (ibid.). Crucially, the parties must attend to each other’s contributions and modify their responses in light of those contributions; their exchange is neither a quarrel nor a one-sided monologue. Its purposes are constructive, didactic, and rationally persuasive. This does not mean that a dialogue necessarily leads to a reasoned resolution or an agreement amongst the parties. Rather, their engagement in this kind of interaction implies a certain mutual orientation toward progress in common understanding.

The fourth condition is *fairness and equality*: Implicitly, the parties are represented as equals in the relevant sense of being equally active and equally empowered participants to the exchange; this implied equality arises not only from the first condition of reciprocity in the parties' reciprocal recognition of each other's communicative rights and duties, but also from the third condition of the reason-giving, argument-based nature of the exchange where each party is addressed as a reasoning agent. Being equally active does not mean that each party must have and make use of equal space in the communications. Rather, it means that each has an equal right not only to speak when he wishes (provided that he respects the equal rights of the other), but also to be heard and to be understood (ibid.).

Finally, dialogue requires *willing participants*. In order that what each party communicates may be credibly taken to be what each party wishes to communicate, the parties to a dialogue are necessarily broadly willing (if not enthusiastic) participants to the exchange in the sense that they are not subject to duress or manipulation. 'Such non-subjugation is necessary for their exchange to be genuinely and truly reciprocal, and for them to be and to be seen to be equal in the relevant sense' (ibid.).

Brownlee believes the fourth and fifth conditions to be the most problematic when attempting to construe trial and punishment in dialogical terms. These are presented as problems whether the offender is remorseful or not, applying to dialogical accounts in general. I shall attend to these first, before proceeding to Brownlee's remorse-specific concerns.

Unwillingness poses a problem for a dialogical model of punishment

Brownlee states that the parties to a dialogue must be broadly willing participants to the exchange in the sense that they are not subject to duress or manipulation. Such non-subjugation is necessary for their exchange to be genuinely and truly reciprocal, and for them to be and to be seen to be equal in the relevant sense. Unwillingness could occur on either side of the would-be exchange.

In terms of the offender, Brownlee argues that ‘willingness...cannot be presumed to apply to the offender confronting punishment’, thus precluding dialogue (ibid.: 58-9). However, I do not think this is really a problem for the dialogical model. Failure, on occasion, to achieve an aim does not invalidate the aim. I agree that if the offender refuses to engage then no dialogue occurs, but the state should not then fail to engage if an offender does invite such dialogical interaction. The state should be ideally hoping for this dialogical scenario, even if it does not always – or, indeed, often – materialise. As I have argued, this makes the state’s communicative aims and mode of imparting censure as legitimate as possible. It is the state’s duty to be ready to heed and to interpret correctly the offender’s would-be communications so that dialogue can occur if it turns out the offender is willing to engage.

Indeed, Duff himself makes a similar point:

The moral possibility of trials and punishments does not, of course, depend on their actual success in bringing wrongdoers to engage in the communicative enterprise, or to answer for, to repent, or to make amends for their crimes: we must address the wrongdoer as someone who *could* respond appropriately, else there is no sense in seeking a response from him; but the value and importance of the attempt to engage him in a penal dialogue does not depend on its actual or likely success. [Duff 2009: 91, emphasis in original]

In fact, I suggest that it is actually a good thing for a liberal theory of punishment that offenders can choose not to engage in the dialogue. It respects their autonomy. But,

again, this does not mean that the state should not be ready and willing to engage when the offender does make an attempt to communicate.

Thus, Brownlee's concern that the unwillingness of the offender to engage poses a problem for a dialogical model is allayed. Dialogue should be aimed for as the ideal, and this is unaffected by the fact that sometimes – maybe most of the time – it is not attained. The dialogical attitude of the court should be maintained so that if the offender does communicate his remorse, the censure is not deaf to it. Brownlee is right that this means that sometimes a particular trial and punishment fails to be dialogical, but not that this poses a problem for dialogical models generally. In fact, it was argued, room for non-compliance may be a strength.

Condemnation is overtly performative, altering the offender's status such that she is prevented from engaging in dialogue

The second of the two general concerns relates to the fourth of the five conditions set out by Brownlee (2011): fairness and equality. The crux of her argument is that the state's act of condemning an offender disrupts the conditions of rough communicative equality and reciprocity necessary for genuine moral dialogue. Brownlee maintains that condemnation in legal (and religious) contexts operates differently from the way in which it does in ordinary interpersonal dialogue.

The conditions for ordinary interpersonal dialogue do not rule out the communication of condemnation by one person to another because, in interpersonal interactions, condemnation is relatively innocuous. Brownlee (2011: 62) argues that an ordinary person's condemnation of another does not alter the latter's status as a party to any dialogue between them unless that condemnation signals a termination of relations rather than an invitation to discuss the charge. When the condemnation does

signal a termination of relations, then the condemnation is a *performative act* because all subsequent would-be communications by the condemned party become, in the first instance, mere acts of expression (ibid.). When, by contrast, condemnation signals an invitation to discuss the charge, the condemned party retains her ability to respond to the charges on a footing of rough communicative equality and reciprocity because her interlocutor has not exercised his limited power to alter her moral standing in relation to him (ibid.).

Whilst there seem to be two options for the ordinary condemner – to invite discussion or to terminate relations – Brownlee argues that only the latter, performative mode is available to the state in the process of criminal justice. She seems to think that this is particularly true on Duff's account, where the quasi-religious language of condemnation is employed, with its connotations of being doomed to punishment and even damned to hell. With respect to the institution of criminal justice, Brownlee suggests that although the notion of *condemnation* is most apt in capital punishment cases, where the sentence really does doom the convicted person, condemnation and hard treatment even in much milder contexts 'radically demotes the offender's legal standing, social standing, rights, and duties, in a way that undermines the present possibility for dialogue about the offending conduct' (ibid.: 63).

However, it is the focus on Duff's communicative theory that gives this objection such force. I shall argue that the quasi-religious language is peculiar to Duff and that other communicative theorists, such as von Hirsch and Ashworth would be more inclined to substitute the word condemnation for censure. The connotations of 'censure', as Brownlee notes, are not as heavily loaded as those of condemnation. Further, I shall argue that even if censure does have an effect on the ability of the

offender to engage in dialogue, this does not threaten my account, as I envisage the offender's remorseful input occurring pre-censure. The censure then *responds to* this input.

Duff's communicative theory is infused with quasi-religious language and metaphor: 'penance'; 'atonement'; 'condemnation'. His use of examples involving monks in religious communities demonstrates that the borrowing of this language is intentional and meaningful. He wishes to draw instructive comparisons between religious penance and secular penance. This makes a particularly strong reading of 'condemnation' understandable, with its connotations of negating a member's standing within his religious community. Such a reading informs Brownlee's perspective that condemnation has the effect of cutting a person off from all interaction, including dialogical interaction. However, even within Duff's penological theory, we must be careful not to take this to an unintended extreme. Central to Duff's theory are the three 'R's of punishment, one of which is reconciliation with the offender's community. Duff argues that reconciliation is what those who the offender has wronged 'must seek if they are still to see her as a fellow citizen' (2001: 109). A complete exiling of the offender is irreconcilable with this objective.

Bennett (2008) is another penal philosopher who endorses condemnation expressed through ostracising treatment of the wrongdoer. In his non-criminal examples he envisages a man appropriately condemning his carousing, noisy neighbour by shunning him on the street and 'cutting him dead' (ibid.: 104). Similarly, he endorses a university communicating condemnation to a extremely slipshod teacher by temporarily suspending her status as well as requiring her to give her own time to help the students she let down. This, however, is in tension with Bennett's theory overall. Whilst arguing for the duty to withdraw recognition and respect from

the offender he also seeks to retain as much from restorative justice as is possible within an overall retributive framework. For Bennett, the ideal censure and sanction scenario is that the victim and offender agree to mediation where, although the severity of the sanction will be determined by a desert-based consideration of proportionality, the precise form this will take will be a product of the conference. Such a scenario seems to maintain a level of respect for and recognition of the offender, rather than subjecting him to exile.

It does seem somewhat of a paradox in Bennett's work that the way to respect and recognise the qualified status of the member of the enterprise is to withdraw respect and recognition.⁵⁸ He argues that it is because all qualified members deserve respect as such that they must be adequately condemned for what they do. For the focus of this condemnation to be withdrawal of respect seems strange. When he begins to outline how state punishment should work on his theory, however, the emphasis on withdrawal of respect and recognition diminishes in favour of putting as much restorative practice in as possible. Even before this stage, ostracism does not always seem to be his preferred mode of condemnation: 'we respect someone's identity by engaging them in dialogue, not by making them suffer' (2008: 74).

Censure, in comparison with condemnation, more decisively seeks a response from the censured, which is incompatible with cutting him off. Whilst censure is not a technique for *evoking* specified sentiments, 'some kind of moral response is expected' (von Hirsch 1993: 10).

This being said, the dialogical model I have developed envisages offender's communicating their remorse *pre-censure*. This would render any change in the offender's status post-censure unproblematic for my account. Brownlee does

⁵⁸ There are, however, degrees of respect and recognition. The respect lost could pertain to just this particular relationship or to a particular element of it.

acknowledge that the specific nature of Duff's theory fuels the worry that, being condemned, the offender suffers a demotion in status, which affects his present capacity to enter into a dialogue with society about his conduct: 'in his view, the aims and purposes of the trial are continuous with those of punishment, and hence the process as a whole is intended to be viewed in the reciprocal terms of moral dialogue' (Brownlee 2011: 65).

For Duff, the offender's input into the dialogue is his apology, expressed through his secular penance. This input *is* post censure. However, compatible with the communicative theory of von Hirsch and Ashworth is the suggestion that the remorseful offender's input occurs at an earlier stage. This being so, the offender's standing is actually *maintained* through responding to him with moderated censure. Through engaging in a dialogical approach to censure, the state demonstrates respect. The offender's undergoing the sanction (dictated by the particular degree of censure) is then to be conceived of in more one-sided, communicative terms. The particular degree of censure, however, is influenced by the offender's input into the dialogue concerning his wrongdoing.

Despite her overall conclusion that 'it is implausible to think of lawful punishment in dialogic terms', she does accept that 'such a forum, which precedes condemnation, does seem broadly acceptable as an analogue of interpersonal dialogue, if we put aside cases of unwillingness to plead' (ibid.: 66). It was previously argued, however, that the refusal of some offender to engage in dialogue does not undermine its value as the correct aim.

The script is already written so there is no chance for genuine expression of remorse (it appears the same as those forced to offer ritualistic apology)

Brownlee (2011) presents a remorse-related objection, which she believes further undermines the plausibility of thinking of punishment in dialogic terms. The objection is a specific instance of the more general ‘generic-script’ problem. The generic-script problem is generated by the ritualistic nature of the state-offender ‘dialogue’ which is dismissive of the offender’s actual attitudes and would-be communicative efforts. The state still tells the offender how he should communicate to society and victims in the way of penance and reparation, and then makes him go through the motions of taking those communicative steps.

This problem is marked in the case of a remorseful offender. Brownlee argues that:

in particular cases – such as full repentance prior to punishment or suitably constrained conscientious disobedience – the convicted person should not wish to recite the formal script of apology and commitment to self-reformation; and to require her to do so through the ritual of punishment and apology not only disrespects her, but makes a mockery of the ideal of genuine moral dialogue undertaken voluntarily and reciprocally on a plane of roughly equal footing by parties who are responsive to the communicative contributions of others. [ibid.: 60]

There are two responses to this. First, it should be pointed out that this is Duff’s script and as such is not essential to communicative theories generally. Whilst, on Duff’s theory, the offender (whether he wants to or not) ‘communicates’ repentance and apology to his victim and community through his penance, the sanction in von Hirsch and Ashworth’s theory is not conceived as communicating anything *from* the offender. Thus there is no generic script, and the offender can input whatever he likes, or remain silent.

