

**MORAL REFORM AND THE DESIDERATA OF RESPONSES TO  
WRONGDOING: the production of a “morally autonomous person freely attached  
to the good”.**

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

of

RHODES UNIVERSITY

by

**HEATH FREDERICK WALLER**

February 2004

## **Abstract**

Moral reform is a neglected response to wrongdoing that has been incorrectly portrayed as a practice involving illegitimate treatment of wrongdoers and as totally unsatisfying to those theorists advocating backward-looking practices such as retributive punishment. A clear explanation of the ethical legitimacy and practical necessity of the reformative techniques moral reform involves has been missed, and this paper details the design of moral reform proper in order to fill this gap in punishment theory. The moral reform of an offender is identified as a desideratum of responses to wrongdoing and it is explained what moral reform ought to entail.

The claim that moral reform qualifies as the overriding aim of responses to wrongdoing is argued for on the grounds that this practice is capable of achieving all the established ends of responses to wrongdoing. The legitimate desiderata of our practices are identified as those usually selected as the ends of punishment practices, and moral reform must accomplish these if it is to be accepted. Moral reform is shown to realise the goals of punishments as the fortunate effects of what is done to achieve an offender's moral improvement and of what reformees do in taking responsibility for their actions. The suffering involved in moral reform receives particular emphasis since the practice will never satisfy unless it accommodates the widely-held intuition that the offender must suffer sufficiently as a consequence of his wrongdoing.

Moral reform is further portrayed as the most meaningful practice for its ability to satisfy the appropriate needs and desires victims have in response to their victimization. A central claim of the thesis is that moral reform best serves the victim, since it most effectively relieves the victim's emotional responses to wrongdoing and is as adept as punishment at the expression of these same emotions. Reformers advocate a constructive response to wrongdoing that benefits all affected parties.

## TABLE OF CONTENTS

<b>Acknowledgements</b>	vi
<b>Introduction</b>	p. 1
<b>Chapter One: Three stages in moral reform.</b>	
§1. Introduction	p. 6
§2. Moral reform is a desideratum of responses to wrongdoing	p. 6
§3. What is moral reform?	p. 8
§4. The Kantian maxim	p. 10
§4.1. Moral reform is not an “influenceability theory”	p. 11
§5. Moral reform is a three-stage process	p. 13
§5.1. Stage One	p. 13
§5.1.1. Coercing wrongdoers	p. 14
§5.1.2. Inflicted suffering is a reformatory technique	p. 16
§5.1.3. Reformatory punishment theories	p. 17
§5.1.4. Punishments tailored to reform offenders	p. 22
§5.1.5. Moral reform begins with reformatory punishment	p. 24
§5.2. Stage Two	p. 25
§5.2.1. Only morally responsible wrongdoers are reformed	p. 26
§5.2.2. Guilt, shame and remorse - the “reformatory emotions”	p. 29
§5.2.3. Emotional suffering is not reformatory punishment	p. 32
§5.3. Stage Three	p. 34
§5.3.1. What is atonement?	p. 35
§5.3.2. Atonement is not punishment	p. 37
§6. Conclusion	p. 38

## **Chapter Two: Expressing and relieving reactive attitudes.**

§1.	Introduction	p. 41
§2.	Emotions inform our practices	p. 42
§2.1.	Strawson on why the reactive attitudes ought to be expressed	p. 43
§2.2.	How the reactive attitudes are expressed	p. 46
§2.3.	Why the reactive attitudes ought to be relieved	p. 48
§2.4.	How the reactive attitudes are relieved	p. 49
§2.4.1.	What do victims, their relations and society want?	p. 50
§2.4.2.	The so-called retributive emotions	p. 51
§2.4.3.	The relief of the so-called retributive emotions	p. 54
§3.	Conclusion	p. 62

## **Chapter Three: Desert and deterrence.**

§1.	Introduction	p. 64
§2.1.	Desert and desert bases	p. 64
§2.2.	Retributive ideas are indelible	p. 67
§2.3.	Reformees pay their debts	p. 68
§2.4.	Reformees relinquish unfair benefits	p. 69
§2.5.	Reformees do not deserve more suffering	p. 70
§3.	Deterring potential wrongdoers	p. 71
§4.	Conclusion	p. 72

## **Chapter Four: Three objections to moral reform.**

§1.	Introduction	p. 75
§2.	The indeterminate sentencing problem	p. 76
§2.1.	Internal sentencing principles	p. 76
§2.2.	External moral constraints	p. 78
§2.3.	When moral reform fails	p. 80
§3.	The unreformable	p. 81

§4.	Moral reform does too much to benefit the wrongdoer	p. 84
§5.	Conclusion	p. 87
<b>Conclusion</b>		p. 88
<b>Bibliography</b>		p. 91

## **Acknowledgements**

Many thanks to my supervisor, Dr Ward E. Jones of the Rhodes University Philosophy Department. The time and effort that he spent discussing my thesis with me is greatly appreciated. Special thanks to Samantha Vice for her invaluable assistance and constant encouragement. Thank you Francis Williamson and the Rhodes University Philosophy Department for being listeners and advisors. I am especially grateful to Rhodes University for granting me a postgraduate scholarship without which I would have been unable to pursue my research.

## Introduction

The literature on responses to wrongdoing creates the impression that our primary concern is to see the wrongdoer suffer, and the job of philosophers is then to justify the fulfillment of this demand. This summation gains some plausibility from the fact that the intuition that the offender must suffer for his wrongdoing is so deeply entrenched. We cannot find satisfaction in a response to wrongdoing that does not involve for the offender a sufficient measure of suffering. Punishments are a popular sort of practice because they guarantee the suffering of the offender by inflicting it. Other goals of responses to wrongdoing are apparently subordinate to this universal demand and must be achieved by means of inflicted suffering (punishment) alone. For instance, reformatory punishment is a practice that aims to reform offenders yet relies solely upon inflicted suffering to accomplish this end, thereby omitting myriad other reformatory techniques (psychoanalysis, rational debate) that improve the odds of success.

Reformatory techniques are needed because offenders struggle to *take responsibility* for the wrongs they are guilty of. To take responsibility for an action is to be moved to redress one's wrongdoing by an understanding of one's culpability for a *wrong* done. It is a difficult task demanding humility and sacrifice. Responsible moral agents possess the psychological capacities to avoid thinking ill of themselves or their actions, and these mechanisms of casuistry must be counteracted if the wrongdoer is to adequately appreciate her culpability for an immoral act. Moral reform involves treatments acting as catalysts for the process of self-examination through which responsible moral agents come to take responsibility for their own actions.

From a young age we are taught that our best response to our own wrongdoings is to accept responsibility for them and pay the price associated with learning from our mistakes. Mature human beings take responsibility for their actions because they are moral agents with a capacity for self-reflection, an ability to evaluate the quality of their own will, identify its failings, and effect change for the better. What we do out of conscience comes from the source of our ability to morally value anything, and it is this response from the conscience of the wrongdoer that moral reform encourages. Of course, moral reform does not aim only to produce an offender who is sorry. A reformee is a morally improved citizen with an understanding of the role of morality in his life and,

consequently, a determination not to transgress again. Wrongdoers emerge from moral reform having acquired commitment to an upright existence, and society benefits from a citizenry who are less inclined to perpetrate immoral actions.

With their attention focused on securing the suffering of the offender, punishment theorists have found it difficult to portray their practices as constructive responses addressing the suffering of the victim. The harm suffered by the victim is, after all, the reason a response is called for, and it is quite fitting that responses to wrongdoing be tailored to redress the victim's discomfort. It is not the case that all victims want or need from a response to wrongdoing is to see the offender have suffering inflicted upon her. The goal of the reform theorist is to develop a practice that constructively sets attention on the needs and desires a suffering victim may have in response to her victimization. For instance, victims demand apology or some meaningful form of recognition from their attackers, something, I argue, punishment does not reliably motivate an offender to do. Apology and sincere recognition come from remorseful offenders who are motivated to atone for their wrongdoings by an appreciation of the immorality of their conduct. Such people also understand that they need to undertake to improve the quality of their will, and moral reform programmes provide the assistance offenders require to accomplish these challenging tasks.

Moral reform is a practice prioritizing the victim's relief, but does not do so to the exclusion of what are commonly identified as the desiderata of responses to wrongdoing. The practice aims to satisfy the victim, and victims want the offender to suffer, they want 'debts to society' to be repaid, they want moral reform, they desire that potential offenders be deterred, and they desire the expression of those emotions they feel in response to wrongdoing. In this way, the ambition of moral reform becomes the achievement of all the desirable goals of a response to wrongdoing, including those ends traditionally identified with punishment practices, through the production of reformees who are motivated to take responsibility for their own actions.

Chapter One is concerned with distinguishing moral reform from punishment practices, particularly reformatory punishment, and explaining why a practice that is not a punishment practice ought to be pursued at all. This is essentially a motivation for the inclusion of those 'extra' treatments and techniques (psychological treatments, rational

debate) that identify full moral reform, and an explanation why inflicting suffering on offenders is not enough to effect their proper and lasting moral improvement. The chapter describes the defining characteristics of a reformee and the three stages of 'producing' a reformee.

In Chapter Two the emotional reactions we have in response to wrongdoing are exploited for their invaluable guidance concerning what we ought to do to wrongdoers. It is a neglected question what victims, their relations and society truly desire from a response to wrongdoing, and emotions prove excellent informants. There are two parts to the chapter, the first concerning the expression of emotions through inflicted suffering. That our practices ought to express valued emotions is an assertion motivated for by P. F. Strawson's emotion-based account of meaningful moral interaction,<sup>1</sup> and moral reform is shown to include inflicted suffering enough for this purpose.

The second argument in Chapter Two is that we can discover what victims, their relations and society want from a response to wrongdoing by examining what relieves those morally valuable emotions we feel in response to an offence. In favour of moral reform, I explain that the so-called "retributive emotions" - those emotions typically identified with punishment (resentment, indignation, moral hatred) - are better relieved by moral reform than by punishment. Moral reform is thus better able than punishment to deliver what our emotions tell us we want from a response to wrongdoing, and thus presents itself as a more meaningful practice.

Analysis of the emotions does not provide an exhaustive list of the desiderata of responses to wrongdoing. Chapter Three describes how moral reform is capable of achieving further recognised ends of responses to wrongdoing, in particular the goals of retributive and deterrence punishment.<sup>2</sup> Moral reform is able to deter, see that offenders suffer as much as deserved, repay moral 'debts' and re-establish the balance of 'benefits and burdens' in the moral economy, by producing remorseful reformees who undertake atoning actions. These goals of punishment are accomplished by the modes of suffering involved in moral reform, which are identified in Chapter One and feature throughout the paper as a strong argument for punishment theorists to see the value in moral reform.

---

<sup>1</sup> Peter Strawson, 'Freedom and Resentment', 1962, pp. 59-80.

<sup>2</sup> Retributive punishment is punishment justified on the basis of meting out just deserts. Deterrence punishment justifies the infliction of suffering on offenders on the basis that it deters potential wrongdoers.

The final chapter addresses three central objections to moral reform. Not all worries faced by reform theorists can be addressed here, and I have selected three core objections for their specific impact on claims that are central features of the practice described in Chapter One. The ‘indeterminate sentencing’ problem argues that moral reform permits more than the amount of suffering required to satisfy the basic demand for the offender to suffer, and thus may involve too much suffering. This casts doubt on whether moral reform involves an amount of suffering that is deserved, or constitutes a debt repayment, or which is just.

The second problem concerns the ‘unreformable’, people who cannot respond to reformatory techniques and so are not legitimate subjects for moral reform, yet who may be morally responsible wrongdoers whom we would see punished. In question is whether moral reform targets only those offenders who have the cognitive and affective capacities of morally responsible persons. Finally, the objection that moral reform inappropriately benefits the offender presents an opportunity to highlight how the production of reformees most benefits the victims, their relations and society and accomplishes all the recognised goals of responses to wrongdoing. The idea that the best way to help the victim is to help the wrongdoer is the core theme of the paper.

## **Chapter One**

### **Three stages in moral reform.**

## **§1. Introduction**

Moral reform is distinct from other responses to wrongdoing for its use of reformatory techniques, including inflicted suffering, psychological treatments and rational debate, and insistence on the offender's eventual atonement. The design of the practice begins with its goal, the reformee, whose emotional and cognitive condition is deliberately influenced by morally acceptable reformatory techniques. This chapter describes who a reformee is and how moral reform is a practice producing reformees through appeal to the cognitive and affective capacities of morally responsible offenders. Moral reform is in competition with 'reformatory punishment', a practice which, I argue, is unable to match moral reform proper in influencing wrongdoers to take responsibility for their conduct and improve the quality of their will. Throughout this chapter the effects on the wrongdoer that are particular to moral reform are identified and their value to the victim, her relations and society is explained. Though it may not be apparent at first, moral reform is dedicated to the victim's satisfaction. Chapter One offers the first arguments in what is a gradual defense of this claim. Victims want the moral reform of offenders and Chapter One explains why and how moral reform best achieves this desideratum of responses to wrongdoing.

## **§2. Moral reform is a desideratum of responses to wrongdoing**

'Forward-looking' accounts of the right responses to wrongdoing, those which are directed towards influencing the offender's future conduct (reform and deterrence theories), have tended to be based on utilitarian or consequentialist ethics. I shall not base my theory of moral reform on any particular ethical approach, though it remains forward-looking. Where reference is made to particular ethical theories this should not be read as an endorsement of that ethical approach as the way reform ought to be justified. An avoidance of broad ethical commitments means that I merely describe a practice that competes with other responses to wrongdoing without addressing any concern that moral reform be consistent with a larger theoretical edifice. Moral reform satisfies the needs and desires the victim, his relations, society and even the offender

ought to have, and it is these effects of reform that recommend the practice to us. In place of forging a commitment to a particular ethical view, I rely on an empirical count of how victims react to captured offenders who do not become morally reformed.

Victims, their relations and society ordinarily find it very unsatisfying if the moral reform of a wrongdoer is not accomplished by our response to his offence. Most forms of punishment are not designed to affect the moral character of wrongdoers in positive ways, but merely to mete out an appropriate amount of suffering as 'just deserts', a deterrent measure, fair play, and so on. A convict may emerge from prison having met his just deserts, yet display no sign that he appreciates he has been punished for doing what is *wrong*, no indication that he realizes he ought to feel remorse and be motivated toward moral self-improvement. Though an unreformed person may have undergone his punishment, victims, their relations and society cannot regard him as having established any 'distance' from his wrongdoing because his present character is readily identified with his past transgressions.

When an offender has not properly 'learned his lesson' we are left with a feeling that our response to wrongdoing has been inadequate. This feeling is not based on a concern that the unreformed offender may re-offend, though that is certainly a worry we would have. Rather, it is acutely dissatisfying to think that an offender may emerge from our treatment of him just the same, morally unimproved, and this is particularly so for the victim. A victim or the victim's relations find great value in a sincere verbal apology or some other form of moral recognition of the victim as someone whom an offender ought not to have victimized. Punishment practices which do not produce offenders who want to offer the victim what comforts they can are sorely unsatisfying, regardless of whether the offender has suffered his due or not. We want an offender to demonstrate that he has come to accept responsibility for his offence, we want the offender to recognize that he has wronged us, and these are appropriate desires that moral reform aims to satisfy.<sup>3</sup>

There is a sense that punishment that does not reform is an ineffective infliction of suffering that has merely been tolerated by the wrongdoer. When our response to wrongdoing has no effect on the offender's character we tend to identify the offender

---

<sup>3</sup> A more detailed account of what victims want from an offender is given in Chapter Two. There I describe the emotions felt by a victim towards a wrongdoer. From these emotions we learn what victims want from offenders.

with his wrongdoing; we think, “They should never have let him out”. An offender who has been punished but who is not reformed has slim chance of re-establishing relationships with the people against whom he has offended. There is no reason for these people to believe that he has changed, that he will not do it again, or that his conscience confronts him with the knowledge that he has unjustifiably harmed another human being. There is no reason to believe that there has been any *effect*. If offenders are not reformed, we feel that our response to wrongdoing has failed.

### §3. What is moral reform?

Responses to wrongdoing are partly defined by the goals at which they are directed, and partly defined by the treatments used to attain those goals. The common denominator among all practices commonly considered to be punishments is that they only inflict suffering to achieve their objectives. For a response to wrongdoing to count as a punishment practice it *must* involve *only* the infliction of suffering. Any practice which employs treatments other than inflicted suffering is not a punishment practice. The distinction is necessary because there is a response to wrongdoing, namely moral reform, which includes treatments not considered to be punishments, such as psychological counseling, rational debate, and moral education. The distinction implies, and I shall argue, that there are reasons for doing *more* than inflicting suffering. Effective moral reform requires treatments that do not inflict suffering on the wrongdoer.

Apart from the treatments involved, a further identifying feature of a practice is the end towards which it is directed. Reformatory punishment employs only inflicted suffering *for the purpose* of moral reform. Moral reform is a practice employing more than inflicted suffering *for the purpose* of morally reforming wrongdoers. The significance of the distinction between these practices is that it draws attention to the fact that there is disagreement over what is required to successfully effect moral reform in an offender. Indeed, the reasons for subjecting wrongdoers to more than just reformatory punishment are a central theme of this first chapter.

My theory of moral reform has as its defining goal Herbert Morris's "morally autonomous person freely attached to the good"<sup>4</sup>. A "morally autonomous person freely attached to the good" has come to, "appreciate the nature of the evil involved for others and for oneself in one's doing wrong"<sup>5</sup>. 'Appreciation' of the immorality of wrongdoing is signaled by one's feeling remorse, shame and guilt, and moves one to, "reject the disposition to do what is wrong and commit oneself to forbearance in the future"<sup>6</sup>. Successful reform requires that a wrongdoer come to, "possess and vividly retain a conception of [herself and the victim] as individual[s] worthy of respect"<sup>7</sup>. This, coupled with the aforementioned emotions, motivates the reformed wrongdoer to perform atoning actions which communicate to the victim and society that a reformee accepts responsibility for her offence.<sup>8</sup>

Morris's description of the newly reformed "morally autonomous person freely attached to the good" rings true because he simply points out those features observed in people who have truly 'learned their lesson' - the cognitive and emotional properties we commend offenders for having in reaction to their wrongdoing and which motivate future 'rightdoing'. It is a very Kantian goal; a "morally autonomous person freely attached to the good" possesses the sort of moral autonomy that Immanuel Kant described as, "the ability to freely decide, out of respect for a moral demand, to act morally, independently of any external incentives"<sup>9</sup>. Reformees are ordinary moral agents who take responsibility for what they have perpetrated and who are motivated by the moral value of good conduct.

Though it is as yet unclear how, its being directed at producing the complex being that is the "morally autonomous person freely attached to the good" described above allows moral reform to satisfy, in order of importance, the needs and desires of the victim, the victim's relations, the community and the offender. Moreover, the virtues of various punishment practices may be found in a practice aiming to attain the goal Morris has provided. A practice producing reformees can accomplish myriad legitimate aims of

---

<sup>4</sup> Herbert Morris (1981), p. 265.

<sup>5</sup> Ibid.

<sup>6</sup> Ibid.

<sup>7</sup> Ibid.

<sup>8</sup> Ibid.