Further, even within Duff's theory, there can be departure from the generic script. Brownlee highlights this in relation to conscientious disobedience. Duff demonstrates a departure from the generic script in the case of an offender who believes she did the right thing in going through with a mercy-killing. Duff explains that her punishment must embody the more complex message that even if, for respect-worthy reasons, she dissents from the content of the law, she ought to obey it out of respect for the law and as a matter of her duty as a citizen. He concludes that her punishment 'will probably be lighter than that imposed on someone whose crime did not flow from respect worthy values' (Duff 2001: 122).

This shows that even within Duff's theory, departure from his script is possible and stems from dialogical concerns. Whilst Brownlee raises concerns over Duff's solution to the conscientious offender, these concerns actually support mitigation on the grounds of remorse. She questions whether the message communicated to the conscientious offender should be censure at all, albeit reduced censure.

Further, she argues that the punishment of a conscientious offender might misrepresent as *remorseful* a person who wishes to engage in dialogue of about the merits of her cause (2011: 61). This objection does not apply to departure that responds to a remorseful offender with mitigation – indeed, the implication of Brownlee's argument is that this is the correct response to remorse and that doing likewise for other types of offenders might misrepresent them as remorseful. Although this is plausible, and shows how genuine dialogue can be established, it should be remembered that Duff himself does not advocate departure on the grounds of remorse.

So, it would seem that even within Duff's theory, remorse does not pose a problem for satisfying the conditions of dialogue. Genuine remorse, if communicated prior to sentencing, can be attended to and responded to uniquely through censure, thus paying attention to the particular offender's attitudes and communicative attempts.

Condemnation insensitive to the efforts of repentant offenders to engage in genuinely reciprocal interaction fails to respect the moral significance of their attitudes

However, Brownlee (2011) concludes that condemnation *is* insensitive to offenders' remorse, and thus dialogue is not achieved. In her conclusion, she says that, for wholly repentant offenders, 'condemnation that is insensitive to their efforts to engage in genuinely reciprocal interaction fails to respect the moral...significance of their attitudes' (ibid.: 66). This objection, in addition to the performative quality of condemnation (discussed earlier), are seen to speak to the general concern that it may be impossible in criminal justice processes for both society and offenders to engage in dialogue.

However, I do not think the fact that condemnation *can* be insensitive to the would-be communications of remorseful offenders rules out the possibility that it could become sensitive. Indeed, it was argued above that moderation of censure could be the result of such sensitivity to the offender's input. But Brownlee seems to assume that the state is unresponsive. She argues that the generic script misrepresents the repentant offender 'by presenting her as being of a similar mind and attitude as unrepentant offenders who need to be brought to appreciate the wrongness of their acts and the reasons for reparation and repentance. As such, her punishment, when

comparable in harshness to that of an unrepentant offender, is indefensibly dismissive of her sincere emotions of remorse, regret, and repentance, and her fervent desires to remedy relations' (ibid.: 60). The simple response at our disposal seems to be that the state should moderate the censure of the repentant offender so that it ceases to be indefensibly dismissive.

Indeed, this would be entailed by her third condition for dialogue, requiring that the parties *attend* to each other's contributions and *modify* their responses in light of those contributions. 'If the state, through its criminal justice system, does nothing to alter its communications to such an offender in light of her attempted responses, then the offender's communicative efforts seem to be, in the eyes of the state, mere acts of expression and not communication' (ibid.).

Brownlee's argument seems to be that (within Duff's theory) the state does not heed the offender's attempts to communicate and thus the conditions for dialogue are not met. But this disregard is not presented by Brownlee as principled and thus practice can change if given good reason. My account provides such good reason, and renders plausible a dialogic model of censuring.

Summary

Before moving on, it would be useful to summarise the development of my positive thesis so far. We began in the preceding chapter with the familiar notion of culpability. It was argued that since culpability is determined by the extent to which the offender is accountable for the criminal act, remorse could not justify mitigation with relation to culpability, as culpability for the offence is temporally static. It was then argued that blameworthiness as responsibility for the crime could not be altered by remorse, as acknowledgement of responsibility is necessary for remorse. Although

blameworthiness understood as an assessment of the offender's character – his attitudes towards others – could coherently change as a result of remorse, it was seen not to justify mitigation. This was because we want to censure the offender in connection to her offence and this understanding of blameworthiness focuses only on an assessment of the present person, resulting in sentencing on character.

Moving on to newer territory, this chapter began by drawing a distinction between blameworthiness and the extent to which an agent should be blamed, arguing that this was a distinction that might have relevance to official censure. Having established the centrality of censure to communicative retributive accounts, I then argued that the censuring is best situated within a dialogical model, as it anticipates and hopes for particular responses and thus should be attentive to any input the offender wishes to make. After demonstrating compatibility with the principal theories (with their emphasis on addressing the offender as a rational moral agent, on seeking a response from the offender, and on his participation in the activity) I explained why remorse would be the relevant input an offender could make to the dialogue. Since remorse is the hoped-for response, censure in a dialogical model should respond in a way that reflects this. I then returned to von Hirsch and Ashworth's quasi-retributive grounds for mitigation to highlight their commitment to censure being sensitive to the offender's capacity for reflective process, and the consequences this has for the severity of censure. Having established that this commitment served to support mitigated censure for the remorseful offender, I explained why mitigation of censure results in mitigation of punishment. Finally, I defended my account against a challenging argument that held that dialogue was an inappropriate model for sentencing and that remorse in particular could not be incorporated into such a model.

Conclusion

In light of my arguments, it would seem that communicative retributive theories such as von Hirsch's and Duff's should adopt a dialogical model of censure, receptive to communications of remorse, where forthcoming. In order to deny this, they would have to argue that remorse was not relevant to the sentencing dialogue or that the communication should anyway be only one-way – from the state to the offender. I have argued that both these contentions would make for a less preferable state of affairs.

Within von Hirsch and Ashworth's work, their claims relating to the significance of the offender's capacity to engage in moral reflection were argued to make this requirement even more pressing. I therefore suggest that mitigation on the grounds of remorse is not only compatible with von Hirsch and Ashworth's theory, but required by it.

Compatibility was seen to be less clear within Duff's theory. The complexity stems from the idea that punishment not only communicates censure to the offender, but also communicates his repentance by serving as a penance: punishment is a 'communicative, penitential system' (Duff 2001: xix). From a censure perspective, my arguments for the superiority of a dialogical model - and the consequences for the remorseful offender that fall out of it - should apply. However, even if the offender deserves less censure, what he has to do to make up for his offence - to be reconciled with the community - may not be affected by remorse. Since remorse does not reduce the seriousness of the offence, the burdensomeness of the penance needed to express the appropriate repentance remains the same, despite remorse. Although I argued that a sufficiently burdensome self-imposed penance might go some way to reducing what the state needs to impose, remorse *alone* does not have the expressive quality that

Duff requires of penance. Therefore, Duff's vacillation between emphasising the censuring and penitential functions of punishment result in a corresponding vacillation with regards the mitigating potential of remorse. If punishment communicates both censure and the offender's ideal apology, but only the severity of the censure deserved is reduced, which communicative function takes primacy?

Regardless of the answer to this question, my arguments that seek to explain why the severity of the censure deserved is reduced hold just as forcefully for Duff's account as for that of von Hirsch and Ashworth.

The next chapter will contrast my account with others, ultimately arguing for its preferability for providing a justification for the mitigating role of remorse at sentencing.

Chapter Nine

Comparing my account with others

Having presented and argued for my justification for the mitigating role of remorse I will now elucidate it further through comparing it with two alternative justifications found in the literature. In doing this, I will also argue that my account is the most preferable. The first comparison will be made between my account and one that offers a justification that appears similar to mine – that the offender’s remorse calls for recognition. However, I will explain how the arguments involved in reaching our conclusions importantly differ. The second comparison I make involves accounts that seek to show the judicial exercise of mercy as a justified state response to the remorseful offender. I will show that, at least for the sentencing factor of remorse, mercy is not the preferable justification.

Comparison one: an account with similarities

Steven Tudor (2008) makes an argument that has parallels with mine. His main thesis is that remorse should be *recognised* or *acknowledged* by mitigation, with the effect that the sentence’s ‘communicative import’ is morally deepened and more ‘attuned to the person to whom it is primarily directed’ (ibid.: 249). Although this argument sounds similar to mine, I am going to explain how my account is different and how we reach our conclusions in different ways.

In order to support his thesis that mitigation should recognise or acknowledge ‘the offender *in* his remorse’ (ibid.: 248), Tudor sets out to show two things. He asks ‘*how...*does mitigation of sentence “recognise” or “acknowledge” the offender’s

remorse? And, moreover, *what is the value* in such recognition, especially to permit or indeed require a reduction of sentence?’ (ibid., emphasis added)

So, first, Tudor wishes to show that there is a plausible way in which we can conceive of sentence mitigation as recognition of the offender’s remorse. He shows how sentence mitigation can acknowledge the offender’s remorse by analogising informal with judicial communication. He suggests that ‘formal punishment may – and normally should – take notice of what an informal reproach would properly note’ (ibid.: 253). Informally, an interlocutor will respond in various ways to communicate acknowledgement of a wrongdoer’s remorseful response to what he has done. In the same way that an informal interlocutor may modulate his speech to acknowledge the other’s remorse, so a judge can modulate his communication by reducing the severity of the sentence. However, so far, this is simply a matter of demonstrating something to be possible, with no value judgement attached to it (although he begins – with the word ‘should’ – to make a social-norms-based argument, to which I return below). We might try to show how, for example, intentionally torturous punishment might communicate a hatred of the offender to him. However, the mere possibility of such communication does not justify it and, of course, in this torture example it would not be justified.

Having provided an account of sentence mitigation as a plausible means of communicating recognition, Tudor then offers an answer for his second question – the second stage to his argument supporting his overall thesis. Whilst he seems to find *prima facie* support for value in the fact that informal reproach appears to function in an analogous way, he confirms this intuition by offering an independent argument for the value of the judicial practice.

Tudor rhetorically asks ‘...what is the *value* in such recognition of the remorseful offender? And more pointedly, is that value enough to allow the otherwise just sentence to be reduced’ (2008: 250). The implication seems to be that the sociological norm of modulating one’s speech in response to remorse is not sufficient to justify the analogous judicial modulation of sentence. According to Tudor, value independent of any found in the reflection of social norms needs to be identified in order to provide this justification. It is telling that he emphasises that for his argument, any value that he might identify needs to be of a great enough magnitude to outweigh, or must defeat, the opposing value of justice. This situation seems to require a ‘weighing up’ of values: justice on the one hand and the value of recognising the offender’s remorse on the other hand.

So what is his argument for the value of recognising the offender’s remorse? Tudor’s argument seems to be part rights-based and part consequentialist. Recognising the offender’s remorse pays moral respect to him (to which we assume he has a right), fosters psychological harmony, as it confirms the offender’s self-conception, and consequently helps to shape the offender’s future conduct in relation to this self-conception. Tudor quotes Charles Taylor as writing ‘misrecognition shows not just a lack of due respect. It can inflict a grievous wound, saddling its victims with a crippling self-hatred. Due recognition is not just a courtesy we owe people. It is a vital human need’ (quoted on page 251). For Tudor, avoidance of these adverse psychological consequences through acknowledging a person’s self-conception is valuable and it is this value which he suggests justifies mitigation as a form of recognition.

However, since Tudor derives these arguments from a general requirement to recognise and take seriously the self-conceptions of others, it is hard to see how this

would not have further implications for sentencing given the myriad conceptions various offenders might have of themselves. Recognition, Tudor says, is ‘essentially one person’s acknowledgement and confirmation of another person’s self, especially her self-conception’ (ibid.: 251). What if, for example, an offender saw herself not as an offender, but as fighting a cause? Perhaps she was strongly committed to the idea that sports cars were significantly destroying the environment. As a self-conceived eco-warrior, she had been searching out such cars and slashing their tyres under the cover of darkness. Let us imagine she has slashed all four tyres of 23 sports cars. In the eyes of the law, this offender has caused extensive criminal damage to property. In her eyes, however, she has been doing her bit to save her planet. Her self-conception is as a friend of the environment, fighting against those who wish to destroy it.

Given Tudor’s argument, it is hard to see why this alternative self-conception (eco-warrior rather than remorseful wrongdoer) should not equally be recognised. Does she not also have the right to acknowledgement of her self-conception? Tudor supports this right with his consequentialist argument that misrecognition can do great psychological harm. It strikes me that the eco-warrior may well suffer the same frustrations and feelings of worthlessness as the unacknowledged remorseful wrongdoer.

I should stress that Tudor’s argument is simply that the remorseful offender’s self-conception should be *acknowledged* and not that it should be rewarded. His convincing argument on this point focuses on the inappropriateness of the tone of celebration implied by rewarding remorse (see pp 247-8).⁵⁹ Thus there is no value judgement in simple recognition and acknowledgement – in just acknowledging that

⁵⁹ The main thrust of Tudor’s argument is that we may ‘object to the whole idea that remorse deserves any sort of reward, compensatory or congratulatory, on the basis that feeling remorse is what a wrongdoer ought to feel, and it is not appropriate to reward people for simply doing or feeling what they ought to do or feel’ (2008: 247-8).

the offender has a particular self-conception, we are not saying whether we judge it good or bad, or right or wrong – and so there seems to be no reason why the eco-warrior should not be recognised in the same way, on the grounds that this avoiding psychological harm.

Further, if we return to Tudor's face-to-face conversation analogy, we can see why value-neutral acknowledgement in general may lead to misunderstanding. A person can communicate in a way that acknowledges his interlocutor's expression of remorse – perhaps he just wants the wrongdoer to stop trying to convince him that he is remorseful – but still not want to communicate that he attaches any value to it. Mitigation of punishment in such a case – where there is a change in tone to prevent the offender continuing to insist on his remorse – would be akin to sending a false message if this acknowledgement were to be interpreted by the wrongdoer as something like respect. If punishment is understood communicatively, then it always *says* something. In my dialogical account, the quality of the censure changes to respond to the offender's remorseful input – the input that censure anticipates and hopes for.

We can now see the difference between my argument and Tudor's. Tudor's argument is driven by social norms and utilitarian benefits, whereas mine is derived from a conception of censure which takes place in a dialogical framework. Tudor begins with the sociological description of how remorse might be recognised through mitigation which he then argues for the implementation of on the basis of the value of confirming an offender's self-conception. My argument, in contrast, begins with the claim, for which I argued, that a dialogical model of communicative sentencing is superior to a one-shot 'yelling at'. The obligation to mitigate the censure delivered – to engage with the offender by responding to his remorseful contribution to the

dialogue – then falls out of commitment to this model. Whereas Tudor weighs up his purported value of acknowledging the offender’s self-conception against the value of justice, my argument makes mitigation on the grounds of remorse integral to correctly delivering censure. Less severe censure becomes deserved. If there are further utilitarian benefits in confirming the offender’s self-conception then this is a bonus but the justification of mitigation, on my account, does not rest on these benefits.

Comparison two: dialogical censure and the exercise of mercy

My account can usefully be contrasted with the alternative principal justification for mitigation on the grounds of remorse – that of the justified exercise of mercy. I am going to show how the two accounts differ, explain why my ‘censure in a dialogical context’ account is preferable and note how it avoids some of the most serious challenges that the ‘mercy’ account faces.

A prominent example of the mercy account has been developed by Tasioulas. His central argument is that there is room within communicative theories of punishment for the exercise of mercy to be a justified judicial response to an offender’s remorse. According to Tasioulas, ‘mercy...embraces reasons for leniency that arise out of a charitable concern with the well-being of the offender, in particular, the compassion we rightly feel towards him as a potential recipient of deserved punishment given various other facts about his life and circumstances whose salience is not captured by the retributive norm’ (2006: 312). Remorse is considered a relevant fact about his life: ‘...a decent concern with the latter’s welfare can justify us in tempering the punishment deserved in order to take account of charitable reasons furnished by his repentance.’ (ibid.: 318).

In line with my account above, Tasioulas sees repentance as relevant to the communicative process (see *ibid.*: 317), yet he departs from my arguments in his claim that this justifies a merciful response which functions outside the domain of retributive justice. I am going to argue that my account is preferable.

Showing an offender mercy and engaging dialogically with his remorse are two different scenarios. As quoted above, mercy stems from a charitable concern with the well-being of the offender. This involves feeling compassion towards him on the basis of his prospective punishment – compassion evoked by some relevant fact (his repentance). Firstly, it is not obvious that a concern for the wellbeing of the remorseful offender is even intuitively the driving force of the impulse to mitigate his punishment. The concern for the offender's wellbeing cannot simply be a response to the offender's impending sanction, otherwise all offenders would be candidates for mercy, since the wellbeing of all offenders is diminished by their punishment. There might therefore be something about the *remorseful* offender's situation that makes his experience particularly burdensome. It could be argued that an offender's punishment *plus* his painful experience of remorse diminishes his wellbeing to a greater extent compared to that of the unremorseful offender and consequently we are more concerned for him and therefore feel that mercy becomes justified. However, I do not think this is Tasioulas' argument. Rather, it seems that we do not have a charitable concern for the offender's wellbeing *unless* he repents, not that we have a greater concern for the remorseful than the unremorseful.

I am going to argue that compassion and charitable concern for wellbeing do not adequately capture the justification for mitigating the punishment of a remorseful offender. Feeling compassion for an offender as a recipient of punishment, given facts about his life and circumstances, has strong overtones of pitying the offender.

Such a conceptualisation of mercy is shared by Murphy (1988c: 166, emphasis added): ‘mercy is often regarded as found where a judge, out of *compassion for the plight of a particular offender*, imposes upon that offender a hardship less than his just deserts’. A parallel perspective was present in Manson’s (2011) argument for mitigation justified by ‘cultural demands’, where *sympathy* justified leniency.

Whilst such responses might be appropriate for an offender who is, say, terminally ill or very old, they are not appropriate attitudes towards the remorseful offender. We do not (or should not) pity an offender his remorse⁶⁰ (although we might recognise that it can be emotionally very painful), rather, we expect (or at least hope for) his remorseful response as we engage with him in communication. Mitigating on the basis of mercy might send the message that we, to an extent, feel sorry for the remorseful offender, not that we want to be receptive to the offender’s input into the dialogue – his appropriate moral response to his wrongdoing.

Such responses as compassion, concern, and sympathy in the face of remorse may actually devalue the offender’s remorse if they were the driving force for mitigation – as if remorse was simply suffering. It is true that we probably do feel more compassion towards the remorseful, in virtue of their appropriate response and their consequently seeming more ‘likable’. However, mitigation is justified because we engage with their response *per se*, not because the response inspires compassion and thus an impulse to want to see them suffer less.

My argument is strengthened if we compare how well the contrasting accounts fit within retributive theories founded on communication and censure. Within my account, the offender’s remorseful input into the dialogue about his wrongdoing results in the mitigation of the censure he receives. I argued that the remorseful

⁶⁰ We might pity an offender who experiences an excessive, pathological, amount of remorse since he suffers emotionally to a degree far greater than we would consider appropriate.

offender should be censured less severely, as the dialogical model requires engaging with his relevant input. Thus, mitigation on these grounds functions within the realms of deserved censure and retributive justice. Engaging with the offender's remorse is internal to the enterprise of delivering censure. Mercy, on the other hand, is external to the censure the offender receives – it has nothing to do with blaming. On my account remorse mitigates censure, on the mercy account remorse justifies leniency.

Based on concern for the offender's wellbeing, mercy was said to temper the punishment the offender could justifiably receive. Exercising mercy does not form part of the censuring dialogue, rather, it distorts it in the direction of leniency. Whereas remorse alters the appropriate content (degree of) censure – given a need to respect rational moral agency and engage in dialogue – mercy has no influence on the content of the deserved censure. What mitigation on the grounds of mercy might communicate is that compassion is felt towards the offender, evoked by his remorse. This is not a substantive response on the part of the state within the dialogue (following the offender's communication of his remorse), but is an emotional, humane reaction to his upcoming suffering, given his moral distress. On my account a substantive response *is* given through mitigation: 'we hear your communication and respect that you have responded in the appropriate way to your wrongdoing'. On the mercy account, what could be *inferred* by the offender (as mercy does not constitute a direct message to him) is that his remorse has evoked the charitable concern of the judge.

This is not to say that I think there is no room at all for mercy within sentencing. Rather, I hope I have shown how it is not required within communicative theories as a justification for mitigating the remorseful offender's punishment. Mercy may, however, be much more plausibly justified in cases where an offender is, for

example, terminally ill. In such cases, the language of compassion, concern for welfare and sympathy are entirely fitting. Moreover, the message that the offender's circumstances invoke such a response *is* the appropriate message. It would be absurd for the offender's illness to influence the substance of the censure as it is not relevant to the dialogue about the wrongdoing. Remorse, however, as the offender's response to his wrongdoing, has been argued to be *the* relevant contribution the offender can make. Remorse is censure-relevant, illness is not. Mercy is not the most appropriate response to remorse, but might be to terminal illness.

However, an objector might remain adamant that remorse is a legitimate source of reasons for merciful leniency, suggesting that we might still be merciful to the remorseful offender who has independent (censure-based) reasons for mitigation. This would then have an additive effect when it came to reducing the severity of the offender's sentence. In response to this I would emphasise that receipt of mercy is not a right and the fact that remorse already serves to mitigate the censure (and hence punishment) the offender deserves might remove a (compassion-based) obligation (see Tasioulas 2006) to exercise mercy that would otherwise be present. If a judge were to nonetheless exercise mercy and reduce the offender's sentence further, he would still be acting within the justifications mercy provides, but without obligation. So, I would argue that my account, having justified mitigation before a sentencer gets to external considerations of mercy, removes the obligation to further mitigate on independent merciful grounds.

Whether mercy is an appropriate response to remorse is not, then, of consequence overall for my argument as, if my justification is accepted, it effectively pre-empts the need for mercy, rendering it unnecessary and without obligation. This argument requires that my justification is indeed internal to calculations of deserved

censure, and that mercy is external, so that the mitigation that I justify is logically prior to that which mercy theorists propose. I will now further defend this argument.