<sup>9</sup> *Penguin Dictionary of Philosophy* (1999), p. 53.

responses to wrongdoing, such as moral reform, deterrence, the expression of our emotional responses to wrongdoing, the repaying of a 'debt to society', or even the satisfaction of the widely-held intuition that the offender must suffer as a consequence of her wrongdoing.

#### §4. The Kantian maxim

Responses to wrongdoing are constrained by moral principles, we do not consider ourselves free to subject offenders to whatever treatments we wish.<sup>10</sup> Our treatment of people responsible for immoral actions ought to be respectful of the dignity of these offenders as rational, moral beings by not violating one of Kant's well-known maxims. Kant formulated a fundamental moral principle stipulating that persons or moral agents ought not to be treated as merely instrumentally valuable:

Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end.<sup>11</sup>

The worst criminals may tempt us to deny them the respect Kant's maxim demands, but Kant insists in his *Metaphysics of Morals* that, "I cannot deny all respect to even the immoral man as a man".<sup>12</sup> It is in virtue of one's having the cognitive capability to recognise the good and make rational decisions that we are endowed with an inalienable dignity. Responsible moral agents are held responsible in virtue of their reasoning capacities, and therefore it must be that responsible moral agents are to be treated in ways respectful of their special dignity. Roger Sullivan makes it clear that:

[A]ll contingent facts about individuals – and our subjective, affective relationships with them – are completely irrelevant both to their inherent value and to the respect we owe them.<sup>13</sup>

There is nothing responsible offenders can do that legitimates their being treated merely as means to some social end.<sup>14</sup>

---

<sup>10</sup> The principles determining how much offenders ought to suffer are discussed in Chapter Four.

<sup>11</sup> Quoted in Roger J. Sullivan (1989), p. 195.

<sup>12</sup> *Ibid.* p. 199.

<sup>13</sup> *Ibid.*

The Kantian injunction plays a strong role in setting the limits of acceptable treatment of wrongdoers if we accept that we ought to reject any justification of a practice permitting wrongdoers to be treated merely as means and not as ends in themselves. Generally speaking, Kant's idea that we must always act with respect for the dignity of persons as rational, self-determining moral beings appears to be a nearly ubiquitous presupposition in moral and political philosophy. I accept the principle as right, as it seems to be, and shape my theory of moral reform around it.

#### §4.1. Moral reform is not an “influenceability theory”

The goal of morally reforming offenders has been closely associated with responses to wrongdoing that do not respect the moral status of offenders. In particular, what Hilary Bok calls “influenceability theories” have encouraged negative perceptions of moral reform.<sup>15</sup> Influenceability theorists like Moritz Schlick argued that, from a utilitarian point of view, two facts about a wrongdoer form the basis of holding a person culpable for an immoral act: that she has committed a crime, and that she will respond to treatments that modify her future behaviour, where these treatments included brainwashing and other manipulative methods.<sup>16</sup> The central objection to these theories is that they violate the Kantian maxim, treating morally responsible perpetrators as mere objects of social policy and not as moral agents with ends of their own. The Kantian worry is that we may *directly* inculcate a person's good behaviour, with no room at all for *responsible* moral agency of the kind that signifies rational and emotional engagement with internal moral knowledge on the part of the subject.

The practice I am advocating is not an influenceability theory. True moral reform should produce agents who are, “freely attached to the good”,<sup>17</sup> where “freely” implies that, at some point in the process, the reformee has taken a decision to improve her conduct out of a heightened appreciation of the moral and practical reasons for doing

---

<sup>14</sup> There is a difficult debate here between utilitarian and Kantian ethics that I shall not pursue, I wish simply to make my position clear. It is my project to describe a practice of moral reform that obeys the Kantian injunction and yet still accomplishes all that can be expected from a response to wrongdoing.

<sup>15</sup> Hilary Bok (1998), p 184.

<sup>16</sup> Moritz Schlick, *Problems of Ethics*, 1939.

<sup>17</sup> Herbert Morris (1981), p. 265.

so. Moral reasons and good character are supposed to motivate us in performing right action, while moral motivations, and therefore responsible agency, seem entirely absent from the stimulus-response ‘machine’ that influenceability theories produce. A manipulative practice does not produce responsible moral agents, whereas moral reform produces *more responsible* (better behaved) moral agents. (I am using the term ‘manipulative’ in the strongest sense possible, referring only to extreme treatments like brainwashing, and not other forms of manipulation like deception).

True moral reform in fact *cannot* be an influenceability theory. The goal Morris sets as definitive of moral reform brings with it a conceptual limitation on treatments that will count as morally reformative. One cannot manipulate or brainwash someone into becoming a “morally autonomous person freely attached to the good” because the concept ‘manipulation’ is a contradiction of the concepts ‘autonomy’ and ‘freedom’. Manipulation entails conditioning a person into responding automatically to particular stimuli. Autonomy and freedom imply precisely that one is not a manipulated automaton, but that one chooses and controls one’s actions. In other words, moral reform cannot in principle achieve its goal without respecting the Kantian maxim. The respect for personhood that Kant demands in practice translates into the stimulation of the cognitive and affective capacities possessed by responsible moral agents. If one is manipulated, these capacities are overridden and one loses one’s status as a morally responsible agent.

What is missing from agents who have been conditioned into behaving in line with the good is the “appreciation” of moral wrongness and rightness that Morris saw as a success criterion for moral reform. “Appreciation” is a synonym for “taken to heart”, as Jennifer Church describes propositions that have been incorporated into one’s, “automatic, habitual reasoning” (practical and theoretical).<sup>18</sup> Church cashes out the metaphor of ‘depth’ used to describe certain beliefs, arguing that ‘deeper’ beliefs are those that operate unreflectively as premises in our reasoning.<sup>19</sup> Assuming, as seems correct, that our moral commitments can be similarly ‘deep’ or ‘taken to heart’, Church has provided a useful way of describing reformees who have properly internalised values

---

<sup>18</sup> Jennifer Church (2002), p. 366.

<sup>19</sup> Ibid.

they previously violated such that they are more committed to upholding and promoting those values. A person who has been subjected to manipulative treatments will not have taken any moral values to heart. A value 'taken to heart' has become deeply embedded in one's motivational set and, as a reason consistent with one's beliefs, is an internal reason for action because at some time one has appreciated its value. Manipulated subjects react merely to a stimulus that is not part of their internal motivational set because it is not, for them, a reason for acting. Brainwashed people act on an impulse borne of conditioning rather than understanding, they are compelled by something they do not value and so cannot have 'taken to heart'.

#### **§5. Moral reform is a three-stage process**

The practice designed to produce a "morally autonomous person freely attached to the good" has three distinct but connected stages. The first stage aims to break down an offender's resistance to even basic participation in a reform programme and communicate to the wrongdoer the wrongness of her conduct through the infliction of suffering of some sort. Stage Two involves the application of further reformatory treatments designed to produce an offender who properly appreciates that she is responsible for a harmful immoral action. This reformee will signal her appreciation by feeling those emotions (guilt, shame, remorse) connected with any moral agent's coming to accept that they are responsible for a wrong done. The third stage features atoning actions voluntarily undertaken by remorseful wrongdoers bent on communicating sincere contrition and appreciation of the gravity of the offence to the victim, her relations and society.

##### **§5.1. Stage One**

Inflicted suffering has two roles in the first stage of moral reform. Firstly, through being subjected to some form of inflicted suffering, such as incarceration, work or the removal of some 'privilege' (e.g. exercise, reading material), uncooperative wrongdoers are given incentive to actively participate in and avoid disrupting moral reform programmes. Secondly, inflicted suffering may be an effective catalyst for moral reform, igniting those

self-evaluating processes that an offender engages in in trying to make sense of the reasons why she is being subjected to moral reform.

#### §5.1.1. Coercing wrongdoers

It is certain to happen that reformers will encounter wrongdoers who are very unwilling participants. Many wrongdoers react defensively when being ‘brought to book’, and so require the infliction of certain types of emotional frustration and physical suffering to open the lines of communication between the reformer and the offender, and to clearly establish who is in charge and who is not. Not least of the inflicted suffering will be the arrest of offenders’ liberty since it is highly improbable that they will voluntarily agree to be morally reformed. Furthermore, serious offenders like rapists and mobsters are likely to prove uncooperative, aggressive, and generally resistant to any initial attempts to engage with them. Or it may be the case that an offender is trying to fool the system, pretending to be receptive to reform programmes merely to get out. In any practice dealing with wrongdoers, there are bound to be lapses of discipline and emotional outbursts that are best handled by inflicting some form of suffering.

Suffering is a better tool for the purposes of Stage One than would be a reward system in which offenders are rewarded for any decision to cooperate with reformers. I must defend the use of suffering here since one might think that a reward system, which does not involve imposed discomfort, might do the work in a more morally acceptable way. If a practice *aims at* morally reforming offenders, as opposed to inflicting deserved suffering, reasons must be given why painful methods are preferred.

Belligerent or recalcitrant wrongdoers may require ‘cooling off’ periods in which they are forced into isolation until heightened emotions have died down and they are calm enough to focus their attention on their situation. Offering an incensed offender some reward if he calms down is unlikely to prove effective. The lines of communication with such a person are distorted by his emotions and a reward proffered for cooperation will easily be refused. The prospect of suffering, however, is less easily ignored. Offenders will not always care about rewards offered to entice them to do what they do not want to do, but the threat of inflicted suffering more often presents to them a greater

evil and is therefore a more powerful motivator. Wrongdoers are to suffer so that they feel they *must* do better in their conduct. The coercive power of deprivation and the use of force are what will most reliably and effectively obtain the discipline and cooperation of recalcitrant wrongdoers.<sup>20</sup>

To inflict suffering is to coerce the wrongdoer, producing an offender whose reason (in Stage One at least) for becoming an active participant in a reform programme is the fear of further suffering. Is this not a violation of the Kantian maxim? Coercing an offender into taking the first step on the path to moral reform *is* a violation of the autonomy of the offender, but this is tolerable for the same reason that it is not a violation of the Kantian maxim.<sup>21</sup> The instrumental infliction of suffering in Stage One is done partly in the *offender's* interest, meaning that the offender's ends are precisely what are being served by preventing him from exercising his autonomy in a manner that does not serve him or anyone morally well.<sup>22</sup> Inflicted suffering is a means to an end and is justified because the benefits (moral reform) outweigh the drawbacks (the suffering itself). The sort of suffering that is inflicted for instrumental purposes and which serves the ends of the sufferer, even though it is inflicted without the consent of the individual, does not violate Kant's maxim.

In bringing about the moral reform of an offender reformers are concerned to inflict suffering to the minimum degree necessary to move particular offenders away from stubborn resistance and toward active participation. Treatments causing suffering are only permissible insofar as and to the extent that they are necessary to the pursuit of the goal of reform. Moral reform carries with it the demand that the offender eventually come to see the point behind the suffering that was inflicted upon him and behind his engaging with reformers such that he will become a willing, as opposed to merely coerced, participant. We inflict suffering to produce offenders who are more receptive to moral reform, where an offender's being receptive means that she is seeking to

---

<sup>20</sup> This is not to claim that *all* wrongdoers respond well to inflicted suffering or that inflicted suffering is always necessary. I am providing reasons why a practice aiming to morally reform offenders would inflict suffering on them at all.

<sup>21</sup> Restricting someone's freedom does not automatically imply that they are being used merely as means without regard for their own ends. Thus not all violations of autonomy are violations of the Kantian maxim.

<sup>22</sup> There is an element of perfectionism in moral reform, but I shall not reply to liberal political worries more than to say that moral reform is an *attempt* to convey a moral message to offenders who may choose

understand what is happening and why, making an effort to pay attention to beliefs and attitudes towards her and her conduct which may not match her own.

### §5.1.2. **Inflicted suffering is a reformatory technique**

In addition to combating the recalcitrance of particular offenders, inflicted suffering itself may have reformatory effects. Morally reforming someone sometimes requires inflicting suffering on wrongdoers for the purpose of ‘igniting’ those processes of self-evaluation which do the real work involved in moral reform. Christopher Bennett explains how receiving a painful response to his wrongdoing works on an offender:

He not only comes to see that he must have offended the others, or must have done something to awaken their indignation, but in trying to find out what he has done, he begins to review his own conduct, to see it in a new light. He asks himself how his colleagues must see him. He tries to find out what he has done to provoke their reaction. He sees his behaviour ‘from the outside’; he imagines how it would have appeared to someone else. In doing so, though, he comes to take their point of view more seriously. He comes to imagine how his actions affect others. He comes to think about his conduct from a perspective dislocated from his immediate aims and desires. And this brings him to see that what he has done is something that he does not really think he ought to have done. He sees his action as one for which others could rightly blame him. He comes to see that he has harmed other people, people he cares for. He *accepts* their blame. Accepting blame represents a kind of moral progress... He becomes aware of the damage he might be doing to people’s feelings, to his relationships with them, and starts to think about how he really ought to treat people, what they have a right to expect of him. Rather than laughing off his responsibilities when these clash with what he feels like doing, he comes to think that he ought to be taking them a bit more seriously.<sup>23</sup>

Bennett describes the positive activities the offender engages in after receiving a sufficiently painful response to his wrongdoing. The ‘shock’ of a painful response to his action causes the offender to review his conduct in a way that has him *reform himself*. Getting hurt induces the offender to focus his cognitive and affective capacities on the reasons why he has been subjected to some suffering, with the end result ideally being that the offender commits to not repeating his past conduct because he acknowledges its

---

to listen or not. A wrongdoer remains free to protest reform if the presiding laws are considered unjust. Moral reform is not an influenceability theory.

<sup>23</sup> Christopher Bennett (2002), p. 153.

damaging immorality. If Bennett has described a true scenario, then the power of inflicted suffering to deliver an initial, sharp message causing an offender to review his conduct ought to be harnessed by reformers. Indeed, inflicted suffering is a potent technique for communicating to an offender that he has done something truly wrong.

Moral reform must communicate the seriousness and wrongness of the offence to the wrongdoer. A reward system is poorly suited to the expression of such messages; one cannot convince someone that he ought to feel contrition because he has done something seriously *wrong* by offering him rewards, nor can one express one's resentment, indignation or anger that way.<sup>24</sup> A reward system does not inspire the self-evaluation Bennett's protagonist engages in and it is seriously at odds with the ordinary attitudes toward wrongdoers of indignant society and resentful victims, people whose emotions and beliefs proscribe the offering of pleasures in response to wrongdoing. A rape victim would have serious reservations about a practice that offers the rapist some rewards in return for his participation, but find it a fitting and just expression of her attitudes if the rapist were made to suffer for his crime.

### §5.1.3. **Reformative punishment theories**

Christopher Bennett's example (above) offers an account of a plausible scenario, but to think that merely inflicting some measure of suffering on an offender is enough to effect genuine moral reform is to underestimate what producing a "morally autonomous person freely attached to the good" entails. The practice of moral reform I am describing will not involve exposing an offender to forms of deprivation and suffering, such as incarceration or work, in the often vain hope that doing so *alone* will reliably produce a "morally autonomous person freely attached to the good". Some theorists, though, have offered theories that do rely on the reformative effects of inflicted suffering in an attempt to capture the virtues of moral reform for punishment practices.

What I shall call "reformative punishment theories" attempt to incorporate the benefits of moral reform into punitive practices by arguing that punishment (minimally

---

<sup>24</sup> That moral reform is not a practice unable to express the emotions felt towards wrongdoers is argued for in Chapter Two.

defined as the infliction of suffering on a wrongdoer) has reformatory effects.<sup>25</sup> These theories are meant to advocate punishment to those who believe moral reform to be the proper goal of responses to wrongdoing by pointing out that punishment can produce an offender who has genuinely learned his lesson. In response, I describe the limited efficacy of inflicted suffering as a reformatory treatment. There is no particularly strong connection between the infliction of suffering on an offender and that offender's coming to properly appreciate the wrong done such that she is moved to feel guilty and remorseful, take responsibility for her past action, and commit to a more morally upright existence.

Reformatory punishment theorists aver that inflicted suffering (most commonly incarceration) alone can communicate to the wrongdoer the seriousness of the wrong done, why the act was wrong and the reasons why the offender ought to modify his conduct to avoid such actions in future, and even ensure that the offender commits to improved conduct in future. This is too much to expect from punishment. Reformatory punishment theories appear to be excessively optimistic about cases like Bennett's example; they overestimate the strength of the connection between inflicted suffering and moral reform. Jean Hampton writes, "Punishment is the state's attempt to teach a moral lesson, but whether or not the criminal will listen and accept it is up to the criminal himself".<sup>26</sup> If punishment is merely inflicted suffering, it seems that reformatory punishment theorists are content to leave it open to the offender to extract a variety of meanings from her treatment at the hands of the authorities. Inflicted suffering of a kind that respects the moral status of the offender leaves far too much to the wrongdoer, naively allowing the cognitive and affective capacities of responsible wrongdoers to respond to the suffering inflicted upon them in a number of ways that are not those leading to moral improvement.

Under reformatory punishment the offender is given no guidance on *how* to suffer her sentence, what meaning or moral message to extract from it. Steven Tudor explains that, "the voluntariness of the offender is to be exercised in deciding *what mode*

---

<sup>25</sup> See Jean Hampton, 'The Moral Education Theory of Punishment', 1984, pp. 208-238, Herbert Morris, 'A Paternalistic Theory of Punishment', 1981, pp. 263-271.

<sup>26</sup> Jean Hampton (1984), p. 232.

of subjectivity she is to adopt toward that to which she is compulsorily subjected”.<sup>27</sup> A reformatory punishment theory leaves an offender free to reconcile herself to her punishment without becoming morally reformed or seeing any validity in the resentment and indignation directed towards her for her conduct. There is nothing about punishment that prevents the wrongdoer simply “copping” her punishment stoically,<sup>28</sup> accepting her lot with equanimity and merely tolerating a complaint that receives no further thought.

The capacity for casuistry is common to all responsible moral agents; they are naturally endowed with the psychological capacities to avoid thinking ill of themselves and their actions, and may even require this talent for their mental and emotional well-being.<sup>29</sup> Raskolnikov, the protagonist in Dostoyevsky’s *Crime and Punishment*, displays his powers of casuistry even before he has perpetrated his crime:

Meanwhile it would seem, as regards the moral question, that his analysis was complete; his casuistry had become as keen as a razor, and he could not find rational objections in himself.<sup>30</sup>

Afterwards, having murdered Alyona and Lizaveta Ivanovna in cold blood, Raskolnikov peddles his rationalisations to himself and others:

“I’ve only killed a louse, Sonia, a useless, loathsome, harmful creature.”  
“A human being - a louse!”  
“I too know it wasn’t a louse,” he answered, looking strangely at her. “But I am talking nonsense, Sonia,” he added. “I’ve been talking nonsense a long time. . . That’s not it, you are right there. There were quite, quite other causes for it!”<sup>31</sup>

And finally, having served some months of his prison sentence, Raskolnikov reveals that his attitude towards his punishment and his crime is not that of someone undergoing moral reform:

But it was not his shaven head and his fetters he was ashamed of: his pride had been stung to the quick. It was wounded pride that made him ill. Oh, how happy he would have been if he could have blamed himself! He could have borne anything then, even shame and disgrace. But he judged himself severely, and his exasperated conscience found no particularly terrible fault in his past, except a simple *blunder* which might happen to any one. He was ashamed just because

<sup>27</sup> Steven Tudor (2001), p. 599.