A closer look at the internal-external distinction

In a recent paper, Tasioulas (2011) examines the boundaries of the internal-external distinction that I have employed when making a case for my account over a mercy-based account. His argument is that, seemingly contrary to my claims, mercy operates *internally* to the logic of punishment. The truth or falsity of this would seem to be crucial to the success of my account. This is because it was argued that one of its key strengths is that it is superior to a mercy account, as it provides a justification for the mitigating role of remorse that is intimately connected with the operation of censure-based desert. If mercy were to also operate thus, then my account would fail to have this advantage. However, Tasioulas is careful to explain what he means by ‘internal’, defining the domain to which it relates in broader terms than I do.

Having carefully outlined his position, I will explain why my account provides a justification that is more intimately ‘internal’ than does Tasioulas’ mercy account. I will also take issue with a couple of features of his account, questioning whether mercy is even as internal to his concept of penal justice as he believes, showing how his focus on remorse as a ground for mercy makes his case seem stronger than it perhaps is. Instead, the intimate connection with punishment that he establishes for remorse demonstrates its unique status within the set of potential grounds for mercy. It will be argued that the reasons for its uniqueness actually lend support to the contention that it operates differently to other potential grounds for mercy, in a way consistent with my account.

In his paper, Tasioulas takes issue with Duff's positioning of mercy as operating externally to the concerns of justice. He seeks to show why Duff is mistaken, pointing to a conflation of two meanings of 'justice' – 'retributive' and 'penal'. Whilst mercy *is* at odds with the former, Tasioulas (2011) argues that it is compatible with, indeed, internal to, the latter. We must be very clear about the distinction he draws between retributive justice and penal justice.

In defining *retributive* justice, Tasioulas (ibid.: 39) explains that the content of the retributive norm is that it requires punishment when it is deserved as a way of communicating censure for wrongdoing. The quantum of punishment deserved as censure is proportionate to the gravity of the wrong-doing in question, where the latter is a complex function of the harm that has been caused or risked and the degree of culpability manifested. This is the conception of retributive justice shared by most communicative theorists, including Duff, von Hirsch and Ashworth.

Tasioulas then contrasts this conception of justice with a broader one, which he calls '*penal* justice' saying:

sometimes 'justice' refers to the whole of morality, or the whole of interpersonal morality, or (in a more authentically Aristotelian vein), that part of interpersonal-morality that may be aptly legalised. Let us, following in the precedent set by translators of Aristotle, call this the concept of 'universal justice', focussing on the third specification just given. [ibid.: 40-1]

He argues that it is this 'universal' sense of justice, rather than retributive justice that is alluded to in expressions such as 'Criminal Justice' and 'Penal Justice'. Penal justice extends to further considerations, outside of retributive desert, that are thought to have a legitimate bearing on punishment. Thus, if mercy is apt for legal embodiment, then it falls within – is internal to – penal justice: '[f]or on the understanding of 'justice' as interpersonal morality, mercy (like charity generally) *is* part of justice' (ibid.: 41).

Whilst Tasioulas agrees with Duff that mercy does not operate inside retributive justice, he wants to show that there is still an important sense in which mercy does operate internally to ‘justice’ – or, the ‘logic of punishment’ – when justice is understood in the sense that is more naturally implied when people speak of criminal or penal justice.⁶¹

Mercy and its relationship to retributive justice

Tasioulas explains his conceptualisation of mercy, emphasising that it is concerned only with the treatment of those who deserve punishment. As in his previous work (see above), he describes mercy as ‘a source of genuine but defeasible reasons, grounded in proper concern for the welfare of the offender, for leniency towards him, i.e. for punishing him less severely than he deserves to be punished according to retributive justice’ (ibid.: 40).

Crucially, in relation to retributive justice, mercy’s reasons do not affect the assessment of the punishment the offender deserves, nor do they confer on the offender a right to more lenient treatment. But, he argues, they influence our judgement as to how much punishment, if any, is all-things-considered *justified*.

In order to understand mercy’s relationship to penal justice, we need to establish what it is that makes a particular value count as *internal* to the institution of criminal punishment (penal justice), and thus able to influence the judgement of how much punishment is *justified* (cf. deserved).

⁶¹ Von Hirsch and Ashworth (2005: 168) also point out a similar distinction between a narrow understanding of justice and a broader all-relevant-things-considered understanding: ‘When one speaks of the “just” sentence, this may denote either the sentence that is just according to all the normal criteria (but which may subsequently be altered to take account of exceptional considerations, such as mercy or equity factors); or else it may refer to the sentence that is just, taking account of every factor with a good claim to affect sentencing (which would include mercy or equity factors). The former is a defeasible concept of justice, whereas the latter is a conclusory concept.’

What makes a value internal to a conception of justice?

Tasioulas claims that we should ‘relate the idea of values internal to the institution of punishment to our best account of the justification of that institution’ (ibid.: 43). Although retributive justice is at the core of Duff’s communicative theory, and defines its logic, Tasioulas argues that this does not preclude other considerations from playing a role within this logic (being internal to it), but only on the condition that they are appropriately related to the over-arching norm of retributive justice. For Duff, this involves two dimensions: first, the other considerations must bear an ‘appropriate constitutive relationship to retributive justice’ (ibid.: 44). The second dimension is that the other considerations ‘do not, insofar as they pertain to the inner logic of punishment, licence punishing offenders either more or less severely than is required by retributive justice’ (ibid.). However, Tasioulas believes that these criteria are only true for what is internal to *retributive justice*, and sets out to show how only the former is required for a value to be internal to his broader concept of *penal justice*.

What makes mercy internal to penal justice?

With reference to Duff’s conditions for internality presented above, Tasioulas (2011) argues that mercy satisfies the former but not the latter. It satisfies the former because mercy is a source of reasons that are dependent on retributive justice: ‘mercy is not a self-standing value, but constitutively parasitic on retributive justice...and can play a significant role *within* the boundaries set by retributive justice’ (ibid.: 46). This is because, he argues, mercy is a dependent value: it can only generate reasons of its

own when independent reasons of retributive justice to inflict a punishment already exist. If no punishment is deserved, there is nothing to be merciful over.⁶²

However, the possibility that mercy might justify a less severe punishment than that required by retributive justice means that mercy does not meet the second condition. Tasioulas explains that this interpretation makes sense of why Duff believes mercy is an ‘intrusion’ into the inherent logic of criminal punishment, and hence external to it.

Such a conclusion would be valid if *retributive* justice were the domain of relevance. However, within the broader notion of *penal* justice, Tasioulas argues that only the former condition needs to be met. Tasioulas expands on the first condition (that considerations can play a role within penal logic if they bear an appropriate constitutive relationship to retributive justice): other facts can be taken into account when determining the level of punishment justified as censure provided ‘these further facts...display a requisite connection to the wrongful act to make them bear on the question of what is justified *as* censure for wrongdoing’ (ibid.: 48). He shows how this is true for repentance as a ground for mercy, explaining that the connection is forged by the fact that repentance is precisely the appropriate personal response to the offender’s wrongdoing *qua* wrongdoing. He emphasises:

Indeed, it is precisely the response that the institution of punishment should ideally elicit from criminal offenders. The upshot is that repentance has place *within* the two-way communicative process that is punishment, rather than something that intrudes upon that process from the outside. [ibid.]

⁶² I am not sure this is entirely true. I think it is not a mistake to describe the villain who spares his victim, at the last minute, in response to his pleas, as having shown mercy despite his victim deserving none of what had and could have befallen him. However, if we *specify* the context of retributive justice, then mercy only makes sense when an offender receives less punishment than he deserves, necessitating that a particular quantum was indeed deserved. I will therefore concede Tasioulas this relationship of mercy to retributive justice.

Mercy, then, with the parasitic nature of its relationship to desert – one cannot show mercy unless there is deserved punishment to waive – and the role its grounds (such as repentance) play within penal logic, make it internal to penal justice. The exercise of mercy more generally Tasioulas seems to justify by emphasising its status as a virtue. According to Tasioulas, mercy is a ‘manifestation of a more general value that may be called charity, compassion or impersonal love, a value that characteristically issues in reasons to advance the wellbeing of others, especially by ministering to their needs or relieving their suffering’ (ibid.: 40).

Contrasting mercy (operating internally) with leniency (operating externally)

Tasioulas (2011) further demarcates mercy’s internal position by contrasting it with the operation of simple leniency, which he argues occurs externally to the logic of punishment. He says that there are certain grounds for leniency – for punishing people less severely than they deserve according to retributive justice – that are not aptly regarded as forming part of the logic of punishment. Rather than being integral to the ‘tailored justification for punishment’, they belong to the ‘diffuse class of other “all-things-considered” factors that might defeat, in whole or in part, a *pro tanto* case for punishment’ (ibid.: 45). He gives examples of possible grounds for leniency citing, amongst others, the offender’s having saved thousands of innocent lives threatened by a retaliatory attack; the offer of leniency in exchange for a plea of guilty in order to avoid an expensive and protracted trial; gratitude, as when an offender’s punishment is moderated in recognition of their outstanding record of community service; and even perhaps if the offender’s medical research were to be on cusp of finding a cure for cancer.

He argues that the intuition that motivates the internalist view of mercy is that the standard grounds for mercy in criminal punishment (which he lists as repentance, grief, disadvantaged up-bringing, and illness) are categorically distinct not only from the considerations that bear on the level of punishment deserved under the retributive norm but also from the paradigmatically extraneous considerations. The reason he gives for this is that, although they do not affect the punishment deserved by the wrongdoer, where this is proportioned to the gravity of their wrong-doing, they nonetheless appear to have an intimate connection with the logic of punishment, unlike such considerations as gratitude, social stability or democracy promotion.

Does Tasioulas succeed in positioning mercy internally?

With Tasioulas' internal-external topography mapped, I will now show that my account still positions the justification for remorse-based mitigation more internally than does Tasioulas'. Further, I will argue that Tasioulas' account does not even achieve what it claims. I will question whether mercy is as specially internal to the logic of punishment as he believes, and show how his focus on remorse as a ground for mercy makes his case seem stronger than it perhaps is. I will argue that the intimate connection with punishment that he establishes for remorse demonstrates its unique status within potential grounds for mercy. The reasons for its uniqueness actually lend support to the contention that it operates differently to other potential grounds for mercy, in a way consistent with my account.

Whilst Tasioulas argues for the operation of mercy as being internal to penal justice but not retributive justice, I argued in my account that remorse-based mitigation operates internally to deserved censure. A dialogical approach to censure, which I argued maximises the legitimacy of the censure, must moderate the censure

delivered if an offender has communicated his remorse. Since it is *censure* that is deserved and not punishment *per se* (although punishment is the preferred means of communicating the censure) then remorse operates within retributive justice on communicative retributive theory. Importantly, it is not that Tasioulas is working with a different conception of retributive justice. Recall: ‘the content of the retributive norm is that it *requires punishment when it is deserved as a way of communicating censure* (or, as I would rather say, more specifically, blame) for wrongdoing’ (ibid.: 39, emphasis added). So we are not talking at cross purposes: for Tasioulas, too, it is actually the *censure* that is deserved and which must be proportionate to the wrongdoing, communicated through the medium of punishment.

Problems

But, there are problems with Tasioulas’ account as it stands. Let us recall the key elements of his argument:

- 1) Mercy is ‘internal to the logic of punishment’ as it is necessarily parasitic on the concept of desert.
- 2) Mercy on grounds of repentance is particularly justified as repentance is precisely the appropriate personal response to the offender’s wrongdoing *qua* wrongdoing. Thus, it displays the requisite connection to the wrongful act as it bears on the question of what is justified as censure.

Is mercy’s parasitic relationship to desert the foundation of its special internal status? Whilst Tasioulas is correct that mercy does bear this relationship – the exercise of mercy to benefit an individual necessarily requires that that individual deserves a certain amount of punishment – I do not see that this is any different from the way leniency in general is related to desert. To me, it seems that *any* species of leniency is

parasitic on the notion of desert, since you need to know what the default severity is before you can amend it the direction of moderation. Take, for example, the guilty plea *discount*. It can only function as a discount if there is a full price.

The other option is that it is not mercy's relationship to desert that is special, but the relationship of the *grounds* for mercy to retributive punishment that confer mercy's special internal status. Tasioulas makes a compelling case for the intimate relationship repentance has with deserved censure as an example of how these grounds operate. However, on closer inspection it seems that, of all the grounds Tasioulas lists, it is only remorse that bears this intimate relationship. Whilst repentance does, as he argues, bear an appropriate constitutive relationship to retributive justice, grief, disadvantaged up-bringing and illness do not seem to. Repentance displays the requisite connection to the wrongful act to make it bear on the question of what is justified *as* censure for wrongdoing. Grief, disadvantaged up-bringing and illness, however, do not display any such connection.

Perhaps arguments could be advanced in the case of disadvantaged up-bringing, such that the offender is somehow less culpable or less tied to the social contract, but this is not the connection Tasioulas has in mind. The relationship repentance has with the wrongful act places it '*within* the two-way communicative process that is punishment, rather than something that intrudes upon that process from the outside' (ibid.: 48). Grief and illness have no connection to retributive justice. Tasioulas provides no arguments as to how they might. He simply states that they have 'an intimate connection with the logic of punishment' (ibid.: 45).

Further, he generally seems unprincipled in his approach to what the legitimate grounds for mercy even are. He cites Card's (2002: 192) proposition of giving 'a bit of compensatory good fortune' to offenders who have suffered

‘extraordinary severe undeserved misfortunes’ in their lives. He states that this consideration does not fall under mercy, explaining that ‘unlike the classic grounds of mercy, it is not a consideration integral to the logic of punishment’ (2011: 47). But, it is not clear how this ground for leniency is less intimately related to the logic of punishment than is disadvantaged upbringing. Indeed, the latter appears to be a species of the former.

The ‘internal’ group certainly is not homogenous. Even though Tasioulas goes some way to accepting this⁶³, it is still not clear what is special about the grounds for mercy, except, perhaps, that they tend to elicit sympathy, which ties in with Tasioulas’ conception of mercy as concern for the offender’s welfare, born out of charity, compassion or impersonal love.

Compounding this confusion is Tasioulas’ vacillation between citing *mercy* as being integral to the logic of punishment and citing the *grounds for mercy* as bearing this relationship. I suggest he makes a kind of category mistake when he relates both repentance and mercy to the first condition for internality (that the consideration must bear an appropriate constitutive relationship to retributive justice). As argued above, none of the other ‘classic grounds’ that he lists (grief, disadvantaged up-bringing and illness) bear the same constitutive relationship as remorse does so, if he wants to rely on his argument, it is the merciful response to the grounds, and not the grounds that must bear this relationship to punishment. It was also argued, however, that mercy has no closer relationship to the logic of punishment than does leniency more generally.

⁶³ In a footnote he states: ‘in an additional layer of complexity, one should allow that some grounds for mercy (e.g. repentance) might be situated closer to the core than others (e.g. disadvantageous upbringing)’ (ibid: 48, note 15).

The unique nature of remorse

As argued above, remorse is the only of the grounds for mercy that does have an intimate connection with the logic of punishment but, I will argue, this connection is to do with deserved censure, not reasons for mercy. Remorse might be a legitimate ground for mercy for other reasons but it cannot be its connection to censure that is necessary for this, or else no other grounds would qualify.

For Tasioulas, repentance is ‘a reason for tempering the deserved punishment not on the basis of what the offender deserves or has a right to, but out of a due regard for their welfare’ (2011: 50). Indeed, he says that since repentance ‘does not come within retributive justice, it seems highly unlikely that it can be anything else [than a merciful consideration]’ (ibid.). But I have argued that it does come within retributive justice, within communicative theories. In fact, the unique way in which Tasioulas argues repentance is connected to censure – being within the two-way communicative process that is punishment – supports this view. The fact that it is only repentance that operates thus – singling it out from other grounds for mercy – lends support to the idea that it might play a different mitigating role. Contra Tasioulas, my account shows that repentance is a reason for tempering deserved punishment on the basis of what an offender deserves.

I think that Tasioulas’ account at first seems plausible because of his focus on repentance. This helps him because (as I have argued) repentance *does* have an intimate relationship with censure and hence punishment, which he wrongly extends to the grounds for mercy in general. However, I have shown that repentance is anomalous amongst Tasioulas’ grounds for mercy and that the reasons for this also render my account particularly plausible.

Summary

I argued that one of the key strengths of my account in comparison with mercy accounts was that remorse operates internally to censure-based desert. Mercy, I argued, operates externally to this domain. Tasioulas' arguments seemed to challenge this, threatening my account's preferable status. However, I have shown that the domain to which Tasioulas argues mercy is internal (penal justice) is more extensive than that within which I situated the mitigating role of remorse (retributive justice).

Further, I have argued that his account is not as persuasive as it first seems, as he focuses on an anomalous ground for mercy and extends its unique characteristics erroneously to all the other grounds for mercy. To save his argument, he would need to show that it is mercy, and not the grounds for mercy, that bears the important relationship to the logic of punishment. I showed that this relationship is not any truer for mercy than it is for leniency in general and that recourse to demonstrating mercy's special relationship based on the relationship of the grounds for mercy to punishment was circular, as neither relationship was established independently. Finally, it was argued that the anomalous nature of remorse instead invites the possibility of an account such as mine, which I conclude remains preferable to a mercy account.

Further reasons to prefer my account

Not only is my account theoretically preferable to the mercy account, it also avoids some of the serious objections that mercy theorists have to deal with. Mercy is often seen as a 'soft' response which is in tension with the demands of justice (see Murphy 1988c: 162-83). The general argument put forward is that if mercy operates outside the domain of justice then exercising it will produce injustice. Within retributive theories, the seriousness of the offence is set by the harm caused and the offender's

culpability. The offender is censured in relation to this offence in the context of dialogue which, I have argued, should be receptive to any relevant input from the offender. The sanction will then be in proportion to the deserved censure. Remorse, on my account, affects the overall censure deserved, whereas it does not on the mercy account.⁶⁴ Operating therefore to deviate from what is deserved, mercy appears to prevent retributive justice. On my account, mitigation on the grounds of remorse is required by respectful, dialogical censure.

Further concerns over mercy arise when we consider its application. Whereas on my account mitigation would operate systematically (since it is required), allowing mercy to guide mitigation does not obviously produce this result. Murphy has argued that mercy, as a manifestation of an individual's sympathy, will be exercised randomly (1988c: 183). The judicial exercise of mercy is seen as the pursuit of 'some private, idiosyncratic, and not publicly accountable virtue of love or compassion' (ibid.: 167-8). As a judge, I may be moved in some cases but not others and certainly other judges will differ from me in their propensities to be moved by remorse. The general argument is that mitigation on the basis of individuals' varying sensibilities is likely to produce unjustified discrepancies in sentencing. Mitigation based on a consistent principle, derived from a dialogical rationale for censure, however, should be applied across the board.

Why we need a principle rather than merciful discretion

I have argued that remorse influences the censure an offender deserves, which is established prior to considering reasons to exercise mercy. So, in relation to repentance, mercy-based reasons are redundant (even if they might seem justified

⁶⁴ 'Mercy is a source of *pro tanto* reasons, defeasible in the context of an all-things-considered judgement, for punishing the offender less severely than they deserve. It is not a source of reasons showing that they deserve to be punished less severely' (Tasioulas 2006: 312).

otherwise – perhaps non censure versions of retributivism might recourse to mercy to justify remorse-based reductions). Even if mercy is justified in relation to other factors, like terminal illness, a consistent principle will always be always better if available.

Ashworth (2010: 191) speaks to this theoretical preference for articulating a principle if possible: ‘The problem is that the judicial use of the concept of mercy suggests that the sentence reduction is discretionary [in a way that is not within the expected, fair operation of discretion]’. Having cited a case in which the offender suffered from a serious medical condition that was difficult to treat in prison, he argues decisively that it is wrong that sentence reduction should be regarded as a matter for the discretion of the court. Rather, he argues, it:

must be decided whether, in principle, [a particular factor is one] that should or should not be allowed to affect sentence. If the answer is yes, it should then be for the court, as with other aggravating and mitigating factors, to assess its strength and to give appropriate effect to it. Referring to this as ‘mercy’ suggests a broad discretion, and that is only suitable for really extraordinary cases with unusual features, such as *Schumann*. [ibid.]⁶⁵

My account has provided such a principle and so, for the sentencing factor of remorse, the exercise of merciful discretion is not needed.

Could mercy constitute a principle?

However, it could be argued that there is an *obligation* for judges to exercise mercy when faced with a remorseful offender (even though the offender would not have a corresponding right to receive mercy). It could, as Tasioulas would argue, be an

⁶⁵ In this case a clinically depressed woman jumped from the Humber Bridge, carrying her young child with her. Although her intention was to kill them both, they survived and the mother kept her child from drowning. Despite pleading guilty to the attempted murder of her child, Lord Phillips CJ held that ‘there are occasions where the court can put the guidelines and the authorities on one side and apply mercy instead’. The sentence of imprisonment was quashed and replaced with a community sentence with a supervision requirement. [2007] 2 Cr App R (S) 465.

obligation created by the virtuous nature of the charitable concern embodied in mercy; or perhaps simply in the humane concern to be less punitive whenever possible.⁶⁶ This would give merciful sentence reduction on the grounds of remorse the status of a principle in the sense that Ashworth requires: it should operate systematically and should not be variable from judge to judge. However, further reasons would then have to be given to justify why remorse generates an instance of this obligation. The reason could not be that remorse evokes charitable concern and compassion, as this simply bases the justification for exercising of mercy on its antecedent feelings. This would not explain why mercy should on principle always be exercised for the remorseful – ‘you should be lenient as some people respond to remorseful offenders with compassion’ does not suffice.

The search for adequate reasons for systematic exercise might then lead to an argument such as the one I make, based on respecting the offender’s rational moral agency as a contributor to the dialogue concerning his offence. But, if this were the case, it is not clear what extra work mercy does, other than suggest that the already-justified mitigation (justified by an account such as mine) might be accompanied by feelings and messages of compassion and concern which would then only serve at most to present the institution of criminal justice as less soulless (which may not be a bad thing).

Conclusion

The discussion in this chapter has explored alternative justifications for the mitigating role of remorse at sentencing. In particular, I have attempted to show that, even if mercy-inspired leniency were to be a justifiable state response to the remorseful

⁶⁶ This is often referred to as the doctrine of ‘penal parsimony’.

offender, this justification is redundant due to the logical priority of establishing deserved censure. I have argued that censure is at its most legitimate when it is not deaf to the offender and so should be sensitive to his communications about his offence. When it is remorse that is communicated – the response censure seeks – the quality of the censure must change in response. This is reflected in the sanction imposed to communicate the censure. Estimating the harshness of the censure deserved is prior to any considerations more external to retributive justice. Whilst mercy may still have a role to play in exceptional cases, principled sentencing should be the rule, where possible. On the account I have developed, the remorseful offender receives principled mitigation.