<sup>28</sup> Ibid. p. 596.

<sup>29</sup> See Shelley E. Taylor, Jonathan D. Brown, ‘Illusion and Well-Being: A Social Psychological Perspective on Mental Health’, 1988, pp. 193-210.

<sup>30</sup> Fyodor Dostoyevsky (19--), p. 72. Thanks to Peter Goldie for suggesting this novel.

<sup>31</sup> Ibid. p. 404.

he, Raskolnikov, had so hopelessly, stupidly come to grief through some decree of blind fate, and must humble himself and submit to “the idiocy” of a sentence, if he were anyhow to be at peace.<sup>32</sup>

Like Raskolnikov, we rationalise our wrongdoings, portraying ourselves as the victims of circumstance, projecting the blame onto others, failing to examine the quality of our own will. In another instance, an offender might accept his guilt but, instead of committing to better conduct in future, he could identify himself with his past conduct and form an image of himself as a person who is just not meant to follow the rules, a person for whom good conduct is simply unnatural.

Casuists are discussed in philosophy as “motivated believers”, people whose beliefs are corrupted by powerful psychological mechanisms in favour of their own positive self-conceptions and vested interests.<sup>33</sup> The mechanisms of motivated believing operate when persons are confronted with evidence of a state of affairs that is not desired. Shelley Taylor and Jonathan Brown write:

These strategies may succeed, in large part, because both the social world and cognitive-processing mechanisms impose filters on incoming information that distort in a positive direction; negative information may be isolated and represented in as unthreatening a manner as possible. These positive illusions may be especially useful when an individual receives negative feedback or is otherwise threatened and may be especially adaptive under these circumstances.<sup>34</sup>

An offender facing punishment or moral reform surely counts as a situation in which the overwhelming presence of “negative feedback” elicits a strong response from the mechanisms of motivated believing. Any practice that has as its aim the moral reform of wrongdoers must be designed to break down the illusions casuistic offenders conjure for themselves.

Returning to Bennett’s example, there is little reason to believe that, after suffering some punishment, the protagonist will reliably, “come to take [his victims’] point of view more seriously”, no reason to think that inflicted suffering alone will bring him to, “think about his conduct from a perspective dislocated from his immediate aims

---

<sup>32</sup> Ibid. p. 525.

<sup>33</sup> A classic example is of the mother who continues to set the table for a son whom she knows is not coming back from the war.

<sup>34</sup> Shelley E. Taylor, Jonathan D. Brown (1988), p. 193.

and desires”, and induce in him the thought that, “he has done something that he does not really think he ought to have done”.<sup>35</sup> Instead, it seems more reasonable to believe that the mechanisms of motivated believing will conjure their illusions in defence of the emotional and mental stability of the prisoner. Indeed, modern penal systems produce an alarming percentage of recidivists, testament to the failure of incarceration as an effective instrument of moral reform.<sup>36</sup>

Russ Shafer-Landau puts the point succinctly by observing that offenders know what they are doing is wrong when they do it, “they know the rules, and they know that most people feel strongly about them”.<sup>37</sup> Responsible wrongdoers expect the expressions of the “reactive attitudes”<sup>38</sup> (in the form of inflicted suffering, ostracism, and so on) in response to their offences that seem to be a surprise for Bennett’s protagonist, yet they perform their immoral actions anyway. It then becomes difficult to see how it is that inflicted suffering effects moral reform in those who knew the possible consequences would be those usual to an immoral act, yet were not dissuaded. In the context of a punishment practice, offenders will view inflicted suffering as an expected, and in some sense accepted (though certainly not desired), consequence of their having being caught doing wrong. An offender may find no indication in his punishment that what he is supposed to glean from the experience is a lasting lesson for *his* moral character, rather than, like Raskolnikov, a regret at being caught.

In the context of a moral reform programme, the use of inflicted suffering will not be as open to Shafer-Landau’s objection as punishment practices are, simply because it will be made clear to an offender that he is being made to suffer not as retribution for his having done wrong, but because he is failing to participate in a process that serves his (and the victim’s) interests and because inflicted suffering expresses our justifiably strong

---

<sup>35</sup> Christopher Bennett (2002), p. 153. I do not mean to destroy all faith we have in Bennett’s scenario. It seems to me that he describes an event that, under certain conditions, does happen.

<sup>36</sup> “Our recidivism rate in this country is an astronomical 70 percent for juveniles and 63 percent for adults...As taxpayers, we could ask why we have a system that is failing 63 percent or 70 percent of the time to rehabilitate people and put them back in the community”. From ‘The US Penal System: Restorative and/or Retributive Justice?’ in *Woodstock Report*, No. 61, March 2000. Effective moral reform programmes will do much to combat recidivism.

<sup>37</sup> Russ Shafer-Landau (1991), p. 200.

<sup>38</sup> Peter Strawson (1962), p. 63. The term refers to those emotions normally felt in response to a wrongdoing. Resentment and indignation are examples.

feelings of moral disapprobation toward the wrongdoer. All treatments in moral reform are ostensibly directed toward inspiring a Bennett-like process, making it clear to an offender that an appropriate response to his suffering is sought from him. The inability of inflicted suffering to direct offender's attitudes and counteract the mechanisms of motivated believing is the central practical reason why a full practice of moral reform should be preferred to reformatory punishment. More must be done with offenders than simply punishing if we are to reliably produce "morally autonomous persons freely attached to the good". It is the task of reformatory punishment theorists to provide reasons why they do not advocate a response to wrongdoing that does all it can to achieve its goal.

#### §5.1.4. Punishments tailored to reform offenders

Reformatory punishment theorists may advocate sentences tailored to suit the specific temperament of each wrongdoer and the nature of the crime committed, "Thus we might require a rapist to work in a battered women's shelter, or an arsonist to work in a burn victim ward".<sup>39</sup> Tailored sentences enhance the reformatory effects of inflicted suffering by providing the wrongdoer with insight into the real suffering of the victim, thereby offering explicit counter-evidence against motivated beliefs or illusions an offender might have created for herself. This is an attractive sentencing policy for any practice that aims to morally reform wrongdoers because it helps produce empathetic knowledge in the wrongdoer. Being in a position that confronts a wrongdoer with the vivid consequences of her action stirs her conscience and imagination so that she may become more cognizant of the wrong she has perpetrated.

If they adopt more than a tailored sentencing policy, reformatory punishment theorists are in danger of losing contact with a practice that can reasonably still be called punishment. A response to wrongdoing that is to remain a punishment practice cannot include treatments beyond inflicted suffering. Conceived of as being on a continuum, treatments that are more and more designed for the purposes of effective moral reform

---

<sup>39</sup> Russ Shafer-Landau (1991), p. 198. Tailored sentencing entered punishment theory in response to objections that punishment is not a constructive practice. This type of sentencing policy is the hallmark of what is known as 'Restorative Justice', a punishment theory that aims to benefit the victim.

are less and less the inflicted sufferings that constitute punishment. When we begin rational debate or psychological counselling with an offender we move away from punishment, even if these extra treatments are done to enhance the reformatory effects of inflicted suffering. Tailored sentencing is the most reformatory punishment theorists can offer to combat wrongdoers' casuistry.

Even for reformatory punishment that has adopted a tailored sentencing policy there remains a concern that the powers of casuistry possessed by any wrongdoer have been underestimated. Just as in Bennett's example, there appears to be no sufficiently strong connection between a rapist's being made to work in a rape counselling centre and a change in his opinion of women as weak, lesser beings to be objectified and abused. Serving a tailored sentence, the rapist may receive an initial shock but soon have it fade into an esteem-building context in which he 'finds evidence' to support a conception of himself as a good worker helping bitter, man-hating women recover from their abuse by other men.<sup>40</sup> With their perceptions left unchecked, offenders serving tailored sentences may still conjure evidence for the legitimacy of their wrongdoings and the illegitimacy of their punishments rather than the other way around. Self-deception and the repression of emotions (guilt, shame, remorse) associated with accepting responsibility for one's having done wrong remain powerful barriers to the full and lasting moral reform of wrongdoers serving tailored sentences. Taylor and Brown write:

Some potentially contradictory information never gets into the cognitive system. Pre-existing theories strongly guide the perception of information as relevant... Interpretational biases also mute the impact of incoming information.<sup>41</sup>

The power of the mechanisms of motivated believing should not be underestimated. People are able to live their lives in a world 'coloured' by the mechanisms of casuistry, as Sartre thought of those self-deceived persons living in "bad faith".<sup>42</sup>

Taylor and Brown cite empirical studies revealing the impermanence of changes in self-perception following exposure to information casting one in a negative light.<sup>43</sup>

---

<sup>40</sup> This is an empirical possibility meant to demonstrate that moral reform will likely be more effective than tailored sentences.

<sup>41</sup> Shelley E. Taylor, Jonathan D. Brown (1988), p. 202.

<sup>42</sup> Jean-Paul Sartre, *Being and Nothingness: A Phenomenological Essay on Ontology*, 1993.

<sup>43</sup> Shelley E. Taylor, Jonathan D. Brown (1988), p. 202.

Motivated believers may deceive themselves into thinking that they have accomplished lasting moral improvement, not realising how difficult it is to effect a real change in the quality of one's will. Along similar lines, Iris Murdoch makes the excellent observation that an offender may accept his suffering as just yet not undertake the difficult task of improving his will:

The ideas of guilt and punishment can be the most subtle tool of the ingenious self. The idea of suffering confuses the mind and in certain contexts (the context of 'sincere self-examination' for instance) can masquerade as purification.<sup>44</sup>

The "masquerade" refers to the offender's substituting suffering for any efforts toward good practical change. An offender might take in the messages of tailored sentencing, yet view his suffering as evidence that he need not undertake any further efforts at moral reform. In contrast to reformatory punishment, moral reform must be a practice that aims to bridge the gap between an offender's witnessing the damage done and an offender's accepting responsibility for his wrongdoing *and* for the improvement of the quality of his will. Hilary Bok writes, "Changing our character for the better...is long, slow, and often tedious work", and we must be careful not to see our suffering as a relief from undertaking the difficult task of moral improvement.<sup>45</sup>

#### §5.1.5. Moral reform begins with reformatory punishment

Stage One of moral reform is reformatory punishment, suffering inflicted for the purpose of morally reforming offenders. This does not make *moral reform* a punishment practice since it does not only inflict suffering. Nor does it make of moral reform what might be called a "combination theory" – a response to wrongdoing that combines two or more practices that have *distinct* aims. Though it contains treatment that can be described as punishment, this technique and all others used in moral reform are directed toward the same end. There is nothing inconsistent about this; it is merely that the methods for attaining that end differ enough for them to fall under the ambits of separate practices.

Theorists will find that an inconsistency often occurs within practices that *aim* to achieve more than one goal of responses to wrongdoing. For instance, if a practice

---

<sup>44</sup> Iris Murdoch (1970), p. 68.

aims both to deter and morally reform wrongdoers conflict will arise as the treatments that best deter (e.g. boiling oil) are not those that best reform. Where lies potential for this conflict, one aim must be given priority and the subordinate aim catered for as much as possible thereafter. Though Herbert Morris declares his scepticism of any justification proposed as an “exclusive” justification of a response to wrongdoing,<sup>46</sup> this should not be read as a demand for combination theories. Morris is well aware that the justification for a response to wrongdoing does not only lie with what Hart calls the “General Justifying Aim”,<sup>47</sup> but with all the desirable effects that come from the attainment of that primary goal. Realising that the debate between theorists of responses to wrongdoing can only be settled if we discover a “General Justifying Aim” from which flows the realisation of the other desiderata at which responses to wrongdoing could aim, Morris declares that advantage should be taken of, “the rich over-determination of goods promoted by the practice of punishment”, but that “determinative weight” be given to “paternalistic ends”.<sup>48</sup>

I agree with Morris’s assessment that the debate can only be settled by a practice that produces the results aimed for by the many responses to wrongdoing that have been suggested. I also agree that moral reform is the goal that should be given the most weight, because I believe that victims find little satisfaction in practices that do not affect their subjects’ attitudes towards their wrongdoing, and, I argue, because moral reform produces many of the desiderata of responses to wrongdoing. However, contrary to what Morris believed, punishment is not the best method of producing a “morally autonomous person freely attached to the good”.

## §5.2. Stage Two

In the second stage of moral reform the offender is subjected to and participates in reformatory treatments that do not inflict suffering. Stage Two furnishes the offender with persuasive evidence inducing in her the emotions of shame, guilt and remorse

---

<sup>45</sup> Hilary Bok (1998), p. 177.

<sup>46</sup> Herbert Morris (1981), p. 271.

<sup>47</sup> H. L. A. Hart (1968), p. 3.

<sup>48</sup> Herbert Morris (1981), p. 271.

signaling her acceptance of culpability. This stage builds on the influence that Stage One's inflicted suffering has had on the offender. Stage Two features treatments designed to bring about the genuine and lasting realisation of the goal of moral reform in a manner more reliable than is done by inflicting suffering. It is in this stage that treatments designed to compensate for the shortfalls and enhance the advantages of inflicted suffering and tailored sentencing play their important role. Reform distinguishes itself from punishment practices by aiming to dispel the rationalisations casuistic offenders generate in their own defense. A comprehensive moral reform programme is required to foil the mechanisms of motivated believing that allow an offender to suppress emotions like shame, guilt and remorse, and develop distorted beliefs about himself and his situation.

I shall not endeavour to offer positive accounts of the reformatory techniques that are best suited to producing an offender who is genuinely reformed. The task of selecting these treatments and explaining their workings is better left to criminal psychologists. However, the various treatments employed in moral reform *must* (for moral and logical reasons) belong to that class of techniques that appeal to the cognitive and affective capacities of responsible wrongdoers, and may range from deep psychoanalysis to rational debate on ethical issues. In Stage Two the offender is led to feel and be confronted with those emotions felt by a person who has some understanding of the moral significance of his wrongdoing. I shall describe the goal of Stage Two in terms of the emotional responses it is aimed at eliciting from wrongdoers, highlighting the fact that these emotions are particularly *meaningful* modes of suffering more strongly connected to moral reform than to punishment.

#### §5.2.1. **Only morally responsible wrongdoers are reformed**

Reformers are not concerned to *begin* the moral education of offenders, they are concerned to help offenders better appreciate the normative force behind actions they already know to be morally valuable. People who are morally responsible for immoral acts are not thought to be ignorant or incapable of the good at the time, they are thought to have failed, for some reason, in the proper exercise of moral capacities they already

have. Certainly among the reasons mature persons do the wrong thing may be ignorance of, say, the correct ranking of moral values or a difficulty reconciling morality with self-interested pursuits, but these are problems we think they ought not to have because they possess the cognitive and affective capacities of persons who are morally responsible for their actions. This is not to declare that moral reform does not contain education of some kind. Hart writes:

Reforming methods include the inducement of states of repentance. Or recognition of moral guilt, or greater awareness of the character and demands of society, the provision of education in a broad sense, vocational training and psychological treatment.<sup>49</sup>

Many wrongdoers may struggle with the competing demands of moral and nonmoral values and moral reform ought to be a practice that helps offenders integrate morality into their lives without compromising the benefits to the victim of producing a “morally autonomous person freely attached to the good”.<sup>50</sup>

Reformees qualify themselves for moral reform by being the kinds of people we reasonably expect to have done the right thing and to have known what the right course of action was when they instead did wrong. What Hampton writes of reformatory punishment is true of moral reform:

[Reformatory] punishment is not a way of treating a “sick” person for a mental disease, but rather a way of sending a moral message to a person who has acted immorally and who is to be held responsible for his actions.<sup>51</sup>

Morally responsible wrongdoers possess a cognitive and affective maturity which provides an understanding of what the good and bad actions in particular instances would be, and this awareness, *ceterus paribus*, ought to guide their decisions. Reformees are people with some moral commitments and knowledge, people who have capacities developed enough to understand the rational justification for, and appreciate the normative force behind, the wrongness of their actions. The subjects of moral reform are offenders who have an appreciation of the good and yet have voluntarily performed an immoral action, and can, therefore, knowingly and voluntarily decide not to do it again.

---

<sup>49</sup> H. L. A. Hart (1968), p. 12.

<sup>50</sup> An objection that moral reform confers too much benefit on the wrongdoer is answered in Chapter Four.

<sup>51</sup> Jean Hampton (1984), p. 214.

Moral reform targets only those who have performed immoral actions and who are morally responsible for their wrongdoings for the practical reason that it *cannot work* on someone who does not possess the emotional and mental maturity that qualifies her as a responsible moral agent.<sup>52</sup> The cultivation of autonomous moral agency in a person responsible for his wrongdoing requires *improving* his moral agency through exposing him to treatments that engage with his mature affective and deliberative capacities. The person being reformed must be free from any mental and physical incapacities (madness, mental slowness) that would excuse him from responsibility and inhibit his moral reform. The practice relies on the agential (moral) capacities a morally responsible person has to strengthen attachments to values that have been violated by that person, by having reformees come to *hold themselves* responsible for their actions. What happens to the wrongdoer in Bennett's example (above) is that he turns his moral capacities in on himself in order to review his conduct and, ultimately, modify it. In a similar way, but with alternative techniques, Stage Two of moral reform aims to influence morally responsible offenders that they take responsibility for redressing their past conduct and improving their future behaviour.

Moral reform will not be so immediately effective that a reformee will emerge without having to consciously resist the urge to transgress. The practice is an appeal to morally responsible offenders to do better than they have, but does not have such strong effects as to rule out even normal akratic behaviour.<sup>53</sup> Certainly the hope is that good moral values do become internalised to the point of being part of an offender's "automatic, habitual reasoning",<sup>54</sup> but this can take years of practice (on Aristotle's plausible account of moral development) and perhaps even continued psychological treatment. At the very least successful reform means that an offender feels the normative force behind a value sufficiently for it to be more difficult for her to be moved to violate that value in future.

---

<sup>52</sup> In making the practical point here I am omitting mention of the ethical reasons for morally reforming only morally responsible people.

<sup>53</sup> Thanks to Samantha Vice for this observation.

<sup>54</sup> Jennifer Church (2002), p. 366.

### §5.2.2. Guilt, shame and remorse - the “reformatory emotions”

When an offender comes to appreciate the good or ‘take to heart’ a value previously violated she naturally feels guilt, shame and remorse – what I shall call the “reformatory emotions”. An offender who does not come to feel the reformatory emotions is not reformed. Stage Two requires that a reformee feel these emotions as evidence of its success, as constructive forces in the ongoing reform process and as a motivating force behind the atoning acts undertaken by a reformee in Stage Three.

On the most plausible account of the nature of these emotions they contain a cognitive element (a belief) as well as an ‘affect’ or feeling. The affect is the unpleasant experiential component suffered by the subject. The cognitive element is constituted by the beliefs that one is responsible for a wrongdoing and that the respondents’ responses are legitimate and legitimately directed. It is in coming to believe in her culpability for an immoral act that an offender comes to feel the uncomfortable affects of guilt, shame and remorse. Reformers are concerned to ensure that an offender acquires those beliefs along with which the affective elements of the reformatory emotions appear.<sup>55</sup>

Shame and guilt are different emotions that can be distinguished by their psychological characteristics and what it is that they motivate the subject to do to relieve their discomfort. For a wrongdoer to experience shame, she must come to regard herself as having been seen, “inappropriately, by the wrong people, in the wrong condition”, as the subject of a negative judgement of her *self*.<sup>56</sup> Guilt, on the other hand, is elicited by negative responses from personal authority figures to the immoral *action* the wrongdoer has perpetrated.