Chapter Ten

Conclusion

I began this thesis by highlighting the significant effect that remorse can have on an offender's sentence, and the general appeal remorse has as a mitigating factor amongst members of the public and legal practitioners. I also drew attention to the fact that this practice, although widespread and cemented in policy in many jurisdictions, has no formal justification. Moreover, the theoretical literature on this question to-date is sparse and much of it *denies* that remorse should be allowed to have the mitigating effect that it often does. This thesis therefore set out to redress this lacuna, exploring various potential avenues for justification within broadly retributive theories. Necessary for this project was an enquiry into the nature and value of remorse. A clear concept was developed, and its value seemed particularly to lie in its being essential for, even constitutive of, full understanding of the wrongfulness of the conduct to which it is a response. This particularly involved understanding what it means to harm another person.

I explored various potential ways of justifying the mitigating role of remorse, finally developing an argument that centred on the idea that censure should be responsive to the offender, and that his remorse is something to which it should respond in particular. This dialogical model of censure was argued to be superior to a model in which censure delivers the same message, regardless of whether the offender has demonstrated through his remorse that he already has a clear understanding of the wrongfulness and harmfulness of his offence. It is superior because responsive censure is more legitimate and morally forceful than that which does not attend to the offender's specific level of understanding. If censure seeks a response and this

response is already forthcoming, then to nonetheless seek this response as if it was absent devalues the censure. Since punishment communicates the deserved censure, and this deserved censure is mitigated as a consequence of remorse, then the punishment is correspondingly mitigated. Remorse therefore mitigates punishment as a consequence of mitigating censure.

Having found inspiration for these arguments within a particular desert theory (that of von Hirsch and Ashworth), I argued that values to which they claim commitment lent further support to my justification. The idea of: 1) censure as seeking a response from a moral agent capable of deliberation, and 2) quasi-retributive concerns about censure being sensitive to individuals' capacities for reflection (illustrated by their recommendation of reduced punishment in the contexts of time delay and voluntary mitigation), served to lend particular support. However, my argument should hold for any censure-based desert theory.

I sought to demonstrate that my justification for why remorse mitigates was preferable to a mercy-based justification. It was argued to be preferable because it offers a principled approach, which is not susceptible to the vagaries of judicial discretion. Mitigation is deserved. Moreover, my justification was seen to be logically prior to considerations of mercy in the calculation of punishment severity, rendering any additional justification redundant.

Even if my arguments have appeared plausible to the reader, there are still a couple of worries he might have about adopting my justification as *the* justification for the mitigating role of remorse in sentencing practice. I will explain and comment on these.

How can such an abstract theory be translated into everyday practice?

The worry might be that, even if my arguments are internally coherent, it is hard to see what the implications are for sentencing practice. It is true that I have developed my justification with detailed reference to contemporary penal philosophies. As a consequence, the discussion has been almost exclusively theoretical, using concepts of desert, dialogue, censure and the idea that punishment communicates censure. What would accepting my arguments mean for day-to-day practice? The answer (I think) is that, in jurisdictions that already confer on remorse a mitigating capacity, very little would change. England and Wales already list remorse in the various sentencing guidelines as a ‘personal mitigating factor’. My thesis provides a theoretical justification for this practice. This is of great value in itself, as a principled approach to sentencing requires that all mitigating factors are justified.

Perhaps (hypothetically) a brief, succinct version of my justification might be developed. This could be included in guidelines for the reference of the court and anyone else interested in why this particular element of sentencing practice operates as it does. If someone were to question the legitimacy of remorse as a mitigating factor, there would then be a formal response, rather than none.

A further effect of adopting my justification might be a drive, by those in charge of developing the guidance, to reassert the relevance of remorse to sentencers. The elements particularly of value, enumerated in Chapter Two, might be emphasised as clues to the genuineness and depth of any given instance of remorse.

What do your arguments mean for the quantum of mitigation?

There is one practical issue that is of more concern. Whilst my thesis has offered a justification for the mitigating role of remorse, it has said nothing about the degree to

which punishment should be mitigated. Does the remorseful offender deserve less censure equivalent to a third of his punishment, or a quarter, or less? It is difficult – and would be unwise – to try to think about the quantum of mitigation appropriate for the remorseful offender in isolation. Rather, the issue needs to be assessed in the context of other sources of mitigation. If we were able to establish roughly how much remorse should mitigate in relation to other mitigating factors, then we would be better able to give an estimation of the appropriate quantum.

Whilst this is an issue better tackled by sentencing guidelines authorities, there are some tentative suggestions I can make. An enquiry into the quantum of mitigation appropriate for remorse might begin by considering the most robust sources of mitigation, thereby establishing a ceiling. In England and Wales, the guilty plea discount and first offender discount (progressive loss of mitigation) are the most consistently applied and attract the greatest amount of mitigation. The guilty plea discount is set at one third. These sources of mitigation, then, might suggest that the ceiling for the quantum of mitigation is about a third.

But remorse is not as categorical as pleading guilty or being a first time offender. Whereas an offender either pleads guilty or does not, remorse can be present to a greater or lesser degree, and the nature of the offence may change the way remorse is perceived – is it sufficient given the seriousness of the offence etc? Remorse is therefore not likely to justifiably attract as great a quantum of mitigation as the more categorical sources of mitigation.

We can further test our intuitions about the relative power of remorse to mitigate. We might compare a hypothetical remorseful recidivist with a first time offender, where the cases are identical in all other respects. This enables us to think about how much mitigating work remorse should do in comparison with being a first

time offender. The intuition would likely be that mitigation lost through being a recidivist should not be equivalent to that gained from being remorseful – i.e. the quantum of mitigation resulting from remorse is less than that which being a first time offender confers.

Even if this intuition seems plausible, it is complicated somewhat by the issue of interaction, which would also have to be considered. It might be that aggravating factors and other mitigating factors influence what we think remorse should achieve by way of mitigation. On the account I have developed, it would seem that remorse should always beget some mitigation, as the quality of the censure (and hence severity of punishment) that becomes appropriate is changed by the offender's remorse. However, focusing on the offender's capacity for refection, perhaps repetition of the same offence might require the censure to be more forcefully articulated, resulting in remorse at best preventing the sentence from being aggravated on these grounds. Alternatively, other sentencing aims – such as protecting the public – might come into play, outweighing the mitigating effect of remorse.

Finally, it may be that some notion of a 'totality principle' with regards to mitigation is theoretically appealing. If there are many mitigating factors in a particular case they might not function additively. At least some of these factors might be understood to function concurrently, thereby making it harder to suggest a concrete fraction for the quantum of mitigation that remorse will independently procure.

Given the arguments of my thesis, my intuition would be that remorse must beget a reduction that seems significant (in that it is at least noticeable) if it is to serve to communicate less censure than would be deserved absent the remorse. Despite not being able to provide a comprehensive answer to the quantum question, justifying a mitigating effect of *some* size is significant progress in itself.

Doesn't your justification require acceptance of a censure-based desert theory?

The short answer is 'yes'. A justification that explains how remorse reduces the extent to which we should censure the offender cannot work if punishment is not censure. Someone who subscribed to the benefits and burdens theory of punishment discussed earlier in the thesis would not be able to avail himself of my justification for mitigation. If punishment is an objective burden - in the form of deprivation of liberty - imposed to divest the offender of his unfair advantage, then there is no way for remorse to alter the burden required. The burden is not intended to prompt the offender to reflect on his conduct, seeking a remorseful response. Instead, it puts the offender back in balance with the rest of society.

Utilitarian theories of punishment could make no use of my arguments. The concepts of censure and desert are central to my arguments, but have no place in purely forwards-looking accounts. Crime prevention is central to these, not communication. This does not mean that no justification for a mitigating role for remorse can ever be found within utilitarian theories, just that the arguments would necessarily be different. If robust evidence were to emerge that remorseful offenders are less likely to re-offend, or are easier to rehabilitate, then remorse would have to be considered at sentencing even from a purely utilitarian perspective.

There might, on occasion, be theoretical conflict-of-interest where, as in England and Wales, a plurality of aims is set out for punishment - retributive and consequentialist. As I suggested above, it might be that, within such a scheme, the aim of crime prevention (which is unconcerned with remorse) sometimes outweighs the aim of delivering deserved punishment (which I have argued is affected by the offender's remorse). However, that there might be theoretical conflict does not

remove the justification for remorse-based mitigation where desert is an aim, even if it is not the sole aim. It still needs to be weighed up and may often not be outweighed.

Further, my arguments are important for censure theorists, whose theories are intended to constitute complete accounts of punishment. For theorists such as Duff, von Hirsch and Ashworth my arguments at least give them something to refute, if not consider adopting.

Finally, the fact that there are theories of punishment and sentencing that are not censure-based should not necessarily be a worry. Even the fact that my arguments fit less simply into a penance perspective – is censure or apology the primary communicative function of punishment? – should not be too troubling. One can only accept one theory of punishment (even if this theory is somewhat internally pluralistic, such as von Hirsch's – incorporating the additional prudential disincentive supplied by punishment renders him not an out-and-out retributivist). Rather than having to first adopt a particular censure theory in order to assess my arguments, the direction of support could go the other way. If my arguments for the mitigating role of remorse convince (and if it is agreed that those I ruled out were ruled out correctly), then they might actually serve to support adherence to a compatible theory, such as that of von Hirsch and Ashworth. A theory that cannot account for the mitigating role of remorse, it might be argued, is perhaps deficient for at least this reason.

Bibliography

- ARISTOTLE (2007), *The Nicomachean Ethics*, Filiquarian Publishing, LLC.
- ARONSON, E. (1992), 'The Return of the Repressed: Dissonance Theory Makes a Comeback', *Psychological Inquiry*, 3, 4: 303-11.
- ASHWORTH, A. (1994), 'Justifying the Grounds of Mitigation', *Criminal Justice Ethics*, 13, 1: 5-10.
- _____. (2010), *Sentencing and Criminal Justice*, 5th edn, Cambridge: Cambridge University Press.
- BAGARIC, M., and AMARASEKARA, K. (2001), 'Feeling sorry? Tell Someone Who Cares: The Irrelevance of Remorse in Sentencing', *The Howard Journal*, 40: 364–76.
- BAUMEISTER, R. F., STILLWELL, A. M., and HEATHERTON, T. F. (1994), 'Guilt: An Interpersonal Approach', *Psychological Bulletin*, 115: 243-67.
- BAYERN, S. J. (2009), 'The Significance of Private Burdens and Lost Benefits for a Fair-Play Analysis of Punishment', *New Criminal Law Review*, 12, 1: 1-42.
- BEEMAN, W. O. (1985), 'Dimensions of Dysphoria: The View from Linguistic Anthropology', in A. Kleinman and B. Good (Eds.), *Culture and Depression*, Berkeley: University of California Press.
- BENNETT C. (2008), *The Apology Ritual: A Philosophical Theory of Punishment*, Cambridge University Press.
- BENTHAM, J. (1970), 'An Introduction to the Punishment of Morals and Legislation', reprinted in J. H. Burns and H. L. A. Hart (eds), *The Collected Works of Jeremy Bentham*, Athlone Press: University of London.
- BITTNER, R. (1992), 'Is it Reasonable to Regret the Things One Did?', *The Journal of Philosophy*, 89: 262-73.
- BLUM, L. A. (1980), *Friendship, Altruism and Morality*, London: Routledge.
- BROWN, R. (1958), *Words and Things*, New York: Free Press of Glencoe.
- BROWNLEE, K (2011) 'The Offender's Part in the Dialogue', in R. Cruft, M. Kramer and M. Reiff (eds), *Crime, Punishment, and Responsibility: The Jurisprudence of Antony Duff*, Oxford University Press.
- BRIGHT, W. (1957), *The Karok Language (University of California Publications in Linguistics, Vol. 13)*, Berkeley: University of California Press.
- BUBER, M. (1970), *I and Thou*, Charles Scribner's Sons: New York.

- _____. (1985), *Between Man and Man*, Collier Books: New York.
- BURGH, R. W. (1982), 'Do the Guilty Deserve Punishment?', *The Journal of Philosophy*, 79, 4 193-210.
- BYRD, B. S. (1989), 'Kant's Theory of Punishment: Deterrence in Its Threat, Retribution in Its Execution', *Law and Philosophy*, 8, 2: 151-200.
- CALHOUN, C. (2004), 'An Apology for Moral Shame', *The Journal of Political Philosophy*, 12, 2: 127-146.
- CARD, C. (2002), *The Atrocity Paradigm: A Theory of Evil*, New York: Oxford University Press.
- CASTELFRANCHI, C., and POGGI, I. (1990), 'Blushing as a Discourse: Was Darwin Wrong?', in W. R. Crozier (Ed.), *Shyness and Embarrassment: Perspectives from Social Psychology*, New York: Cambridge University Press.
- CORDNER, C. (2007), 'Guilt, Remorse and Victims', *Philosophical Investigations*, 30: 337-62.
- COX, M. (1999), 'Remorse and Reparation: 'To double business bound'', In M. Cox (ed.), *Remorse and Reparation*, London: Jessica Kingsley Publishers.
- CROZIER, W. R. (1998), 'Self-consciousness in Shame: The Role of the 'Other'', *Journal for the Theory of Social Behaviour*, 28, 3: 275-286.
- DAN-COHEN, (2007), 'Revising the Past: On the Metaphysics of Repentance, Forgiveness and Pardon', in A. Sarat and N. Hussain (eds.), *Forgiveness, Mercy, and Clemency*, Stanford University Press.
- D'ARMS, J. and JACOBSON, D. (2000), 'The Moralistic Fallacy: On the 'Appropriateness' of Emotions', *Philosophy and Phenomenological Research*, 61: 65-90
- DAVITZ, J. R. (1969), *The Language of Emotion*, New York: Academic Press.
- DEIGH, J. (1994), 'Cognitivism in the Theory of Emotions', *Ethics*, 104: 824-54.
- _____. (1996), *The Sources of Moral Agency: Essays in Moral Psychology and Freudian Theory*, Cambridge University Press.
- DEWEY, J. (1927), *The Public and its Problems*, New York: Holt.
- DE WIJZE, S. (2004), 'Tragic-Remorse: The Anguish of Dirty Hands', *Ethical Theory and Moral Practice*, 7: 453-71.
- DICKENS, C. (2003), *David Copperfield*, Spark Educational Publishing.

DILMAN, I. (1999), 'Shame, Guilt and Remorse', *Philosophical Investigations*, 22: 312-29.

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

_____. (1988), 'Punishment and Penance – A Reply to Harrison', in 'R. Harrison and R. A. Duff, 'Punishment and Crime', *Proceedings of the Aristotelian Society, Supplementary Volumes*, 62: 139-151+153-167

_____. (1996), 'Penal Communications: Recent Work in the Philosophy of Punishment', *Crime and Justice*, 20: 1-97.

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

EASTON, S. and PIPER, C. (2008), *Sentencing and Punishment: The Quest for Justice*, 2nd edn, Oxford University Press.

EKMAN, P. (1992), 'An Argument for Basic Emotions', *Cognition and Emotion*, 6, 3: 169-200.

FEINBERG, J. (1970), *Doing and Deserving: Essays in the Theory of Responsibility*, Princeton University Press.

_____. (1983), 'Autonomy, Sovereignty and Privacy: Moral Ideas in the Constitution?', *Notre Dame Law Review*, 58: 445-492.

FESTINGER, L. (1957), *A Theory of Cognitive Dissonance*, Evanston, IL: Row, Peterson.

FINCKENAUER, J. O. (1988), 'Public Support for the Death Penalty: Retribution as Just Deserts or retribution as Revenge?', *Justice Quarterly*, 5, 1: 81-100.

FINNIS, J. (1972), 'The Restoration of Retribution', *Analysis*, 32, 4: 131-135.

_____. (1999), 'Retribution: Punishment's Formative Aim', *The American Journal of Jurisprudence*, 44, 91-103.

GAITA, R. (1991), *Good and Evil: an Absolute Conception*, St. Martin's Press.

_____. (1999), *A Common Humanity: Thinking About Love and truth and Justice*, Text Publishing: Melbourne Australia.

GARDNER, J. (2005), 'Wrongs and Faults', *The Review of Metaphysics*, 59 (1): 95-132

GARVEY, S. P. (1996), 'As the Gentle Rain from Heaven: Mercy in Capital Sentencing', *Cornell Law Review*, 81: 989-1048.

GERBER, E. (1975), *The Cultural Patterning of Emotions in Samoa*, Unpublished doctoral dissertation, University of California, San Diego.

- GLOVICH, T. and MEDVEC, V. H. (1995), 'The Experience of Regret: What, When, and Why', *Psychological Review*, 102: 379-95.
- GOBODO-MADIKIZELA, P. (2002), 'Remorse, Forgiveness, and Rehumanization: Stories from South Africa', *Journal of Humanistic Psychology*; 42: 7-32.
- GORELICK, K. (1989), 'Wielder of Many Swords: remorse and its Transmutations', *Psychotherapy Patient*, 5, 219-34.
- HAMPTON, J. (1988), 'Forgiveness, Resentment and Hatred', in J. G. Murphy and J. Hampton, *Forgiveness and Mercy*, Cambridge University Press.
- HANNA, N. (2009), 'The Passions of Punishment', *Pacific Philosophical Quarterly*, 90: 232–250.
- HARRÉ, R. (1990), 'Embarrassment: A Conceptual Analysis', in W.R. Crozier (Ed.), *Shyness and Embarrassment: Perspectives from Social Psychology*, New York: Cambridge University Press.
- HART, H. L. A. (1959), 'The Presidential Address: Prolegomenon to the Principles of Punishment', *Proceedings of the Aristotelian Society, New Series*, 60: 1-26.
- _____. (1968), *Punishment and Responsibility*, Oxford: Clarendon Press.
- HARTMANN, P. (1984), 'Factors in Aggravation and Mitigation: A Trap for the Sentencing Judge?', *DePaul Law Review*, 33: 357-370.
- HAUCK, P. A. (1989), 'The Remorseful Patient', *Psychotherapy Patient*, 5: 13-18.
- HIATT, L. R. (1978), 'Classification of the Emotions', in L. R. Hiatt (Ed.), *Australian Aboriginal Concepts*, Princeton, NJ: Humanities Press.
- HOEBER, P. R. (1986), 'The Abandonment Defense to Criminal Attempt and Other Problems of Temporal Individuation', *California Law Review*, 74, 2: 377-427.
- HOYT, M. F. (1983), 'Concerning Remorse: With Special Attention to its Defensive Function', *Journal of the American Academy of Psychoanalysis*, 11: 435-44.
- HUPKA, R. B., LENTON, A. P., and HUTCHISON, K. A. (1999), 'Universal Development of Emotion Categories in Natural Language', *Journal of Personality and Social Psychology*, 77, 2 : 247-278.
- JACOBSON, J. and HOUGH, M. (2007), *Mitigation: The Role of Personal Mitigating Factors in Sentencing*, London: The Prison reform Trust.
- JOHANNESSEN, R. L. (1971), 'The Emerging Concept of Communication as Dialogue', *Quarterly Journal of Speech*, 57, 4: 373-382.
- JOHNSON, J. (2008), 'Revisiting Kantian Retributivism to Construct a Justification of Punishment', *Criminal Law and Philosophy*, 2, 3: 291-307.

- JONES, H. (1956), *Crime and the Penal System*, University Tutorial Press.
- KAHEMAN D., and TVERSKY, A. (1982), 'The Simulation Heuristic', in D. Kaheman, P. Slovic, and A. Tversky (eds), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press.
- KANT, I. (2009), 'The Retributive Theory of Punishment' in J. E. White (ed) *Contemporary Moral Problems*, Cengage Learning.
- KEELER, W (1983), 'Shame and Stage Fright in Java', *Ethos*, 11:152-165.
- KLEINIG, J. (1984), *Paternalism*, Manchester University Press.
- KOLBER, A. J. (2009), 'The Subjective Experience of Punishment', *Columbia Law Review*, 109: 183-236.
- LANDMAN, J. (1993), *Regret: The Persistence of the Possible*, Oxford University Press.
- LEBRA, T. S. (1983), 'Shame and Guilt: A Psychocultural View of the Japanese Self', *Ethos*, 11: 192-231.
- LEVINAS, E (1991), *Otherwise than Being or Beyond Essence*, Dordrecht: Kluwer Academic Publishers.
- LEVY, R. I. (1973), *Tahilians*, Chicago: University of Chicago Press.
- LEWIS, H. B. (1971), *Shame and Guilt in Neurosis*, New York: International Universities Press.
- LINDSAY-HARTZ, J. (1984), 'Contrasting Experiences of Shame and Guilt', *American Behavioral Scientist*, 27: 689-704.
- LUTZ, C. (1980), *Emotion Words and Emotional Development on Ifaluk Atoll*, Unpublished doctoral dissertation, Harvard University.
- LYND, H. (1999), *Shame and the Search for Identity*, Routledge.
- MADELL, G. (1981), *The Identity of the Self*, Edinburgh University Press.
- MANSON, A. (2011), 'The Search for Principles of Mitigation: Integrating Cultural Demands', in J. Roberts (ed.), *Mitigation and Aggravation at Sentencing*, Cambridge University Press.
- MARKEL, D. FLANDERS, C., and GRAY, D. C. (2011), 'Beyond Experience: Getting Retributive Justice Right', *California Law Review*, 99: 605-28.
- MATZA, D. (1990), *Delinquency and Drift*, Transaction Publishers.
- MCEWAN, I. (2003), *Atonement*, New York: Anchor Books.

METZ, T. (2007), 'How to Reconcile Liberal Politics with Retributive Punishment', *Oxford Journal of Legal Studies*, 27, 4: 683–705.

MEYER, L. R. (2005), 'Eternal Remorse', in A. Sarat (ed.), *Studies in Law, Politics, and Society: Special Volume – Toward a Critique of Guilt: Perspectives from Law and the Humanities*, vol 36, Elsevier.

MORGAN, M. L., (1992), *Classics of Moral and Political Theory*, Hackett Pub. Co.

MORICE, R. (1978), Psychiatric Diagnosis in a Transcultural Setting: The Importance of Lexical Categories, *British Journal of Psychiatry*, 132: 87-95.

MORRIS, H. (1968), 'Punishment and Persons', *The Monist*, 52, 4: 475 – 501.

_____. (1971), 'Guilt and Suffering', *Philosophy East and West*, 21: 419-34.

_____. (1976), *On Guilt and Innocence: Essays in Legal Philosophy and Moral Psychology*, Berkley, CA: University of California Press.

_____. (1981), 'Reflections on Feeling Guilty', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 40: 187-93.

_____. (1987), 'Nonmoral Guilt', in F. Schoeman (ed), *Responsibility, Character and the Emotions*, Cambridge University Press.