Shame and guilt are morally less valuable emotional responses to wrongdoing than remorse because they are self-interested, lacking in concern for others as the victims of immoral actions.<sup>57</sup> Tudor does not regard shame and guilt as especially valuable modes of suffering because neither of these emotions is directed at recognising the plight and moral status of the victim. He defines what he calls “guilt feelings” as, “a fear

---

<sup>55</sup> I am ignoring instances of nonmoral guilt.

<sup>56</sup> In *Shame and Necessity* Bernard Williams notes that this judgement stems from an individual’s own sense of who they ought to be (personal conviction rather than public opinion), as reformers would prefer. Bernard Williams (1993), p. 78.

engendered by a sense of having disobeyed and angered some authority figure”<sup>58</sup>, and shame as, “a feeling of worthlessness induced by a sense of having prompted the derision or disapproval of some respected audience”.<sup>59</sup> Shame and guilt are not in any sense directed towards recognition of the victim or the fact that what one has done is immoral. Instead, these emotions concentrate on the offender herself, and may motivate her to modify her own conduct and character out of the self-interested need to avoid feeling worthless and to appease the fearful ire of personal authority figures.

The fact that shame and guilt are not other-directed and remorse is leads us to recommend remorse as the emotion involved in what Tudor calls “meaningful suffering”:

The meaning of this kind of suffering is made possible by its intentional object, that is, what it is suffering “for” or “at”, rather than its sufficiently desirable consequences.<sup>60</sup>

Remorse is a mode of meaningful suffering because its intentional object is both, “the Other whom one has wronged and oneself as having done wrong”.<sup>61</sup> It is just that emotion felt out of the belief that one is responsible for having harmed someone intentionally and unjustifiably. In remorse moral reform finds a mode of suffering that is valuable in itself as a fundamental acknowledgement of the moral status of the victim and of oneself as being responsible for a *wrongdoing*. If responses to wrongdoing are to offer victims the moral recognition that is so meaningful when coming from the person responsible for undermining their standing as moral agents of equal worth, then it is the meaningful mode of emotional suffering, namely remorse, that ought to be cultivated in offenders. However, it is only practices that aim at moral reform that can hope to offer the victim this recognition because it is only reformees who will feel remorse.

Though shame and guilt are missing the special moral dimension that endows remorse with its meaning, they remain modes of suffering that are valuable to the reform process. All three of the reformatory emotions have a role in moral reform, motivating atoning actions and changes in conduct and character. Guilt motivates an offender to avoid repeating his offence since it is an emotion identifying those actions that invite the

---

<sup>57</sup> Steven Tudor (2001), p.588.

<sup>58</sup> An authority figure can be anyone whose opinion is respected by the offender.

<sup>59</sup> Steven Tudor (2001), p. 588.

<sup>60</sup> Ibid. p. 586.

<sup>61</sup> Ibid. p. 588.

wrath of personal authority figures. Guilt focuses on, “what I have done” (agency), whereas shame looks at, “what I am” (identity).<sup>62</sup> Shame connects wrong action to the self, helping one understand one’s relations with the world and one’s actions, and motivating one to “rebuild” a self that conforms to one’s conception of a good person.<sup>63</sup> The shame-filled wrongdoer works to alter his character since it is a defect of the self that has invited the disdain of others. And finally, remorse motivates an offender to take responsibility for the redress of his wrongdoing and to undertake the arduous task of effecting genuine moral improvement in himself.

Reformers keep painful emotions present to the reformee to motivate change or bolster an offender’s moral commitments and character. We should not underestimate the power of these emotions to produce a ‘private hell’ that motivates a wrongdoer to undertake the improvement of the quality of his will. In the same way that taking pleasure in an activity can enhance one’s performance, so the uncomfortable affects of guilt, remorse and shame may act to boost the reformees’ participation in the further refinement of their will.<sup>64</sup> It is the task of the reformer to harness the reformatory emotions by confronting the casuistic offender with that evidence which strongly suggests to him he ought to feel them and undertake those changes in character (shame) and conduct (guilt), and the atoning actions (remorse) that relieve them.

As I have argued, offenders are not disposed to thinking ill of themselves and are endowed with psychological mechanisms that prevent feelings of contrition and knowledge of culpability from coming to offenders’ conscious attention. Reformers must be careful not to allow reformees the opportunity to ‘abdicate’ responsibility by suppressing the reformatory emotions or, as Murdoch observed, by viewing this emotional suffering as a substitute for any efforts towards substantive change.<sup>65</sup> Effective moral reform requires treatments designed to keep offenders honest, to have them understand their feelings of guilt, shame and remorse and undertake the projects these useful and valuable emotions demand for their relief. Again, the unfortunate deficiency of reformatory punishment theories that they do not take account of a

---

<sup>62</sup> Bernard Williams (1993), p. 92.

<sup>63</sup> Ibid. p. 94.

<sup>64</sup> Nancy Sherman (1999), p. 185.

<sup>65</sup> Iris Murdoch (1970), p. 68.

responsible wrongdoer's capacity for emotionally skewed perception and the suppression of the reformative emotions, makes it so that this practice does not produce reformees as reliably as full moral reform. Punishment does not harness the reformative emotions, but allow offenders ample opportunity to rationalise away any reasons they may have for feeling shame, guilt and remorse, those emotions a morally responsible wrongdoer ought to feel in response to the plight of her victim.

Furthermore, by not making mention of the emotions a reformee comes to suffer, reformative punishment theorists have overlooked an important source of suffering that contributes to the satisfaction of the demand for suffering that underpins all punishment practices. The reformative emotions arise with an offender's believing in his culpability and seem naturally to increase in discomfort as an offender becomes aware of the relative severity of his offence. Reform theorists can argue that reformees suffer emotionally enough to recommend to punishment theorists those reformative techniques they have traditionally found no room for because they do not inflict suffering.<sup>66</sup>

### §5.2.3. Emotional suffering is not reformative punishment

Any emotional suffering produced by reformative techniques that influence an offender to believe in her culpability for a wrongdoing, like psychoanalysis and rational debate, is not the result of punishment. Punishment is minimally described as 'inflicted suffering'. To inflict suffering is to do things to an offender that cause her suffering, it is to force suffering upon that person. Describing guilt, remorse or shame as "inflicted" pains is misguided because people cannot be brought to believe that they have angered the authorities, performed an immoral action or demonstrated defective character *without* feeling these emotions.

Shame, guilt and remorse are the *necessary* symptoms of a genuine belief or 'deep' acknowledgement that one has performed an immoral act for which one is to take responsibility. It is never the case that one consciously believes one is morally responsible for a wrong done without feeling some measure of the reformative emotions.

---

<sup>66</sup> That moral reform contains the suffering usually demanded for a wrongdoer is argued for in Chapters Two and Three.

The natural occurrence of emotions like shame, guilt and remorse in response to an appreciation of one's responsibility for a wrongdoing makes it so that some degree of suffering is a necessary product of moral reform's success. Successful reform *always* involves suffering for the offender, but this does not mean that this suffering is inflicted. Moral reform confronts wrongdoers with the truth of their status as wrongdoers and suffering comes about as a natural accompaniment to an offender's acknowledgment of the truth.

Respect for the Kantian maxim leaves room for a degree of voluntariness in moral reform, such that it is correct to view moral reform as merely offering reasons why an offender ought to undertake a change in her behaviour. We *attempt* the moral reform of wrongdoers in a way in which punishment is not an attempt to inflict suffering. Reformers do not inflict the beliefs with which these emotions share a necessary connection on the wrongdoer, they do not *make* offenders believe in their culpability and they do not *make* offenders feel the reformative emotions that naturally accompany this belief. In reform, one is confronted with evidence for certain beliefs and will simply "find oneself"<sup>67</sup> experiencing these uncomfortable emotions if those beliefs arise.

The information and stimuli an offender is confronted with do not automatically produce a "morally autonomous person freely attached to the good", but require the participation of the wrongdoer if they are to have any effect. Hampton observed, "[W]hether or not the criminal will listen and accept it is up to the criminal himself".<sup>68</sup> The reform process is mediated by the offender's efforts at contemplation of her situation and introspective examination of her conscience. She must focus her cognitive and affective capacities on her situation, concentrating her attentions such that she comes to acknowledge that she is culpable for a wrongdoing, and consequently feels the reformative emotions. The offender is free to accept or reject the assistance offered by a reform programme, and so bears some responsibility for whether or not she comes to feel the reformative emotions.

It will not do either to say that guilt, shame and remorse are self-inflicted pains because, again, they are not chosen but simply 'well up' in a person who has come to

---

<sup>67</sup> Steven Tudor (2001), p. 593.

<sup>68</sup> Jean Hampton (1984), p. 232.

acknowledge their culpability for an immoral action. It is a phenomenon of which Dostoyevsky was aware:

“But the real geniuses,” asked Razumihin frowning, “those who have the right to murder? Oughtn’t they to suffer at all even for the blood they’ve shed?”  
“Why the word *ought*? It’s not a matter of permission or prohibition. He will suffer if he is sorry for his victim”.<sup>69</sup>

Hilary Bok provides a useful illustration of the point, explaining that recognition of culpability has the same connection with the reformatory emotions as recognition of the loss of love has with the pain of heartbreak.<sup>70</sup> These are not inflicted pains, merely appropriate responses to the violation or loss of something important to us. A person can neither choose to feel the natural conative affects that accompany certain beliefs nor can they directly choose to believe those beliefs,<sup>71</sup> though, as I have argued is required of a reformee, they may participate in the epistemological process (moral reform) that provides evidence for those beliefs. The involuntariness of beliefs and the necessary connection between certain beliefs and the reformatory emotions thus make it incorrect to claim that a person chooses to inflict the pains of the reformatory emotions on themselves.

### §5.3. Stage Three

A reformee who enters the third stage of moral reform is expected to *voluntarily* undertake atoning acts that communicate his sincere contrition to the victim and society. If a wrongdoer does not offer up atoning acts or, *ceteris paribus*, does not perform them when the opportunity arises, we must regard him as still in need of reform. The desire to perform atoning acts is part of what it is to be genuinely remorseful:

The person who professed to feel remorse but felt no urge to do something of the kind I refer to as atoning acts would be thought to suffer from only a distorted sort of remorse.<sup>72</sup>

---

<sup>69</sup> Fyodor Dostoyevsky (19--), p. 259.

<sup>70</sup> Hilary Bok (1998), p. 169.

<sup>71</sup> Bernard Williams has argued the latter claim in ‘Deciding to Believe’ in his *Problems of the Self*, 1973.

<sup>72</sup> Steven Tudor (2001), p. 589.

That a wrongdoer feels remorse to such an extent that it motivates him to perform atoning acts signals the commitment to and appreciation of the good sought by moral reform. It is by atoning that an offender takes responsibility for his wrongdoing in recognition of the victim, and it is here that the victim derives a benefit (moral recognition from the offender) peculiar to moral reform.

### §5.3.1. What is atonement?

Atoning acts are forms of apology, playing a communicative role that (hopefully) impacts on the opinions and feelings of the victim and the community, “[such acts] can create in the mind of the public the belief that he deserves to be forgiven or, at least granted a fresh start with a clean slate”.<sup>73</sup> The metaphor of *distance* is commonly used to explain how performing atoning actions or sincerely offering to do so dissociates the wrongdoer from the wrongs of the past, and confirms her present commitment or proximity to virtuous conduct. Atoning acts express a change in moral character that may have both constructive effects - the re-establishment of relationships with the community and closure for the victim - and redemptive benefits - reacceptance by third parties can only strengthen a wrongdoer’s commitment to responsible moral agency.

Offences are typically self-interested actions, and so we look for an opposite sort of act in atonement. When an offender atones, she undertakes some form of suffering which communicates that she is taking responsibility for an act she knows to have been wrong. Suffering is present in atonement because it conveys the sincerity of contrition and the offender’s appreciation of the harm suffered by the victim. We do not believe that the victim’s moral status and well-being have become important to an offender if that offender is not prepared to make some fitting sacrifice as a token of his remorsefulness. Offenders do not atone by handing out old clothes destined for disposal anyway. It is the *cost* to the offender of atoning acts that convinces us that he possesses a genuine desire to ‘do right by’ the victim.

The form of atoning actions may be varied according to the specific needs of the victim, the seriousness of the crime and the capacities of the wrongdoer. Remorse and

---

<sup>73</sup> Ibid. p. 79.

apology can be expressed verbally and, as cases become more serious, through action. Atoning actions ought to be selected according to their power to communicate an offender's appreciation of the wrongness of the offence of which she is guilty. For example, "the doctor who cheated the Medicare system [volunteers] weekend service in a state-supported clinic",<sup>74</sup> for a period long enough to 'fit' the severity of the crime, thereby expressing his appreciation of the harmfulness of the fraud. We judge atoning acts by their "fittingness" – a demand for the atoning act to be of a suitable *kind* - the rapist serving in a rape counseling centre - and of a suitable *amount* - one does not need to spend one's savings in atoning for an insult, a sincere apology such as, "I'm truly sorry", will do, though it will not suffice if one has murdered.<sup>75</sup> An atoning act that 'fits' is one which offers tangible evidence of the offender's successful moral reform.

It is surely unrealistic to deny that an offender undertakes atonement partly for self-interested purposes; a suffering reformee will be well aware of the benefits that accrue from atonement. The wrongdoer is motivated to atone because it may lead to his forgiveness and the re-establishment of fractured relations with the victim and the community. Furthermore, atoning acts may be done to gain relief from the reformatory emotions, "reducing the pangs of conscience".<sup>76</sup> But these are too self-interested to ever be the primary aims of atonement, the third stage in a practice that aims primarily to serve the victim. There are morally better and morally worse forms of atonement, and those acts of atonement undertaken primarily for self-interested purposes are the morally worse kind.<sup>77</sup>

The most morally valuable form of atonement is done chiefly to communicate to the victim and the community the wrongdoer's sincere appreciation of the wrongness and severity of her conduct and her acknowledgement of the victim's status as someone of equal moral standing. Atoning acts are means of communicating the sincerity of an offender's moral reform, her contrition and appreciation of the material, emotional and moral damage she is responsible for having done to the victim. The best sort of atonement is done for the victim – a tangible apology signalling an offender's changed

---

<sup>74</sup> Jean Hampton (1984), p. 224. Atonement may look very like the tailored sentences mentioned above.

<sup>75</sup> Sentencing principles limiting how much offenders ought to suffer are discussed in Chapter Four.

<sup>76</sup> David B. Hershenov (1999), p. 79.

<sup>77</sup> Thanks to Sam Vice for this.

attitude toward the victim and symbolically recognising the victim's equal moral worth.

Tudor writes:

[W]hat gives meaning to the expressivity available in the suffering [is] its being "for" or "about" (or intentionally directed toward) one's remorse, apology and repentance.<sup>78</sup>

Acts of atonement are endowed with a meaning coming from their deep connection to the offender's appreciation of her wrongdoing and her suffering remorse; they qualify as instances of Tudor's "meaningful suffering" because their intended objects are the victim, victim's relations and the community. Expectations of the benefits of atoning should be the least of a wrongdoer's motivations to atone, and moral reform programmes are precisely what are required to ensure that this is the case. The full complement of reformatory techniques are our best method for producing reformees who understand the purpose behind atoning and who are properly motivated to do it.

Certain offences are of such a nature that the damage inflicted cannot be undone. When one has murdered, there is nothing that can be done to undo the past and restore things to the way they were. Atonement does not share in the inadequacy of suffering inflicted to 'make amends' or 'repay a debt' when it comes to the worst crimes because it is purely an expressive device.<sup>79</sup> Atoning acts that are undertaken out of sincere recognition of the victim and appreciation of the gravity of the offence are all that can be done in those sad cases where what is lost cannot be replaced. It is a failing of punishment practices that they do not aim at producing this response from wrongdoers or do so less reliably than moral reform.

### §5.3.2. Atonement is not punishment

The "repayment theory" of punishment views the offender as having incurred a 'debt' to the victim and to society through his wrongdoing and punishment is inflicted as settlement.<sup>80</sup> Atoning acts can easily be construed as attempts to settle the 'debt' of the offender since they offer sacrifices to the victim and society that sometimes repair the

---

<sup>78</sup> Steven Tudor (2002), p. 600.

<sup>79</sup> This does not preclude atoning acts that do repair the damage done. The point is that, as I have defined atonement, these acts are done as tangible expressions of the wrongdoer's feelings towards the victim.

damage that has been done. Atoning acts may appear to be punishments that offenders inflict on themselves in settlement of what is owed the victim.

John Cottingham correctly states that atonement is *not* what is meant by the “repayment theory” of retributive punishment, “There could not be an atonement theory of punishment; atonement is something voluntarily undertaken, punishment is something exacted”.<sup>81</sup> The suffering *undertaken* by a wrongdoer in atoning is not the suffering *undergone* by the punished offender. The former has the wrongdoer choosing his sacrifices, doing what needs to be done to communicate his newfound recognition of the victim as someone wronged by his own hand, a communication that may lead to repaired relations with the victim and the community. The latter has the punisher leaving the offender no choices, inflicting suffering as the meting out of just deserts, the repayment of a ‘debt’, as a condition of fair play, or the annulment of the crime.<sup>82</sup> The implications of the difference Cottingham points out are that one cannot punish oneself by voluntarily causing oneself suffering and that one cannot be forced to atone. This voluntarism adds a value to atonement that no punishment can have since moral recognition freely given is more valued than what is done under duress or out of obligation.

## §6. Conclusion

Moral reform is a practice that begins with reformatory punishment and, because of the casuistic tendencies of many wrongdoers, continues with reformatory techniques that do not *inflict* suffering and which are therefore no part of any punishment practice. The practice aims to successfully produce a “morally autonomous person freely attached to the good” by appealing to the cognitive and affective capacities of morally responsible offenders. The reformee is a person who feels the meaningful reformatory emotions out of an appreciation of his status as a wrongdoer, and is thereby moved to undertake the improvement of his will and express the symptoms of effective moral reform through meaningful acts of atonement. This reformee offers much of what is sought from a response to wrongdoing – he satisfies the victim’s need for a sincere apology, he

---

<sup>80</sup> John Cottingham (1979), p. 238.

<sup>81</sup> *Ibid.*

<sup>82</sup> *Ibid.* pp. 238-246. The importance of these retributive rationales is critically discussed in Chapter Three.

undertakes atoning acts that may assist the victim with her loss even as they communicate sincere contrition, he is less likely to transgress, he is better educated, and he experiences suffering as a consequence of his wrongdoing.

The basic proposition of this first chapter has been that the moral reform of offenders is a desideratum of responses to wrongdoing and that proper moral reform best accomplishes this goal. The claim that moral reform provides what victims, their relations and society want from a response to wrongdoing is given further attention in the next chapter. An analysis of those emotions typically felt towards a wrongdoer (resentment, indignation, moral hatred) allows us to further discover what victims, their relations and society *want* from a response to wrongdoing and to assess which practice can best provide satisfaction. In this regard, the limited ability of punishment to produce offenders who are properly reformed or reformed at all proves again that punishment is an inferior response to wrongdoing.

## **Chapter Two**

### **Expressing and relieving reactive attitudes.**

## §1. Introduction

The defining aim of moral reform is the production of a “morally autonomous person freely attached to the good”, a reformee who suffers those emotional pains associated with recognition of oneself as responsible for a wrong done, and who is thereby motivated to undertake atoning acts in the form of meaningful sacrifices expressing appreciation of the victim’s moral status and the immorality of his conduct. Moral reform includes inflicted suffering to ensure the cooperation of wrongdoers and to communicate to wrongdoers that they have committed an immoral act inviting moral disapprobation (Stage One). It includes further techniques (psychotherapy, rational debate) designed to counteract casuistry on the part of wrongdoers and to produce wrongdoers who feel guilt, shame and remorse (Stage Two). The practice concludes with the acts of atonement a wrongdoer undertakes out of remorsefulness and sincere acknowledgment of her responsibility for having unjustifiably harmed another (Stage Three).

The moral reform of an offender is ordinarily a desideratum of responses to wrongdoing and is better achieved by moral reform than by any punishment practice. Inflicted suffering (punishment) alone cannot penetrate the defensive mechanisms that casuistic offenders deploy in response to sanction or properly motivate offenders to undertake substantive changes in the quality of their will in response to their suffering. Thus, moral reform better delivers the change in character, the moral recognition and sincere contrition that we want from an offender.

An inability to reliably reform wrongdoers makes punishment practices less adept at providing victims, their relations and society with the apologetic reformee that they desire. This same failing is capitalized upon in this chapter. Punishment practices do not *relieve* particular morally valuable emotional responses to wrongdoing (e.g. resentment, indignation, moral hatred) as effectively as moral reform does, because punishment practices are less able to prevent the motivated believing that inhibits the production of “morally autonomous persons freely attached to the good”. If, as I shall argue, the relief of resentment, indignation and the like is tantamount to satisfying what victims and third parties desire from a wrongdoer, then punishment is less able to provide

victims, their relations and society with what they want from a response to wrongdoing. I explain which emotions are to be relieved, how doing so provides what is desired by victims, their relations and society, and which practice does so most effectively. In addition to their relief, our emotions ought to be *expressed* by our practices and I explain that the inflicted suffering in both punishment and moral reform are evenly matched in this regard.

## §2. Emotions inform our practices

A desideratum of a response to wrongdoing is that it express those emotions we judge people ought to have in response to wrongdoing. To this I add that our responses to wrongdoing ought also to relieve these emotions. In appealing to the emotions to adjudicate the debate between punishers and reformers I am *not* arguing anything more than that the emotions can help in the justification of a practice. I do not wish to enter the difficult debate concerning whether moral truths are discovered via the emotions or whether emotions are at the base of our moral judgements. There may be justifications for moral reform and for punishment that make no mention of the emotions, and there may be justifications of these practices that rely solely on the emotions. I am not committed to either extreme, but find that those emotions typically felt in response to wrongdoing reflect our most common judgments about what a practice should do, whatever the source of those judgements might ultimately be.

We can judge that certain emotions ought to be felt by people, be they victims or third parties, towards a wrongdoer and/or a wrongdoing. There are appropriate and inappropriate emotional responses to wrongdoing. One ought to feel a little resentful at being stood up, but one ought not to feel a terrible fury or nothing at all. Society ought to feel indignation towards the rape of its women because this demonstrates respect for women's moral worth, and wrongdoers ought to feel remorse for their sins because remorse reflects meaningful recognition of the wrong and the victim.

Appropriate emotional responses can be seen to endorse particular practical responses to wrongdoers (and, in the case of remorse, *by* wrongdoers) if we also hold that those emotions we consider indicative of good persons should also inform our normative

claims about actions.<sup>83</sup> For instance, the kinds of emotions we judge that people ought to feel are also the kinds of emotions we judge that people ought to express. Thus, if we judge that a certain emotional response to circumstance warrants expression because, say, it reflects commitment to a moral value (as indignation does), then we are given a guide as to what to do in response to those circumstances. If a valued emotion like indignation is expressed by a certain practice, then that practice gains some moral weight as an expression of that emotion. We ought to express valuable emotions, and practices that do so become things we ought to do. Following this line of argument, I add that the same value must be given to the *relief* of those reactive attitudes we judge are appropriate reactions to wrongdoing. A response to wrongdoing gains added value if it relieves the emotions appropriately felt by victims, their relations and society. In justifying practices by providing reasons in terms of something else (emotions) accepted as valuable, we find a way of using the emotions we ought to feel, and therefore express and relieve, to identify the best response to wrongdoing.

### §2.1. Strawson on why the reactive attitudes ought to be expressed

Peter Strawson provides two reasons for disregarding responses to wrongdoing that are not expressive of the “reactive attitudes” - those emotions we typically feel as part of our everyday interactions with other people.<sup>84</sup> Firstly, only practices that are expressive of the reactive attitudes will properly take account of the great importance of the Kantian maxim as a demand for goodwill in our ordinary interpersonal relationships. Strawson writes:

The reactive attitudes rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves [and towards others].<sup>85</sup>

The demand for “a certain degree of goodwill” I identify with the Kantian demand that the dignity of the other person qua rational agent must always be respected. In harming another human being I am not treating her with her ends in mind or displaying goodwill

---

<sup>83</sup> Peter Goldie (2002), p. 19.

<sup>84</sup> Peter Strawson (1962), p 63.

<sup>85</sup> *Ibid.* p. 71.

towards her, but using her in the pursuit of my own ends. Matching Strawson's account of the demands of the reactive attitudes to Kant's moral principle, we find that persons have an emotional commitment to the Kantian demand that their dignity as moral agents always be respected in interactions with other persons. When the Kantian maxim is violated, that is, when a person displays indifference, contempt or malevolence in treating us or others as mere means, we naturally feel resentful, indignant, indeed a whole range of reactive attitudes.<sup>86</sup> Our experiencing these emotional responses expresses our recognition that the demand for interpersonal goodwill on the part of responsible moral agents has or has not been satisfied.

Any practice that is an expression of the reactive attitudes expresses the value of the demand for goodwill or, as I have it, the Kantian maxim, which is an essential moral feature of decent human interaction. Our responses to wrongdoing express our deep commitment to how the *victim* ought to have been treated by being expressions of the reactive attitudes. The moral status of the victim as a member of the moral community to whom respect is owed will go sorely unrecognized if our response to their victimization does not express the reactive attitudes. For us to respond to wrongdoing unemotionally would betray us as people who do not care about our fellows or who do not see offences against them as of any importance.

The image Strawson paints of moral interaction is one imbued with meaning by the interplay of emotions and attitudes – a wrongdoer expresses a certain attitude toward a victim and society and the victim and society feel and express certain attitudes in response. It matters a great deal to us that certain attitudes be expressed in interpersonal relations, our attitudes towards one another provide interaction with its meaning in the same way that actions done out of goodwill are commonly deemed more valuable than actions done simply out of social contractual obligation. We know from personal experience that even the smallest gesture, perhaps the giving of a bouquet of flowers, is made more meaningful if done out of love or concern rather than out of mere politeness or obligation. Atonement gains much of its status as a “meaningful mode of suffering”

---

<sup>86</sup> The relation between Kant's rational maxim and the emotions will only be explored as much as is necessary for the purposes of this chapter, though it is surely a fruitful topic. My claim is not that the source of the Kantian maxim is, contrary to Kant's thinking, emotional. It is just that we respond emotionally to violations of the Kantian maxim.

from its being an expression of remorse.<sup>87</sup> It is a small step to seeing that we find practices that express our attitudes towards one another more valuable than those that do not, and this is imperative for any practice, like moral reform, that prides itself on offering the victim recognition and satisfaction. If I have offended against my peers and they reprimand me for it, the reprimand is made the more poignant and meaningful if it expresses their resentment than if it, somehow, is not the expression of any emotion at all. Indeed, if their response were not the expression of an emotion it would fail to convince me that I had done anything really *wrong*. Therefore, a practice that expresses our emotions will be more meaningful to both victim and offender.

The second reason to be gleaned from Strawson's *Freedom and Resentment* is that only practices expressing the reactive attitudes will respect the Kantian maxim as a guide to how responsible *wrongdoers* ought to be treated.<sup>88</sup> Responses to wrongdoing that suspend the reactive attitudes are not the right sort of practices when responsible moral agents are their subjects, because they do not respect the moral status of their subjects.

Our expectation that goodwill be expressed in social interaction means that we view other people as moral agents who are part of moral relationships in the moral community. Strawson observes that we know empirically that those instances in which resentment, indignation and other reactive attitudes are *not* expressed in response to some transgression involve 'excusing and exempting' conditions like, "He couldn't help it", or, "He's a hopeless schizophrenic".<sup>89</sup> Those instances in which the reactive attitudes are not expressed are always either instances in which a moral agent lacked control or was coerced or in which the offender, because of some deficiency in her cognitive or affective capacities, cannot be regarded as a member of the moral community. Towards this latter group we adopt what Strawson calls an "objective attitude", in which we see the offender as, "an object of social policy; as a subject...to be managed or handled or cured or trained".<sup>90</sup> There is no meaningful moral interaction with these unfortunates; we regard them from a detached vantage point that is devoid of the reactive attitudes (but not all

---

<sup>87</sup> Steven Tudor (2002), p. 586.

<sup>88</sup> Peter Strawson, 'Freedom and Resentment', 1962, pp. 59-80.

<sup>89</sup> *Ibid.* pp. 64, 65.

emotions). These are people we do not hold responsible yet with which something must be done for the protection of society and the subjects themselves.

A practice that does not express the reactive attitudes in cases where excusing and exempting conditions are *absent* will violate the Kantian maxim. Given that our regard for a *morally responsible* offender as someone whose value, despite their offence, is that of a member of the moral community, is demonstrated by our feeling the reactive attitudes towards her, if our responses to the wrongdoings of responsible moral agents are not expressive of the reactive attitudes, they are not respectful of the dignity of the offender as a responsible moral agent, that is, they cannot but violate the Kantian maxim. When feelings of resentment or indignation are not felt in response to an offence it signals that the offender is not regarded as a member of that class of beings, namely moral agents, towards whose acts we do feel the reactive attitudes. Instead, the offender is regarded as a member of that class of non-responsible beings towards whom we adopt the objective attitude. Adopting the objective attitude towards a responsible wrongdoer is to deprive her of her humanity, to violate the Kantian maxim by regarding her as something to be “treated” or “trained” in the interests of society.

Strawson observes that responses to wrongdoing ought to be expressions of the reactive attitudes because the moral status of the victim as someone unjustifiably wronged is respected by such expressions. Expressions of the reactive attitudes are acknowledgements of the victim’s right to be treated as an end and not as a mere means. Moreover, the moral status of the offender as a morally responsible wrongdoer is respected by practices that express the reactive attitudes.

## §2.2. How the reactive attitudes are expressed

What is ordinarily regarded as an expression of resentment and other emotional responses to wrongdoing is that the victim and society ostracize the offender, or yell at her, or deprive her of her liberty or inflict suffering in some other way. Strawson writes:

---

<sup>90</sup> *Ibid.* p. 66.

So the preparedness to acquiesce in that infliction of suffering on the wrongdoer which is an essential part of punishment is all of a piece with this whole range of attitudes.<sup>91</sup>

We can interpret this as meaning that Strawson believes that the connection between the reactive attitudes and the demand that the offender suffer is just that the suffering of the offender is what expresses people's feelings toward a wrongdoing. We go about expressing resentment or indignation by punishing. In general, inflicting suffering is our way of letting an offender know how we feel, and it follows that we identify the expression of emotions like resentment and indignation with punishment.

Psychoanalysis, rational debate and other reformatory techniques that do not inflict suffering are not expressive of the reactive attitudes even though, as I shall argue, these treatments contribute to bringing about the *relief* of the reactive attitudes. This does *not* mean that these extra reformatory treatments "train" or manipulate wrongdoers to the ends of society. Though psychological treatments and rational debate require a certain *objectivity* on the part of the reformer, these are reformatory techniques helping wrongdoers to focus those cognitive and affective capacities they possess as morally responsible wrongdoers on taking responsibility for their conduct. Psychologists do not violate the Kantian injunction by helping wrongdoers or anyone accept culpability for their conduct and confront their uncomfortable emotional responses to their past. The offender's own ends as a responsible moral agent are exactly the concern of all techniques employed in moral reform, and it is therefore false that 'extra' reformatory techniques adopt the objective attitude towards wrongdoers.

If moral reform is to be a practice that expresses our emotional responses to wrongdoing then it must inflict suffering to communicate to the wrongdoer that she has perpetrated an immoral act about which people have strong feelings. Stage One does just this, inflicting suffering because such sanction communicates to the offender how victims and society feel about the wrongdoer's conduct, making it clear that she has done something of which the victim and society disapprove. As Christopher Bennett's example showed,<sup>92</sup> the expression of the reactive attitudes through the infliction of suffering on the wrongdoer delivers a clear message of disapproval that may have

---

<sup>91</sup> Ibid. p. 77.

reformatory effects.<sup>93</sup> Thus, the inflicted suffering that is the expression of the reactive attitudes fits very well into a practice of moral reform.

### §2.3. Why the reactive attitudes ought to be relieved

As much as morally permissible we should design responses to wrongdoing that deliver what those who have been offended against want, and it is certain valuable reactive attitudes that are our guides in discovering what a response to wrongdoing must provide the victim with to satisfy those desires. Consider resentment, defined as an appropriate emotional response to an action in defense of one's self-respect.<sup>94</sup> We infer from this emotion that persons feeling resentment *want* their self-respect defended, an appropriate desire that ought to be fulfilled. An appropriately negative emotional reaction to a wrongdoing indicates that something undesired has occurred and that, *rightly*, the subject also has a desire to change that state of affairs.

Just as certain valuable emotions warrant expression and so point us in the direction of the right response to wrongdoing, so these same emotions lead us to desires that warrant relief and so again to a correct course of action. We judge that certain valuable reactive attitudes ought to be relieved, *ceteris paribus*, because we recognise that their relief is something that it is appropriate for persons to want. If resentment is judged to be an emotion victim's ought to feel, then it is appropriate for a resentful victim to advocate a response to wrongdoing that provides him with what he wants, which is (I explain) whatever provides relief from his emotion. That is, morally valuable emotional reactions to wrongdoing indicate that victims, their relations and society have certain appropriate desires that ought to be satisfied by a response to wrongdoing.

The nature of emotional responses is that they signal what is important to us, and uncovering what specifically certain emotions are responses to indicates what effects of a wrongdoing on the victim are most unwanted. Thus the reactive attitudes provide insight into what people will feel it is important to gain from a response to wrongdoing. If it is the case that a response to wrongdoing is not dedicated to the relief of our reactive

---

<sup>92</sup> Christopher Bennett (2002), p. 153. See Chapter One.

<sup>93</sup> See Chapter One, Stage One.

<sup>94</sup> Jean Hampton, Jeffrie G. Murphy (1988), p. 16.

attitudes, or if it is the case that a response to wrongdoing is relatively less effective at relieving the reactive attitudes, this counts as a failing of that practice and as grounds for selecting an alternative.

#### §2.4. How the reactive attitudes are relieved

We can infer from our having a negative emotional reaction to some particular state of affairs that something has happened that we do or do not want, that is, that I desire that a particular state of affairs should or should not have come to be. My feeling resentment in reaction to an offence against me is a negative reaction to a particular event, and this prompts the inference that something has happened that I would prefer not to have occurred. In the case of a negative emotional reaction to wrongdoing we can reasonably infer that the emotion signals the presence of a desire to remove or alter the cause of the feeling. A negative emotional reaction indicates that something bad has occurred, and normal people ordinarily desire for bad things and the uncomfortable emotions they cause to go away. If resentment is a negative emotional reaction to an offence against one's self-respect we reasonably infer that people feeling resentment want their self-respect to be defended or affirmed in some way. The desire accompanying this negative emotional reaction motivates me towards its relief; that is, people who feel resentment have a desire to do something or see something done that amounts to an affirmation of their self-respect such that their desire is relieved.

*Satisfying this desire also relieves the emotion.* Our desires are satisfied when we believe that what is desired has come about, and our emotions are relieved when we believe that the cause of our emotional reaction no longer obtains, that is, when we believe that we no longer have reason to feel the way we do.<sup>95</sup> If what is desired is that the cause of our emotional distress should go away, then the belief that it has disappeared relieves both emotion and desire. For resentment, the desire is satisfied when one believes one's self-respect has been affirmed and this same belief informs us that we no longer have reason to feel resentment. It is in providing sufficient evidence for me to

---

<sup>95</sup> Relief is not instant. Emotions tend to linger even after the cause is quite gone. Nonetheless, convincing someone that the cause of their emotion is gone contributes greatly to the relief of that emotion.

believe that my self-respect has been defended that a response to wrongdoing relieves both my resentment and accompanying desire. If I believe that inflicting suffering on wrongdoers is an adequate defence of my self-respect then doing so relieves my resentment and my desire for my self-respect to be affirmed. This is what I mean by the relief of the reactive attitudes.

The project is to follow the lead of those emotions we judge to be valuable responses to wrongdoing to find out what victims', their relations and society rightly want and what practice might offer relief from the uncomfortable reactive attitudes felt in response to a wrongdoing. My task is to find out what the reactive attitudes contribute to the debate about what we really want from wrongdoers. By analyzing morally valuable reactive attitudes we discover what those persons feeling emotions in response to a wrongdoing want. We evaluate responses to wrongdoing according to their ability to give us what we want and thereby relieve our uncomfortable emotions.

#### §2.4.1. What do victims, their relations and society want?

If we concentrate our attention on the expression of the reactive attitudes it appears that victims, their relations and society want to inflict suffering on the offender in response to his wrongdoing. Strawson holds that we acquiesce in the "*infliction of suffering*"<sup>96</sup> (my italics) on the offender, which is tantamount to the claim that the reactive attitudes point out a *specific* sort of practice, namely a punishment practice. However, if the *expression* of an emotion does not bring about the *relief* of that emotion, then it is a mistake to believe that what expresses an emotion also provides people feeling the myriad reactive emotions with what they want. For instance, we can express fear by screaming or pulling back from the snake, but these are not actions that relieve our fear because they are not actions that aim at removing the source of our fear. The relief of fear requires actions that remove the cause of fear, not actions expressing the fact that I am feeling afraid. Thus, we do not get what we want (relief from fear) by expressing our fear. In this way we see that the expression of a reactive attitude (inflicted suffering) *may not* provide us

---

<sup>96</sup> Peter Strawson, (1962), p. 77.

with what we want from a response to wrongdoing, unless that expressive action also removes the cause of the emotion.

We come to know what we want from wrongdoers by examining how the reactive attitudes are *relieved*, as opposed to *expressed*. We must uncover which practice relieves or best relieves reactive attitudes because that same practice produces what people feeling appropriate emotions desire. The question to ask is: Can inflicting suffering alone provide victims and other witnesses to a wrongdoing with what their commendable emotional reactions indicate they desire? That is, can punishment relieve the reactive attitudes we feel in response to wrongdoing? If the answer is negative and the reactive attitudes are relieved by what moral reform delivers, then moral reform is a better response to offences.