_____. (1999), 'Some Further Reflections on Guilt and Punishment', *Law and Philosophy*, 18: 363-78.

MURPHY, J. G. (1979), *Retribution, Justice and Therapy*, Springer.

_____. (1999), 'Shame Creeps Through Guilt and Feels Like Retribution', *Law and Philosophy*, 18: 327-44.

_____. (1987), 'Does Kant have a Theory of Punishment?', *Columbia Law Review*, 87, 3: 509-532.

_____. (1988a), 'Hatred: A Qualified Defence', in J. G. Murphy and J. Hampton, *Forgiveness and Mercy*, Cambridge University Press.

_____. (1988b), 'Forgiveness and Resentment', in J. G. Murphy and J. Hampton, *Forgiveness and Mercy*, Cambridge University Press.

_____. (1988c), 'Mercy and Legal Justice', in J. G. Murphy and J. Hampton, *Forgiveness and Mercy*, Cambridge University Press.

_____. (2003), *Getting Even: Forgiveness and its Limits*, Oxford University Press.

_____. (2006), 'Well Excuse Me! – Remorse, Apology, and Criminal Sentencing', *Arizona State Law Journal*, 38: 371-86.

NATHANSON, D. L., (1992), *Shame and Pride: Affect, Sex, and the Birth of the Self*, New York: W.W. Norton.

NIEDENTHAL, P. M., TANGNEY, J. P., and GAVANSKI, I. (1994), "'If Only I Weren't' Versus 'If Only I Hadn't': Distinguishing Shame and guilt in Counterfactual Thinking", *Journal of Personality and Social Psychology*, 67: 585-95.

NORMAN, R. (1971), *Reasons for actions: A Critique of Utilitarian Rationality*, Oxford: Basil Blackwell.

NUSSBAUM, M. C. (1990), *Love's Knowledge*, Oxford University Press.

NUSSBAUM, M. (2001), *Upheavals of Thought: The Intelligence of Emotions*, Cambridge: Cambridge University Press.

OBEYESEKERE, G. (1981), *Medusa's Hair: An Essay on Personal Symbols and Religious Experience*, Chicago: University of Chicago Press.

OHBUCHI, K., AGARIE, N., and KAMEDA, M. (1989). 'Apology as Aggression Control: Its Role in Mediating Appraisal of and Response to Harm', *Journal of Personality and Social Psychology*, 56: 219–27.

PARFIT, D. (1986), *Reasons and Persons*, Oxford University Press.

PERRY, J. (1978), *A Dialogue on Personal Identity and Immortality*, Hackett Publishing.

RADDEN, J. (1995), 'Shame and Blame: The Self through Time and Change', *Dialogue*, XXXIV: 61-74.

ROBERTS, J.V., HOUGH, M., JACOBSON, J., and MOON, N. (2009), 'Public Attitudes to Sentencing Purposes and Sentencing Factors: An Empirical Analysis', *Criminal Law Review*, 11: 771-82.

ROBERTS, R. (1991), 'What is Wrong with Wicked feelings?', *American Philosophical Quarterly*, 28: 13-24.

_____. (2003), *Emotions: An Essay in Aid of Moral Psychology*, Cambridge University Press.

ROBINS, R. H. (1958). *The Yurok Language: Grammar, Texts, Lexicon (University of California Publications in Linguistics, Vol. 15)*, Berkeley: University of California Press.

ROBINSON, P., and DARLEY, J. (1995), *Justice, Liability, and Blame: Community Views and the Criminal Law*, Bolder, CO: Westview Press.

RORTY, A. O. (1980), 'Agent Regret' in A. O. Rorty (ed), *Explaining Emotions*, University of California Press.

- ROSALDO, M. Z. (1983), The shame of Headhunters and the Autonomy of Self, *Ethos*, 11: 135-151.
- ROSEMAN, I. J., WIEST, C., and SWARTZ, T. S. (1994), 'Phenomenology, Behaviours, and Goals Differentiate Discrete Emotions', *Journal of Personality and Social Psychology*, 67: 206-21.
- ROSTHAL, R. (1967), 'Moral Weakness and Remorse', *Mind*, 76: 576-79.
- RUSSELL, J. A. (1991), 'Culture and the Categorization of Emotions', *Psychological Bulletin*, 110: 426-450.
- SCANLON, T. (2008), *Moral Dimensions: Permissibility, Meaning, Blame*, Harvard University Press.
- SCHEFF, T. J., (1995) 'Shame and Related Emotions: An Overview', *American Behavioural Scientist*, 38: 1053-9.
- SCHIED, D. E. (1983), 'Kant's Retributivism', *Ethics*, 93, 2: 262-282.
- SCHER, S. and DARLEY, J. (1997), 'How Effective are the Things People Say to Apologize? Effects of the Realization of the Apology Speech Act', *Journal of Psycholinguistics Research*, 26: 127-40.
- SCHOPENHAUER, A. (1995), *On the Basis of Morality* (trans. E. F. J. Payne), Oxford : Berghahn Books.
- SHABAD, P. (1989), 'Remorse: The Echo of Inner truth', *Psychotherapy Patient*, 5: 113-33.
- SHAVER, P., SCHWARTZ, J., KIRSON, D., and O'CONNOR, C. (1987), 'Emotion Knowledge: Further Exploration of a Prototype Approach', *Journal of Personality and Social Psychology*, 52: 1061-86.
- SHAW, J. (1988), 'The Usefulness of Remorse', *Psychotherapy Patient*, 5: 77-82.
- SMITH, A. (2002), *The Theory of Moral Sentiments*, K. Haakonssen (ed), Cambridge University Press. (Original work published 1793).
- SOLOMON, R. C. (1976), *The Passions: The Myth and Nature of Human Emotion*, University of Notre Dame Press.
- SPINOZA, B. (2002), *Spinoza: Complete Works*, Hackett Publishing.
- STARK, S. (2004), 'Emotions and the Ontology of Moral Value', *The Journal of Value Inquiry*, 38: 355-374
- STERN, E. M. (1989), 'The Psychotherapy of Remorse', *Psychotherapy Patient*, 5: 1-11.

STEWART, H. (2010), 'The Limits of the Harm Principle', *Criminal Law and Philosophy*, 4: 17-35.

STOCKER, M. (1996), *Valuing Emotions*, Cambridge University Press.

SUNDBY, S. (1998), 'The Capital Jury and Absolution: The Intersection of Trial Strategy, Remorse, and the Death Penalty', *Cornell Law Review*, 83: 1557-98.

TANGEY, J. P. (1990), 'Sharing Shame and Guilt: Another Social Clinical Interface', *Contemporary Social Psychology*, 14: 83-8.

_____. (1991), 'Moral Affect: The Good, the Bad, and the Ugly', *Journal of Personality and Social Psychology*, 598-607.

_____. (1995), 'Recent advances in the Empirical Study of Shame and Guilt', *American Behavioral Scientist*, 38: 1132-45.

_____. (1999), 'The Self-Conscious Emotions: Shame, Guilt, Embarrassment and Pride', in T. Dalgleish and M. Power (Eds.), *Handbook of Cognition and Emotion*, London: Wiley.

_____, Miller, R. S., Flicker, L., and Barlow, D. H. (1996), 'Are Shame, Guilt and Embarrassment Distinct Emotions?', *Journal of Personality and Social Psychology*, 70: 1256-69.

_____, Wagner, P., Fletcher, C., and Gramzow, R. (1992), 'Shamed into Anger: The Relation of Shame and Guilt to Anger and Self-Reported Aggression.

TASIOULAS, J. (2007), 'Repentance and the Liberal State', *Ohio State Journal of Criminal Law*, 4: 487-521.

_____. (2006), 'Punishment and Repentance', *Philosophy*, 81: 279-322.

_____. (2011) 'Where is the Love? - The Topography of Mercy' in R. Cruft, M. Kramer and M. Reiff (eds), *Crime, Punishment, and Responsibility: The Jurisprudence of Antony Duff*, Oxford University Press.

TAYLOR, C. (1985), *Human Agency and language: Philosophical Papers 1*, Cambridge University Press.

TAYLOR, G. (1985), *Pride, Shame and Guilt: Emotions of Self-Assessment*, Oxford University Press.

TEICHMAN, J. (1973), 'Punishment and Remorse', *Philosophy*, 48: 335-46.

TEN BRINKE L., MACDONALD, S., PORTER, S. and O'CONNOR, B. (2011), 'Crocodile Tears: Facial, Verbal and Body Language Behaviours Associated with Genuine and Fabricated Remorse', *Law and Human Behavior*, Online First.

THALBERG, I. (1963), 'Remorse', *Mind*, 72: 545-55.

- THOMAS, A. (1999), 'Remorse and Reparation: A Philosophical Analysis', in M. Cox (Ed), *Remorse and Reparation*, London: Jessica Kingsley.
- THOMPSON, A. (1989), 'Emotional Origins of Morality – A Sketch', *Journal of Moral Education*, 18: 199-207.
- TOMBS, J. and JAGGER, E. (2006), 'Denying Responsibility: Sentencers' Accounts of their Decisions to Imprison', *British Journal of Criminology*, 46: 803-21.
- TONRY, M. (1996). *Sentencing matters*. Oxford: Oxford University Press.
- TUDOR, S. (2001), *Compassion and Remorse*, Peeters.
- _____. (2008), 'Why Should Remorse Be a Mitigating Factor in Sentencing?', *Criminal Law and Philosophy*, 2, 3: 241-257.
- TUNICK, M. (1996), 'Is Kant a Retributivist?', *History of Political Thought*, 17, 1: 60-78.
- VON HIRSCH, A. (1986), *Past or Future Crimes*, Manchester University Press.
- _____. (1991) 'Criminal Record Rides Again', *Criminal Justice Ethics*, 10: 2, 55-56.
- _____. (1993), *Censure and Sanctions*, Oxford: Clarendon Press.
- _____, and ASHWORTH, A. (2005), *Proportionate Sentencing: Exploring the Principles*, Oxford University Press.
- _____, and JAREBORG, N. (1991), 'Gauging Criminal Harm: A Living-Standard Analysis', *Oxford Journal of Legal Studies*, 11, 1 1-38.
- WALKER, N. (1999), *Aggravation, Mitigation, and Mercy in English Criminal Justice*, Oxford University Press.
- WARD, B. H. (2006), 'Sentencing Without Remorse', *Loyola University Chicago Law Journal*, 38: 131-67.
- WASIK, M. (1982), 'Partial Excuses in the Criminal Law', *The Modern Law Review*, 45, 5: 516-533.
- WEINER, B. (1992), 'Attributional Theories of Motivation: Persons as Scientists', in B. Weiner (ed.), *Human motivation: Metaphors, Theories and Research*, Newbury Park, CA: Sage
- WIGGINS, D. (1991), *Needs Values and Truth*, Oxford: Blackwell.
- WILLIAMS, B. (1973), *Problems of the self*, Cambridge University Press.

_____. (1981), *Moral Luck: Philosophical Papers 1973-1980*, Cambridge University Press.

_____. (1993), *Shame and Necessity*, University of California Press.

WILSON, A. N. (1990), *Incline our Hearts*, London: Penguin Books.

ZEELLENBERG, M., VAN Dijk, W. W., MANSTEAD, A. S. R., and VAN DER PLIGHT, J. (1998), 'The Experience of Regret and Disappointment', *Cognition and Emotion*, 12: 221-30.