To find out what it is that satisfies or best satisfies those reactive attitudes we regard as valuable emotional responses to wrongdoing (those we think ought to be relieved), that is, to find out whether we ought to punish or morally reform wrongdoers, I closely analyze a group of reactive attitudes commonly called the “retributive emotions”, which are just those reactive attitudes usually identified with punishment practices: indignation, malicious hatred, moral hatred, resentment and what Jeffrie Murphy calls “retributive hatred”.<sup>97</sup> My intention is to demonstrate that, if we are aiming to relieve the retributive emotions, it is a mistake to identify these emotions exclusively with any form of punishment because they are best relieved by moral reform. I therefore refer to them as the ‘so-called retributive emotions’.

#### §2.4.2. The so-called retributive emotions

As far as the relief of the reactive attitudes is concerned, the so-called retributive emotions cannot be pointed to as motivators for any particular response to wrongdoing until it is established what one desires should happen to the offender or should be done by an offender. We might think that we can find this out by following a certain principle: an

---

<sup>97</sup> Jean Hampton, Jeffrie G. Murphy, *Forgiveness and Mercy*, 1988. Anger is an emotion linked with retributivism. In *Forgiveness and Mercy*, Hampton and Murphy regard all the so-called retributive emotions as ‘forms of anger’, so it would seem that this complex emotion is discussed below in at least some of its manifestations.



emotion is identified as retributive or reformatory if retributive or reformatory practices respectively are *alone* able to relieve that emotion, that is, *if one particular practice* is able to provide what is desired by the person feeling that emotion. If retributive punishment is able to relieve one's emotion by producing the desired state of affairs that makes one's emotion 'go away' in the way that emotions do upon their relief, then that emotion is a retributive emotion. For instance, if with moral hatred comes a desire for the offender to suffer, and the belief that an offender is suffering quenches that desire, then moral hatred points us in the direction of a practice, such as retributive punishment, that makes an offender suffer. In this way we can discover which practices relieve which emotions, or which practices provide victims with what they want.

However, in discussing the so-called retributive emotions I aim to illustrate how there are *multiple* ways of bringing about an emotion's relief and of providing victims, their relations and society with what are recognised as appropriate things to want, and therefore as appropriate things to get. If this is correct it hinders attempts to attach emotions to *specific* responses to wrongdoings. The emotions seem to offer little direct guidance in the way that our principle requires. Though the principle described above can identify so-called retributive emotions as being satisfied by and therefore demanding inflicted suffering, it cannot exclusively attach a desire for punishment to any emotion if, as I shall argue, emotions are on their own "compatible with even polarized practices".<sup>98</sup> The principle aims to identify emotions as reformatory or retributive by singling out the practice that produces what is desired, but struggles to identify an emotion with a particular practice because it is often the case that either practice can produce the desired state of affairs. Put another way, our desires are satisfied when we believe a certain state of affairs obtains, and producing evidence for the necessary belief may be within the capabilities of more than one practice.

To briefly illustrate this point, consider someone who already suffers the reformatory emotions and consequently undertakes acts of atonement that communicate his remorse and improved regard for the victim. It is not obvious that the victim will continue to believe that she has reason to feel or express resentment or any other so-called retributive emotion. It is *prima facie* plausible to suggest that the victim's seeing

---

<sup>98</sup> Peter Goldie (2002), p. 16.

an offender demonstrate the symptoms of successful moral reform relieves these emotions. Remorse and atonement would seem to offer evidence that the victim no longer has reason to feel the so-called retributive emotions.

On the other hand, consider an offender who has served his sentence and yet rejoins society unrepentantly claiming that he has never come to feel remorse for his actions. It seems appropriate and highly likely that our so-called retributive emotions well up in condemnation of this character and his past conduct. One may think that we are simply displaying appropriate disapproval of this man's character, emotional reactions that would be had in response to *anyone* of such character independent of past acts they have or could have done. But there is something more in the case of the punished. Our so-called retributive emotions arise in response to the bad character of the offender and the bad acts an offender has perpetrated, and those same emotions return when that offender exits punishment unreformed. We expect the punished to have learned their lesson and consequently to be suitably contrite, or, finding them unrepentant, we regard them as having squandered an opportunity to start afresh by not being reformed. If punished offenders are not morally improved we believe we still have reason to feel the so-called retributive emotions towards them and their wrongdoing (though we no longer have reason to express these emotions since the offender has served his sentence). Punishment may not be enough to relieve the so-called retributive emotions.

A possible confusion must be avoided here. The fact that emotions like anger, resentment and indignation fade with the moral reform or punishment of offenders does not indicate that the offender has been forgiven, where forgiveness is defined as the forswearing of these emotions.<sup>99</sup> Although there is a strong connection between responses to wrongdoing and an offender's being forgiven, I am referring to the ability of these practices to satisfy what is desired by victims feeling the reactive attitudes, not to their ability to provide moral grounds persuading a victim that she ought to give up these desires. Responses to wrongdoing relieve our reactive attitudes by providing what is desired, they do not just provide grounds for the victim to make the effort to relinquish the demand that these emotions be relieved.

---

<sup>99</sup> Jean Hampton, Jeffrie G. Murphy, *Forgiveness and Mercy*, 1988.

### §2.4.3. The relief of the so-called retributive emotions

In their debate on responses to wrongdoing, Jean Hampton and Jeffrie G. Murphy provide a useful though not exhaustive analysis of those emotions usually referred to as ‘retributive’: malicious hatred, moral hatred, Murphy’s “retributive hatred”, resentment and indignation. My project is to discover what it is that a person feeling resentment, indignation, or a species of hatred desires should be done to address a wrongdoer’s immoral action, and then to ask whether punishment practices best bring about the belief that relieves this desire and, in turn, the emotion (it is this exclusivity that will recommend a particular practice). I have suggested that the retributive emotions do not point to a particular practice and intend to demonstrate this below. If I am correct, then with respect to each retributive emotion we must discover the practice that *best* provides evidence for the belief that satisfies the desire. Moral reform and punishment will be evaluated according to their relative effectiveness in relieving the so-called retributive emotions.

*Malicious hatred* expresses the victim’s personal drive to compete with a wrongdoer in order to restore a moral status they believe to have been lowered by the offence, and is declared irrational and immoral.<sup>100</sup> Its irrationality stems from expressing malicious hatred through the hopeless attempt by the victim to raise her own moral status by lowering the offender’s moral status, and its immorality is contained in the act of lowering another’s (even a wrongdoer’s) moral worth (a violation of the Kantian maxim).<sup>101</sup> I believe this emotion is rightly rejected and so we need not consider whether moral reform or retributive punishment best relieves it.

As mentioned above, Murphy defines *resentment* as an emotional response to an action in defense of one’s self-respect.<sup>102</sup> Resentment is dismissed by Hampton as a “weak” emotion driven by the irrational fear that an offender’s view of one as of low enough moral status to be a victim may be correct.<sup>103</sup> However, I do not think Hampton has been fair to resentment. Someone who did not feel this emotion in response to an

---

<sup>100</sup> Ibid. p. 62.

<sup>101</sup> Ibid.

<sup>102</sup> Ibid. p. 16.

<sup>103</sup> Ibid. p. 148.

offence against her would be regarded as someone who lacks self-respect, a person who believes they are of low moral status. If *all* people who have self-respect feel resentment toward wrongdoings, it is more plausible that resentment is a positive emotion associated with a desire to affirm one's self-respect than with an unconfident desire to defend it because *no-one* really feels entitled to it. If with resentment comes a desire to affirm one's self-respect this is a sign that I do have self-respect and will not countenance attacks upon it. For this reason we find resentment to be an appropriate emotional response to wrongdoing, one that ought to see relief. The desire for one's self-respect to be affirmed (or defended) may be satisfied by seeing an offender made to suffer for having offended against one (punishment). However, moral reform may offer evidence that one is indeed someone worthy of respect by making an offender suffer and taking further steps to endow the offender with that respectfulness.

In both moral reform and punishment the proficiency of inflicted suffering at relieving our emotions is limited. Inflicted suffering alone may leave an offender's view of the victim as someone not worthy of respect intact. A wrongdoer may emerge from punishment with the same low regard for his victim (bearing in mind the weaker connection between casuistic offenders' attitudes towards their victims and punishment argued for in Chapter One), whereas extra reformatory techniques make moral reform more effective at producing offenders who respect the victim as a moral being. The proper defense of one's self-respect requires those further treatments of moral reform that direct offenders to meaningful modes of suffering born out of acknowledging their victims as people who ought to be respected. The emotional suffering, sincere apology and other acts of atonement, which are meaningful products of moral reform that come from the same person who attacked one's self-respect, constitute persuasive evidence that one's self-respect has been affirmed (or defended). Moral reform best relieves resentment because it will have the most success in providing resentful victims with what they want, where resentment is recognised as an emotion we ought to relieve.

Hampton defines *indignation* as an "emotional protest" mounted in defense of a value that has seen violation.<sup>104</sup> Indignation is felt not by victims (who feel the more personal emotion of resentment), but by people who believe that a value they hold has

---

<sup>104</sup> Ibid. p. 59.

been offended against and who therefore desire to defend or reaffirm that value. This emotion is recognized as valuable because one's commitment to a moral value is clearly demonstrated by a need to see that value respected. Part of what it is to value something is to desire to defend or reassert that value when it has been violated.

Though punishment may defend a value by inhibiting offenders' freedom to offend, we may also claim such a defense is mounted by the demand for wrongdoers to sincerely acknowledge the wrong done as *wrong* and undertake efforts to properly atone. Morally reforming wrongdoers can relieve indignation; there is no error in thinking that indignation may be wholly relieved by the belief that an offender has come to sincerely apologise as a sign that she has learned the immorality of her action, and this is to deny any exclusive connection between indignation and punishment. Moreover, in what better way can evidence be provided for the belief that a value is being defended than by having the offender come to hold that same value herself? The defense of a value that has seen violation is not best carried out by only inflicting suffering on an offender (even for reformatory purposes) or protecting society from her, since moral reform does these *and* employs further techniques to better motivate the offender to acknowledge that value as something to be respected. Moral reform offers the best defense of a value because the practice most effectively promotes that value in those who have violated it.

*Moral hatred* is the hatred felt by third parties in reaction to the violation of a value, and is felt towards a person identified with an immoral cause as a desire to overcome that immoral cause.<sup>105</sup> Murphy first casts doubt on there being a strong connection with punishment, "There is no desire to hurt the person *simpliciter* but only a willingness to allow such hurt if unavoidable in the pursuit of victory over his immoral cause".<sup>106</sup> This emotion is valued because it demonstrates commitment to good values; it is felt against racists, sexists, anti-Semites and others who subscribe to immoral ideologies or immoral beliefs.

It is clear that one wanting to believe that an immoral cause is being "overcome" may find some evidence in the fact of an offender's being punished (incarcerated, ostracized) for upholding that cause, but moral reform surely does better by

---

<sup>105</sup> Ibid. p. 88.

<sup>106</sup> Ibid.

removing offenders from the ranks of the immoral and increasing the numbers of those willing to fight the cause. Moral hatred does not call for defense but for attack, a positive drive to see an immoral cause eradicated. When an offender himself rejects his past immoral commitments due to the influence of moral reform it signals a victory more final and authentic than that usually won via the infliction of suffering. The mere infliction of suffering on casuistic offenders shares such a weak connection with overcoming their immoral cause that it may even strengthen wrongdoers' commitment to that cause rather than reverse it entirely as moral reform does. Thus, if it is a belief that moral victory has been attained that relieves moral hatred, then it is moral reform that most convincingly provides evidence for that belief.

Hampton's candidate retributive emotion, that emotion she feels is most representative of the true theory of retribution, is a close cousin of moral hatred because both aim at the "moral defeat" of the wrongdoer and neither requires suffering as necessary to the victory. She associates her quintessential retributive emotion with the desire to affirm the intrinsic value of victim as a moral being through the symbolic "defeat" of the wrongdoer.<sup>107</sup> A belief that the wrongdoer's claim to moral superiority (made through her immoral act) has been shown false, or a belief that the wrongdoer has been morally defeated and the equal value of the victim affirmed, is what relieves Hampton's retributive sentiment.

Hampton recommends that the offender is best 'defeated' by punishment, but offers a unique definition of punishment that would seem to encompass what I have called moral reform, "*any non-painful method, so long as it [is] still a method for defeating the wrongdoer, would still count as punishment*".<sup>108</sup> Terminological differences aside, my argument is again that moral reform supersedes punishment in relieving the retributive emotions. Moral reform provides the better evidence that the victim's equal moral worth has received recognition through the 'defeat' of the wrongdoer. What more convincing sign of 'defeat' can there be than for the offender herself to admit her 'defeat'

---

<sup>107</sup> Ibid. p. 117.

<sup>108</sup> Ibid. p. 126. It is purely a terminological point whether Hampton's definition of punishment is correct and mine wrong for, if Hampton is correct, nothing changes about the argument that the practice I have called moral reform is our best response to wrongdoing. Even so, Hampton appears to be claiming that any reformatory technique, including psychological counseling and rational debate, is a punishment. This is quite at odds with the accepted terminology.

and be motivated to acknowledge a victim's moral status as a direct consequence of the good work of moral reform? Though punishment may sometimes result in the moral reform of a wrongdoer, confronted with the power of motivated believing it cannot be as complete and effective as a full moral reform programme. A practice combating casuistry, as moral reform does, best relieves Hampton's retributive emotion.

The so-called retributive emotions discussed so far have been shown to offer little guidance in nominating a particular treatment of wrongdoers, and I have argued that these emotions are better relieved by moral reform than by punishment. In an attempt to secure at least one emotion for the retributivists, Jeffrie Murphy posits another type of hatred, "retributive hatred", that he defines as moral hatred plus the desire to see the offender receive his just deserts.<sup>109</sup> Retributive hatred features a belief that an offender deserves suffering for his wrongdoing and a desire to mete out that suffering as deserved. Unfortunately this emotion cannot do the justificatory work required of it.

The original project was to discover which response to wrongdoing certain valuable emotions in fact pick out. We already know there to be justifications for these practices that do not rely on the emotions, such as desert claims in the case of retributive punishment. To adjudicate the debate, I am appealing to valued emotions that add some separate justificatory force in favour of a particular practice. This means that the fact that emotions of *independent value* may find their relief in certain practices can be seen as recommending those practices. An emotion of 'independent value' is an emotion that does not rely on the justification of the practices that relieve it for its ratification as something we ought to feel. I ought to feel moral hatred for the racist, but this emotion is not valued because it is a strategy for seeing some practice done. One ought to feel resentment at being wronged because this emotion reflects one's self-esteem, not because its relief takes the form of a practice we value on other grounds. Emotions of independent value do not have value derived from their being relieved by certain practices, but are valued because they reflect commitment to good moral values and stem from good character.

We are looking for some guidance from the emotions in finding the right response to wrongdoing, and so cannot rely on an emotion that does not *add* further value

---

<sup>109</sup> Ibid. p. 90.

to the practice that relieves it. "Retributive hatred" is not an emotion of independent value because it relies on the justification of retributive punishment for its value. Specifically, it is Murphy's stipulation that retributive hatred contains a desire to see an offender receive the suffering he *deserves* that deprives retributive hatred of independent value. Retributivists justify punishment on the basis of desert. To say that punishment receives *further* support from an emotion containing a desire to see deserts meted out is false as it is the notion of desert that makes us value that desire. If it is the case that retributive punishment is devoid of moral worth, then so is retributive hatred because this emotion can only be valued for its attachment to the moral principle which justifies retributivism anyway. In the case of indignation, we can say that this is an emotion of independent value because an emotional reaction in defense of one's values remains commendable regardless of whether the traditional justifications (desert, for instance) of the practices that relieve or express it fall flat. Retributive hatred cannot act as a pointer toward retributive punishment because it is the practice of meting out deserts that acts as a pointer to it.

Retributivists might argue that we have a retributive emotion of *independent value* to which is attached a desire to see a wrongdoer suffer as a response to his wrongdoing. The desire for vengeance might be a good example. The belief that will bring about the relief of this emotion is just the belief that the offender is suffering or has suffered for his wrongdoing. However, it is not clear that my desire to hurt or see hurt a person who has hurt me can comfortably be declared a desire that is in some way good for me or something that should be encouraged in people. M. S. Moore regards it as the strongest objection to retributive punishment that, "the emotions that give rise to retributive judgements are always pathological".<sup>110</sup> The desire for vengeance is not obviously indicative of good moral character because it recommends the infliction of pain on another human being, something we, in many circumstances, regard as an evil.

A retributivist might respond by arguing that the desire for vengeance is valuable because it is a response to an *immorality*. But why should we think that the desire to hurt the wrongdoer is a correct response even to an immoral act? If we answer this question by appealing to desert, intuitions of fittingness, metaphors of debt

---

<sup>110</sup> M. S. Moore (1987), p. 190.

repayment and the like, then it begins to sound like Murphy's mistake, where what are already justifications of retributive punishment serve to legitimate the desire. It appears that the retributivist faces a paradox. Without reference to other justifications for the infliction of suffering on wrongdoers, there seems to be no retributive emotion of independent value. Yet with reference to these other retributivist justifications, an exclusively retributive emotion only has derived value and so cannot do the justificatory work required of it.

The case for the existence of a purely retributive emotion that has independent value is very weak, tempting one to speculate that the emotions we want our practices to relieve all share the feature of being relieved by more than one practice. Assuming there is no virtuous emotion that points exclusively to the infliction of suffering, then the retributivist must argue that there are emotions that are better relieved by punishment. I see no reason to believe that punishment better relieves any emotional response to wrongdoing given that all punishment provides is a wrongdoer's suffering and, as I have tried to show, there is a great deal of suffering present in moral reform. Inflicted suffering, emotional suffering that naturally tracks the severity of the offence, and acts of atonement that 'fit' the seriousness of the wrong – these three modes of suffering provide significant competition for any practice that hopes to satisfy an independently valuable emotion only by the suffering it involves.<sup>111</sup>

If there are virtuous emotions that specifically demand suffering on the part of the wrongdoer then they must desire particular modes of suffering if they are to help choose between the inflicted suffering of punishment and the various modes of suffering in moral reform. We are likely to judge that an emotion relieved by meaningful modes of suffering is the most virtuous of our candidates. Meaningful modes of suffering are those modes of suffering that are informed by the offender's belief that he has violated the Kantian maxim, that he has done wrong by the moral status of the victim. These modes of suffering are emotional (remorse) or emotionally driven (atoning acts motivated by remorse). Thus, if there are emotions and desires that are relieved by the offender's

---

<sup>111</sup> In Chapter Three I shall further defend the point that punishment theorists ought to be content with the suffering wrongdoers experience in moral reform.

distress itself, we ought to meet the demands of those that demand the meaningful modes of suffering delivered by moral reform rather than the inflicted suffering of punishment.

### §3. Complete and meaningful emotional exchange

Guilt and remorse (the reformatory emotions) are amongst the emotions that Strawson calls, “self-reactive attitudes associated with demands on oneself for others”.<sup>112</sup> The self-reactive attitudes are the correlates of the reactive attitudes felt by the victim and society, and it is through them that we discover an emotion-based explanation of why victims are dissatisfied with punished offenders who are not reformed. Unreformed offenders do not display those self-reactive attitudes that mean an offender holds that regard for the victim that is the due of all members of the moral community. Murphy writes:

One reason we do so deeply resent moral injuries done to us is... because such injuries are also *messages* – symbolic communications. They are ways a wrongdoer has of saying to us, “I count and you do not,” “I can use you for my purposes”... It is moral injury, and we care about such injuries.<sup>113</sup>

Strawson identifies our ‘caring about moral injuries’ with the *exchange* of the reactive attitudes. If responses to wrongdoing do not encourage wrongdoers to express their remorse in response to the victim’s expression of resentment, then they do not facilitate the exchange of emotion that symbolically redresses the wrongdoing’s meaning. The self-reactive attitudes and the pains of apology and atonement they motivate as their expression are what relieve the correlative reactive attitudes we feel towards offenders. These modes of suffering participate in the moral discourse Strawson locates at the core of our most meaningful interaction with others; it is in these non-inflicted pains that we find as complete a response to the meaning of a wrongdoing as can be. The reformer who takes responsibility for his wrongdoing by expressing remorse through atonement offers the victim his most morally significant response to wrongdoing. The strong connection with the meaningful modes of suffering constituted by the self-reactive attitudes and atonement, these products of all the techniques of moral reform, makes moral reform the most meaningful response to wrongdoing.

---

<sup>112</sup> Peter Strawson (1962), p. 71.

<sup>113</sup> Jean Hampton, Jeffrie G. Murphy (1988), p. 25.

#### §4. Conclusion

Deciding what response to wrongdoing is best requires knowing what the desiderata of responses to wrongdoing are, that is, knowing what we want a response to wrongdoing to do. One method for discovering these desiderata is to examine those emotional reactions to wrongdoing that we regard as having independent moral value, and therefore as eligible for expression and relief. The ability of moral reform and punishment to both express and relieve the reformative emotions determines their participation in the moral discourse constituted by the exchange of reactive attitudes. Since both practices involve inflicted suffering, they are judged equally capable of expressing the reactive attitudes.

With regard to the relief of the reactive attitudes, moral reform is most effective at providing victims with the meaningful emotional responses that we know they require from wrongdoers. It is the production of Morris's "morally autonomous person freely attached to the good" that best completes the moral interchange begun by a wrongdoing. The inability of punishment to properly reform casuistic wrongdoers has shown punishment to be a practice less adept at providing victims and society with the intimate satisfaction (affirmation of self-respect and moral status, defense of a value, defeat of an immoral cause) they feel most strongly about.

The analysis of the emotions has not provided an exhaustive list of desirable goals for a response to wrongdoing. There remain aims of responses to wrongdoing that appear to be the preserve of punishment, and my argument that moral reform is the superior practice will not convince unless it contributes to the satisfaction of the traditional goals of punishment practices. Punishments aim to mete out just deserts, to settle debts or re-establish a balanced economy of benefits and burdens, and to deter potential wrongdoers. The case for moral reform will be all the stronger if the practice can achieve these ends.

## **Chapter Three**

### **Desert and deterrence.**

## §1. Introduction

The central tactic of my argument that moral reform is a better response to wrongdoing than punishment has been to explain how moral reform best produces the desiderata of responses to wrongdoing. The moral reform of offenders is ordinarily a desirable goal for a response to wrongdoing, and moral reform is a better practice for achieving this goal than even reformatory punishment. Furthermore, the emotions felt in response to wrongdoing act as a valuable guide for uncovering the desiderata of these practices. In relieving these emotions, moral reform provides victims, their relations and society with what is important to them, from the affirmation of their self-respect to the defence of their values and the moral 'defeat' of the wrongdoer. The inflicted suffering in Stage One moral reform also allows the practice to be as adept as punishment at expressing morally valuable reactive attitudes. However, there are other important goals for moral reform to accomplish.

Punishment practices of various kinds have been proposed because the goals these practices hope to achieve are also desired effects of a response to wrongdoing. There are deterrence theories of punishment, arguments that punishment metes out just deserts, that it produces a fair distribution of 'benefits and burdens' or 'repays the debt' owed by the offender. The central tactic of my thesis continues here as I explain how moral reform accomplishes these ends of punishment that are established goals of responses to wrongdoing. Moral reform will compete with punishment practices on their own terms.

### §2.1. Desert and desert bases

In the argot of retribution we most commonly find the economic metaphor in which the offender has, through his wrongdoing, incurred a 'debt' to society and/or the victim. The retributivist debt metaphor is a 'desert base' or "*desert-locution*",<sup>114</sup> an attempt to flesh out the moral concept of desert. Desert bases have developed in response to objections that desert is an insufficient justification for punishment, because scrutiny of the concept

---

<sup>114</sup> Ted Honderich (1985), p. 119.

reveals it to be an appeal to bedrock intuition or a self-justifying claim that 'it is just right' that offenders be punished. Ted Honderich writes, "If *P is deserved for O* just is the proposition that it is right that A get P, then we have no independent reason for its being right that A get P".<sup>115</sup> To add substance to what we mean when we say an offender 'deserves' a certain treatment in response to her wrongdoing, retributivists have appealed to the metaphor of a moral economy where debts are to be repaid and a fair distribution of 'benefits and burdens' re-established by punishment. Unfortunately, desert-locutions, "are hardly less than paradigms of multiple ambiguity and vagueness",<sup>116</sup> and there is not one of these that can make sense of the whole of the practice of punishment.

The debt metaphor has struggled to fit persuasively with what punishment actually entails. It is unclear how inflicted suffering acts as a currency in the debt-ridden economy set up by wrongdoings:

It is far from clear how the serving of a prison sentence constitutes 'payment' for the crime committed. For, as far as the victim is concerned, his costs are in no sense 'paid back' by the imprisoning of an offender. The loss or harm he has suffered still stands.<sup>117</sup>

Though arresting the liberty of a wrongdoer does much to cause a loss to him, it does very little to restore to the rape victim what she has lost. Nietzsche was first puzzled as to how imposing a loss or discomfort on the *offender* could repair or pay back the harm imposed on the *victim*.<sup>118</sup> How does the infliction of suffering on offenders compensate anybody, let alone the victim? What debt has a wrongdoer incurred for a *failed* attempt at a wrongdoing? Who, in each and every crime, are the appropriate creditors in this economy and why are they to be compensated for the failure of others to dutifully obey the law? These questions have vexed retributivists in their attempts to make sufficient sense of punishment under the debt-repayment metaphor that it justify the infliction of suffering on another human being. This is not an adequate justification for punishment.

Herbert Morris has attempted to make better sense of the debt metaphor by articulating it in terms of the fair and unfair distribution of "benefits and burdens".<sup>119</sup>

---

<sup>115</sup> Ibid.

<sup>116</sup> Ibid.

<sup>117</sup> John Cottingham (1988), p. 764.

<sup>118</sup> Dennis Klimchuck (2001), p. 94.

<sup>119</sup> Herbert Morris, 'Persons and Punishment', 1968, pp. 472-501.

Morris claims that offenders obtain an unfair psychological advantage by not restraining themselves as other law-abiding citizens do, that people in society bear the burden of unsatisfied criminal impulses and offenders unfairly tread lighter having satisfied themselves by indulging in immoral conduct.<sup>120</sup> But this cannot be right if we know that most citizens feel no burden of restraint when it comes to immoral action – the majority of husbands are not battling the urge to beat their wives. To say that we all restrain ourselves from molesting children is to attribute to us a desire, and therefore a burden, that not all of us have. Besides, there is something decidedly inappropriate about begrudging a rapist the ‘unfair advantage’ of whatever psychological satisfaction he may have accrued from his act. We do not disapprove of what offenders do because we think it unfair for them to enjoy the spoils, our disapprobation comes from knowing that offences are *wrong* and we expect our responses to wrongdoing to be concerned with that, far stronger, moral judgement.

Variations on the ‘burdens and benefits’ theme depend on how the advantage gained or burden imposed by offending is understood, and finding a convincing account of what exactly is won and lost as a result of wrongdoing proves too difficult for punishment theorists. Shafer-Landau discounts psychological satisfaction and material gain as possible benefits of wrongdoing by simply observing that these versions would allow many acts which we do count as offences to go unpunished because they do not confer either advantage.<sup>121</sup>

Another attempt to cash out what a ‘benefit’ might be regards criminals as having, “arrogated to themselves an “extra liberty” which incarceration takes away.<sup>122</sup> Though it is uncertain what is ‘extra’ about the liberty a poverty-stricken thief might gain from his loot in an unequal society, the proper shortcoming of this account lies in its failure to make sense of the ‘proportionality’ sentencing principle, “its inability to explain why a criminal’s extra liberty increases in proportion to the wrongness of his conduct”.<sup>123</sup> There is no sense of ‘extra liberty’ that can justify harsher sentencing for murder than for child abuse. Furthermore, it sticks in the throat to say that what is wrong with rape is the

---

<sup>120</sup> Ibid. p. 81.

<sup>121</sup> Russ Shafer-Landau (1996), p. 301. Drunk driving is an example.

<sup>122</sup> Ibid. p. 302.

<sup>123</sup> Ibid.

‘extra liberty’ a rapist acquires, implying that this liberty is a good thing which is just unfairly attained.

## §2.2. Retributive ideas are indelible

Explanations of how desert and desert bases struggle to withstand close scrutiny are not meant to dismiss these justifications as irrelevant when evaluating responses to wrongdoing. Moral reform is quite amenable to these weak justifications. There is a good reason why it is an advantage of a response to wrongdoing that it lend itself to description under retributive ideas: the basic retributive intuition - that it is wrong for the victim alone to suffer – is simply beyond us to surrender. John Cottingham writes of retributivists’ attempts to explain their intuition that, “at best these metaphors illustrate, in different ways, the extent to which retributive ideas are rooted in our everyday language and thought”.<sup>124</sup> John Mackie calls it the “paradox of retributivism” that sense cannot be made of retributivist justifications, yet, “a retributive principle ‘cannot be eliminated from our moral thinking’, that retributive ideas... are deeply ingrained, a part of our lives”.<sup>125</sup> Powerful intuitions would bristle and there would be little chance of gaining the support of the majority of people if moral reform were a practice that involved nothing to satisfy desert bases.

This is not to grant too much ground to the retributivists and allow that ‘settling debts’ and ‘balancing benefits and burdens’ should figure as the *aims* of our responses to wrongdoing. “Retributive ideas” like desert and desert-locutions cannot provide the full justification required for a response to wrongdoing and therefore should not be the “General justifying Aim” of a response to wrongdoing. It is not enough that many people share retributive intuitions that appear deeply entrenched, a response to wrongdoing must ultimately make sense, and especially because of the suffering involved. However, the notion of desert is so deeply a part of our thinking that we cannot avoid dissatisfaction with any response to wrongdoing that gives an offender other than what we think she deserves.

---

<sup>124</sup> John Cottingham (1988), p. 765.

<sup>125</sup> Ted Honderich (1985), p. 139.

Following my tactic of building a reform practice that has those properties considered to be the virtues of punishment, I intend for moral reform to be the sort of practice that, without aiming to, delivers reformees of whom we may say, “He paid his debt”, or, “She does not deserve to suffer more”. By merely lending itself to description under retributive ideas, moral reform solves the justificatory problems faced by retributivism. Moral reform can be as retributive as punishment without having to justify this as an *aim* of the practice. In this way moral reform gains great appeal for punishment theorists

### §2.3. Reformees pay their debts

If suffering is the currency of debt repayment, then, given the three modes of suffering involved, moral reform ought to settle accounts. Together, the inflicted suffering of Stage One, the emotional suffering of Stage Two and the voluntary suffering of atonement in Stage Three provide all the suffering even a retributivist ought to demand. One wonders what more an offender owes after having suffering inflicted upon him, suffering the reformatory emotions and undertaking the fitting sacrifices of atonement. An offender’s appreciation of the seriousness of the wrong committed is reflected in the degree of discomfort felt in guilt, shame and remorse. These emotions naturally track the offender’s appreciation of the seriousness of her offence so that she suffers out of an empathetic understanding of the harm she has caused. Moreover, the atoning actions undertaken by a wrongdoer must involve a degree of sacrifice and suffering that reflects the seriousness of the offence if they are to persuasively express an offender’s contrition. These products and processes of moral reform should arguably be all that is required by way repaying a debt through suffering.<sup>126</sup>

David B. Hershenov offers a different interpretation of retributivism’s debt-repayment metaphor, one that turns on a regard for the victim’s needs rather than the offender’s suffering.<sup>127</sup> Hershenov’s debt-repayment theory is about “restitution” - it is only through the restoration of a “moral symmetry” that the victim can be

---

<sup>126</sup> A similar argument that moral reform involves sufficient suffering to satisfy desert is given below.

<sup>127</sup> David B. Hershenov, ‘Restitution and Revenge’, 1999, pp. 79-94.

“psychologically compensated” and “emotionally replenished”.<sup>128</sup> The restitutionary bent on the debt-repayment metaphor implies that debts are settled by practices that focus on the needs of the victim rather than on seeing the offender suffer, since the needs of the victim as victim indicate where the debt lies. It is a controversial version of retributivism, a practice traditionally focused on making the offender worse off rather than making the victim better off.

My theory of moral reform is primarily directed at satisfying the needs and desires of the victim. Moral reform expresses and relieves the victim’s emotions, producing an offender who feels meaningful emotions indicating his acceptance of responsibility for attacking the moral status of another person, and who is thereby motivated to undertake atoning actions that offer material, psychological, emotional and moral consolations to the victim. The meaningful modes of suffering that are produced by moral reform are more suited to the ‘restitutive’ debt-repayment metaphor than merely inflicted suffering, because they are directed towards the victim, offering recognition of the victim’s plight and her status as a moral being.

Suffering may be the only currency in which debts to victims are repaid, but this need not mean that *inflicted* suffering (punishment) is the best means of repayment. If it is a moral debt that the offender owes the victim, then those pains (remorse, atonement) brought on by a reformee’s recognition of the moral status of the victim as someone who ought not to suffer are better currency than the sufferings inflicted by jailers. Restitution is not made by third parties inflicting suffering on the wrongdoer, but by a remorseful reformee taking responsibility for the harm he has caused the victim. As practices increasingly cater for the needs of the victim, they will more and more be directed at producing a “morally autonomous person freely attached to the good”, and thus increasingly look like the practice of moral reform I have described.

#### §2.4. Reformees relinquish unfair benefits

If, as Morris averred, an offender has obtained an unfair psychological benefit, then her coming to believe that she is responsible for an act that is truly wrong, and consequently

---

<sup>128</sup> Ibid. pp. 90, 82 respectively.

suffering the reformatory emotions, must surely count as the most convincing evidence that any such benefit has been lost. The difficulty that punishment practices have in dealing with the mechanisms of motivated believing means that casuistic offenders have a greater chance of retaining any unfair psychological benefits in punishment than they do in moral reform. A self-deceived offender who has not appreciated the wrongness of what he has done may look upon his offence with exactly the satisfaction Morris thought unbalanced the moral economy. And even if wrongdoers have immorally acquired for themselves an “extra liberty”,<sup>129</sup> time spent both in a reform programme and undertaking acts of atonement is time spent surrendering any such ill-begotten liberty.

Clearly moral reform does very well accommodating those metaphors that make sense of retributive punishment, the punishment theory least concerned with the forward-looking effects of punishment on the offender and traditionally sat ‘furthest’ from moral reform on the spectrum of responses to wrongdoing. We might conclude that an analysis of desert bases reveals that moral reform is, *without aiming to be*, as ‘retributive’ as punishment, because we are reasonably able to say of successful reformees that they have paid their debts or surrendered any unfair benefits they may have acquired through their wrongdoing.

## §2.5. Reformees do not deserve more suffering

Turning, finally, to the most basic retributivist position: that a wrongdoer *deserves* suffering and punishment is a good way of seeing deserts meted out. A response to wrongdoing satisfies the demand for a wrongdoer to receive his just deserts only if the offender suffers as much as is deserved. Above I have argued that moral reform contains enough suffering for the reformee that the debt-repayment desert base is satisfied. Given that desert bases are attempts to explain the concept of desert, that argument is tantamount to the claim that an offender does not deserve more suffering than he experiences in the three stages of moral reform. The inflicted suffering of Stage One, the meaningful emotional suffering of Stage Two and the meaningful suffering undertaken through acts of atonement in Stage Three, in total deliver such an amount of suffering

---

<sup>129</sup> Russ Shafer-Landau (1996), p. 302.

that it would be unjust to say that an offender *deserves* more. It is doubtful that we would say of a rapist who has had suffering inflicted upon him, who feels the emotional pain accompanying true understanding of the immorality and harmfulness of being a rapist, and who undertakes a fitting degree of atonement in service at a rape counseling clinic, that more needs to be done to see him receive his just deserts.

### §3. Detering potential wrongdoers

Punishment is commonly defended on the grounds that it reduces the amount of crime committed by deterring people from committing offences. However, there are ethical objections to treating offenders for the purpose of deterring others. John Cottingham observes that punishing just to deter permits the use of excessive punishments, “such as evisceration or boiling in oil”.<sup>130</sup> If the aim of our practice is to deter, then we derive from that aim a permission to use the most effective deterrents available, and these treatments will no doubt be the most terrifying and inhumane. Deterrence-based punishment does not of itself provide reasons for not inflicting morally reprehensible treatments on offenders in achieving its goal.

A second major objection to the deterrence justification of punishment is its lack of respect for the moral status of the wrongdoer as a person with ends of his own. Kant might ask what the justification of our treatment of a particular offender has to do with the effect this would have on the behaviour of *other* people, objecting that deterrence uses offenders as mere means to a social benefit. The Kantian maxim stipulates that persons are not to be treated as mere means but are to have their status as moral agents with ends of their own respected, and it is violated by any practice that justifies the infliction of suffering on an offender for the purpose of deterring potential wrongdoers from harming society.

The above objections only hold against practices that *aim* specifically to deter. If it is the goal of a response to wrongdoing to deter wrongdoers then it cannot but violate the Kantian maxim in its treatment of offenders, nor can we avoid deriving an immoral sentencing policy from that goal. Aiming to deter means using captured offenders as

---

<sup>130</sup> John Cottingham (1988), p. 772.

mere means to an end and permitting any treatment that is an effective deterrent. However, though there are ethical problems with aiming at deterrence, it remains an advantage of any response to wrongdoing that it does deter would-be offenders from immoral activity. If it just so happens that the moral reform process is an experience unpleasant enough for it to deter potential wrongdoers, then offenders are not being treated as mere means and sentencing policy is derived from the goal of moral reform and not deterrence. We can accept deterrence as an advantageous side-effect of the unpleasantness of our responses to wrongdoing. Just as the justificatory difficulties of retributive punishment are solved by moral reform's ability to attain retributive ends without aiming to, so the problems for deterrence theorists will be avoided in the same way.

It follows from the fact that punishment gains its influence as a deterrent from the suffering it inflicts that, if we are to find any similar effects in moral reform, it is the suffering involved in being reformed that will do the work. Moral reform involves three modes of suffering that make the practice an emotionally, psychologically and even physically uncomfortable process that begins by arresting *all* the self-interested pursuits of the wrongdoer (spending time with loved ones, earning money, and so on), continues by confronting the offender with her own powerful emotional suffering, and ends with a substantial sacrifice voluntarily undertaken by the wrongdoer. Incarceration deters because it is a loss of liberty with huge opportunity costs.<sup>131</sup> Moral reform partakes of those same opportunity costs and adds on the unpleasantness (reformatory emotions, atoning acts) that accompany one's accepting responsibility for an immoral action. With all this suffering, there is no reason to think that moral reform will not be as adept at deterring potential wrongdoers as punishment.

#### §4. Conclusion

It is startling that our familiar punishment practices struggle to find sound ethical bases. Retributivists have been unable to provide sound philosophical arguments in defence of a

---

<sup>131</sup> "Opportunity cost" is an economic term that here refers to the lost opportunities for earning money, spending time with loved ones, and so on, that accrue because the offender is jailed.

practice representing such deeply held intuitions. Retributivism is forced to resort to various versions of the attractive, yet ultimately inadequate, metaphor of a moral economy in an attempt to justify the deliberate infliction of suffering on wrongdoers. The failure of desert and desert bases to explain why punishment is warranted is a problem, “grave enough to disqualify [retributive punishment] from philosophical eligibility”,<sup>132</sup> yet our intuitive commitment to desert runs too deep for retributive concepts to ever be eliminated from an assessment of responses to wrongdoing. The result is that our responses to wrongdoing ought not to *aim* at exacting retribution, but they must produce effects on the wrongdoer that lend themselves to description under retributive ideas.

Moral reform caters for our indelible retributive intuitions by producing reformees whom we may describe as having suffered as much as deserved, paid their debts, and relinquished any ‘benefits’ that were unjustifiably obtained. The deterrence of potential offenders is similarly accomplished as a happy effect of the unpleasant treatments and effects of a practice that aims exclusively at the moral reform of wrongdoers. It is a strong argument in favour of moral reform that the practice is able to achieve these most established ends of punishment practices.

The final chapter reveals that moral reform is not without its faults. Three objections are explained and given reply.

---

<sup>132</sup> Russ Shafer-Landau (1996), p. 312.

## **Chapter Four**

### **Three objections to moral reform.**

## §1. Introduction

Moral reform is not without its objectors, and three particular objections are explained and replied to in this chapter. This is by no means a comprehensive defence. Public policy issues, such as the possibility that moral reform is prohibitively expensive, and practical problems concerning the efficacy of reformative techniques and the psychology of adult offenders, are set aside in favour of purely philosophical objections to the ethics and scope of moral reform. I have selected particular objections for their relevance to distinguishing features of the practice I have described. Most difficult is the argument that moral reform permits indeterminate sentences for recalcitrant wrongdoers and so violates traditional sentencing principles. Sentencing principles determine how much suffering is acceptable for an offender and moral reform must be a practice that follows those principles determining an appropriate amount of suffering. The objection has special relevance because it undermines my claim that moral reform provides the suffering needed to satisfy desert and desert bases, those retributive concepts articulating the fundamental intuition that it is only just that the wrongdoer suffer for his offence.

A second worry is that moral reform omits from its scope certain wrongdoers whom we would punish. This objection rests on the claim that there are unreformable people who are morally responsible persons, that there exists a set of people whom we would punish but whom it is useless to even attempt to reform. The groundwork for a response to this objection was laid in Chapter One, where moral reform was distanced from “influenceability theories”<sup>133</sup> and defined as a practice wholly dependent on the cognitive and affective capacities of responsible moral agents for its success. Finally, retributivists find it inappropriate that the wrongdoer ‘benefits’ from moral reform by gaining improved character and a broad education in what is ultimately a valuable experience. This objection presents an opportunity to reinforce moral reform’s commitment to the victim, her relations and society while explaining that benefiting reformees in this way is commendable practice.

---

<sup>133</sup> Hilary Bok (1998), p 184.

## §2. The indeterminate sentencing problem

An objection to moral reform is that the practice involves indeterminate sentencing.<sup>134</sup> If the aim of the practice is to morally reform wrongdoers, there is nothing about that aim that suggests moral reform programmes should release offenders until they are successfully reformed. There is no reason internal to the practice for releasing offenders before they have satisfactorily completed all three stages. The danger is that an offender might stubbornly resist moral reform, leading reformers to inflict an amount of suffering beyond what is ordinarily considered reasonable. It may be that a car thief proves so recalcitrant as to need ten years of moral reform, an amount of suffering far beyond what is ordinarily considered to be appropriate for that crime.

It is certainly a shortcoming of moral reform that it permits indeterminate sentencing, but in the context of deciding whether to punish or reform offenders the problem is of little value to reform's objectors because, *tu quoque*, difficulty with sentencing is not unique to moral reform. For punishment, the concept of desert is too 'underspecific' to guide us in determining exactly how much suffering to inflict on particular offenders for particular offences, and any demand for proportionality between the harm caused and punishment exacted fails to provide substantive guidance as to what sentence an offender should in fact receive. These problems and the appeals to 'external moral constraints' that they prompt will be explained in an attempt to make it clear that all responses to wrongdoing are in equally dire straits when it comes to determining how much offenders ought to suffer.

### §2.1. Internal sentencing principles

There are two ways of determining how much an offender ought to suffer. The first way is to derive an 'internal sentencing principle' from the goal of the practice: wrongdoers are punished as much as is deserved when the goal of the practice is to mete out just deserts, and wrongdoers suffer until they are considered properly reformed when the goal of the practice is moral reform. Internal sentencing principles lead moral reform into

---

<sup>134</sup> Jean Hampton (1984), p. 232.

trouble, since the principle derived permits long periods of suffering for even the smallest offence. In defense of moral reform I point out that internal sentencing is also problematic for retributive punishment.<sup>135</sup>

Desert-based justifications for punishment fail to offer any real guidance in determining how much punishment to inflict on any particular offender. Outside of an established set of rules for doling out rewards and punishments (such as marking systems in education) desert cannot serve the particular aims of just sentencing. Shafer-Landau writes:

There are no determinate moral criteria that can translate the gravity of an offence into even a roughly correlative measure of deserved punishment [and] that kind of underspecificity is fatal to a moral principle.<sup>136</sup>

If one were to conduct an experiment and ask one's peers what they thought an abusive husband deserved as punishment it is almost certain that a consistent reply would not be forthcoming, nor do we have available to us an objective measure of who is correct and who is not. Desert at best provides retributivists with only an *ordinal* sentencing policy: the more serious the crime the greater the amount of suffering deserved. This is not enough since, though we all agree that the rapist deserves a harsher sentence than does the tax dodger, desert cannot provide the non-comparative sentencing measures which scale ordinal sentencing. Retributive punishment cannot rely on desert to explain why an offender received this particular sentence and not another, and sentencing is made arbitrary.

This is a strong objection against the infliction of suffering as a consequence of the perpetration of an offence. For punishment justified by desert or desert bases, there is a specific and not approximate amount of suffering an offender must undergo in order to have met his deserts, settled his debts or relinquished unfair benefits. The amount of suffering undergone by the offender must match the amount of harm suffered by the victim if justice is to be served. Justice demands that sentencing match moral deserts as determined by the wrong committed, and, with desert providing only the weakest approximations, the risk of doing injustice by punishing more or less than is deserved is

---

<sup>135</sup> I discuss only retributivism here, but the same problem affects all punishment practices.

<sup>136</sup> Russ Shafer-Landau (1996), p. 294. Though he uses the word "fatal" here, Shafer-Landau agrees with Mackie that, despite their failings, retributive concepts like desert cannot be done away with.

too high. Retributive punishment aims at serving justice, yet *in every instance* runs the risk of committing serious injustice since we lack a precise measure of when justice is done.

Moral reform, and atonement in particular, is also vulnerable to the underspecificity problem. In atonement there is a demand for 'fittingness', where a fitting atoning act is one that meaningfully communicates a reformee's sincere contrition and appreciation of the wrongness of the offence to the victim, her relations and society. It is not fitting for a reformed rapist to volunteer twenty-four hours of community service and expect to have properly atoned. But how do we judge what are sufficient atoning acts and how do we settle the debate if there is disagreement? Is it fitting for a reformed rapist to volunteer ten or eleven years of public service? One can conduct the same experiment here as with desert: ask one's peers how much atoning ought to be done for bank robbery, perjury or drug dealing and discover that consensus outside of an already established set of rules is too rare. There is no objective way of deciding when an offender has fittingly atoned, yet we must set targets for atonement to avoid seeing an offender suffer too much or too little for his offence. As yet, there are no solutions to the underspecificity problem.

## §2.2. External moral constraints

A second way of determining the amount of suffering it is morally permitted for an offender to suffer for his wrongdoing is to appeal to external moral constraints. These are moral principles that are derived from the basic intuition that the wrongdoer must suffer as a consequence of his wrongdoing. Coming from the same deep source as our retributive ideas, these sentencing principles have an intuitive hold on what we believe amounts of suffering ought to be. Cottingham writes, "These limiting principles have a very strong appeal to all those who share certain fundamental intuitions about justice and individual rights".<sup>137</sup> It is these principles which we expect all practices to abide by. Examples are the principle of proportionality between the harm caused and the suffering

---

<sup>137</sup> John Cottingham (1983), p.425.

of the offender, from which can be derived the ordinality principle mentioned above, and the principle that the same offences warrant the same treatment.

Our most entrenched beliefs regarding the just sentencing of offenders are that offenders should not suffer more or less than their victims (proportionality), sentences increase in severity according to the seriousness of the immoral act (ordinality), and we ought not to treat unequally offenders who are guilty of the same wrongdoing (comparative justice). Our idea of justice will not tolerate *any* practice involving violations of these sentencing principles. A practice like moral reform, which condones immoral sentencing practices under an internal sentencing policy, will appeal to external moral constraints to ensure that offenders do not suffer too much for their offences (proportional suffering), to ensure that recalcitrant offenders guilty of lesser crimes do not suffer more than responsive offenders guilty of more serious crimes (ordinal suffering), and to ensure that the reform process exposes wrongdoers guilty of similar offences to a comparable degree of suffering (comparative suffering).

Moral reform is not alone in appealing to external moral constraints. The heart of the underspecificity problem concerns our inability to establish what amount of suffering on the part of the offender is proportional to the harm she has caused. Rape and incarceration, or car theft and tailored sentences like community service, are acts so different in kind that we cannot know when proportional suffering occurs. There is a lack of commensurability between punishments and offences that has prevented punishment theorists from providing satisfactory answers to questions concerning how much punishment equals the amount of suffering caused by tax fraud or attempted murder.

One solution is to do to the offender what has been done to the victim; thus we should rape the rapist, wound the attempted murderer and molest the child abuser. These treatments of offenders are in fact commanded by the most fundamental retributivist sentencing principle, *lex talionis* ('eye-for-an-eye'). This principle's insistence that the suffering of the offender be commensurable and proportional with that inflicted on the victim is overridden by the imposition of external, 'rights-respecting' moral constraints. To avoid the practice of punishment being one in which punishers engage in those same acts we consider morally pernicious, punishment theorists must appeal to external moral constraints that protect the offender from 'cruel and unusual' treatments or punishments

that are debasing. Both retributivism and moral reform must appeal to external moral constraints to avoid objectionable treatment of wrongdoers.<sup>138</sup>

### §2.3. When moral reform fails

Sentencing is the most problematic issue for responses to wrongdoing and I have tried to show that moral reform and punishment are in equally dire straits in this regard. Moral reform relies on similar sentencing principles to punishment and so, *tu quoque*, is subject to similar difficulties with those principles. But moral reform is also only an *attempt* to teach a moral lesson that may or may not succeed. The indeterminate sentencing problem is partly the result of the nature of moral reform as a practice that does not *force* the realisation of its goal. Moral reform launches persuasive appeals at offenders' capacities as responsible moral agents, but cannot *make* them become "morally autonomous persons freely attached to the good".

In practice we must judge when to stop trying to reform and let the offender go. If I am to retain the claim that moral reform contains as much suffering as is deserved or as repays the offenders debt,<sup>139</sup> then failed moral reform must still mean that the offender has suffered as much as if he had served a retributive punishment. Fortunately, external moral constraints like the proportionality and ordinality principles, those limiting principles that shape all responses to wrongdoing, demand that the attempt at moral reform involves the same amount of suffering as retributive punishment would inflict on a wrongdoer. That is, moral reform should try to reform a rapist for as long as would deliver an amount of suffering that conforms to the proportionality and ordinality principles in the same way that punishment does. If a car thief would serve a four-year sentence then that is the amount of suffering he will experience in our attempt to morally reform him, and no more. The proportionality and ordinality principles are universally limiting, ensuring that moral reform and punishment involve comparable amounts of discomfort whether moral reform is successful or not.

---

<sup>138</sup> The problem deterrence theories of punishment have with internal sentencing is that, as mentioned above, they permit excessive treatments (boiling oil) to accomplish their goal. Deterrence punishment, too, requires limiting by external moral constraints.

<sup>139</sup> As argued for in Chapter Three.

This appears to be an *ad hoc* appeal to retributive sentencing principles in order that moral reform always contain an amount of suffering that is enough to satisfy the basic demand for the offender to suffer. However, the proportionality and ordinality sentencing principles are external moral constraints setting the limits of suffering in *all* responses to wrongdoing that involve suffering. Whether or not the aim of a practice is inflicting suffering, deterrence, the expression of the so-called retributive emotions or moral reform, if the practice involves suffering then we appeal to the proportionality and ordinality principles to discover what amount of suffering is permissible. No practice that violates these principles will be accepted.

In moral reform, indeterminate sentencing raises the worry that failed moral reform involves *too much* suffering. Reform obeys the proportionality principle as an external moral constraint that prevents a recalcitrant offender suffering more than he ought to. Moral reform is not a practice contriving to involve sufficient suffering to satisfy retributivists, it is a practice that must appeal to external moral constraints to prevent the offender suffering more than even the retributivist would demand.

Retributivists subscribe to these same principles of justice, and so any failed attempts at moral reform will involve an amount of suffering which is proportional, or deserved, or which settles a debt and re-establishes the balance of benefits and burdens.

### §3. The unreformable

Moral reform cannot justify inflicting suffering on offenders for the purposes of morally reforming them if it is established that the wrongdoer in question cannot respond in the proper way and be reformed. One is not justified in pursuing a goal one knows cannot be attained, and so we do not attempt to reform the ‘unreformable’.<sup>140</sup> Shafer-Landau identifies it as a problem that “moral educationists” (reformative punishment theorists like Jean Hampton and Herbert Morris) are unable to justify the infliction of reformative punishment on offenders who are “morally incurable”:

---

<sup>140</sup> Of course something must be done with unreformable people to prevent them harming society and themselves. I need not say much about this since I have argued that the unreformable are not the legitimate subjects of any response to wrongdoing. It seems all we can do is keep them in mental hospitals and some prisons.

But this has the unsavoury consequence that those who are irremediably evil should be exempt from punishment. That's hardly what we'd expect from a plausible justification of punishment.<sup>141</sup>

Shafer-Landau's worry is that reformatory practices will hold responsible a different set of people to that which other punishment practices would hold responsible because some kinds of wrongdoers whom we would punish we will not reform. This is a false concern. Those who are reformable must be persons possessing the cognitive and affective capacities of responsible moral agents, and so the unreformable must be those people who lack some or other cognitive or affective capacity, that is, they must be people who are not responsible moral agents.

Unreformables will be psychologically ill people like madmen, people lacking the sensibilities required for appreciating that they are morally responsible for having acted wrongly. They could be evil people, if an evil person is someone who is incapable of coming to understand and be motivated by the good.<sup>142</sup> Anybody who understands that they have done wrong, who was capable of making the distinction between right and wrong at the time of the offence, can come to know or feel the full normative force of the value they have violated. Such people are also morally responsible agents because, unlike children and madmen, they are capable of making choices that are morally informed. We do not regard as a responsible moral agent a person who cannot be guided in action by an appreciation of what a good act is (this seems to be our thinking with regard to young children). On the other hand, if a person can freely choose the good out of an appreciation for its moral value, then they possess the mature cognitive and affective qualities of a reformable person.

The practice of moral reform does not work on people who are not responsible moral agents because it is exactly to the capacities of responsible moral beings that the practice appeals.<sup>143</sup> The morally reformable are the morally responsible. Jonathan Bennett writes, "Accountability is strongly correlated with susceptibility to moral

---

<sup>141</sup> Russ Shafer-Landau (1991), p. 213.

<sup>142</sup> I am not offering a theory of evil. For an evil person to be unreformable yet morally responsible, he would have to understand and appreciate the good, yet be a person who can never be influenced to do good. It is up to those making the objection to prove that such people exist.

<sup>143</sup> See Chapter One.

pressures”,<sup>144</sup> where “moral pressures” refers to the reformatory techniques involved in moral reform. It is no problem for moral reform that there are unreformable people because such people are not morally responsible for their actions, and those who are not morally responsible for their wrongdoings are exempt from *any* response to wrongdoing. Thus, the set of responsible people who would be punished is the same set of people who would be reformed because both moral responsibility and reformability are attributed on the basis of a wrongdoer’s possessing certain cognitive and affective capacities. There is no middle ground. If the basis for attributing moral responsibility to someone was either that they would be responsive to reformatory treatments that are respectful of the Kantian maxim, or that they had acted of their own free will in performing an act they know to be immoral, the set of responsible wrongdoers pointed out on either basis would be the same.

Herbert Morris acknowledged that it followed from his theory of “paternalistic punishment” that the morally responsible and the reformable would be the same, and further:

It also follows that excuses must be recognised and that mitigating factors be taken into account, including as an excuse, of course, reasonable ignorance or mistake of law.<sup>145</sup>

We do not hold responsible persons to whom certain excusing conditions apply, and Morris observes that we do not morally reform these same people. Moral reform targets those who are in need of moral reform, and people who have accidentally become involved in wrongdoing, or who were coerced into transgressing the law, or who were non-negligently ignorant of the consequences of their actions, are not in need of moral improvement. Thus, moral reform does not fail to target those who are morally responsible, and neither does it target those who are not morally responsible because of certain excusing (accident) or exempting (madness) factors.<sup>146</sup>

---

<sup>144</sup> Jonathan Bennett (1980), p. 19.

<sup>145</sup> Herbert Morris (1981), p. 270.

#### §4. Moral reform benefits the wrongdoer

Retributivists believe that reforming an offender is tantamount to benefiting that offender for committing an offence. It is the 'benefit' of acquiring a more sound moral character that retributivists find so objectionable, given that the retributive focus is to make offenders, "worse off than they were status quo ante".<sup>147</sup> The pervading sentiment from the retributivists seems to be that an offender should not be released from a response to his wrongdoing having acquired anything of value. This is a harsh position that runs counter to moral reform's concern to see an offender 'learn a lesson' that helps him move on after his wrongdoing has received redress.<sup>148</sup>

It is a valuable aspect of reform that its subjects are equipped to lead good lives after having taken responsibility for their wrongdoings. Samuel Scheffler emphasises the "resonance" of morality in referring to the pervasive influence of morals within and throughout our social and mental lives.<sup>149</sup> Morals are, "woven throughout the fabric of human emotion and motivation", and, "structure our relations with others", by influencing what we think, feel, and see, and how we behave.<sup>150</sup> Without the correct values, both the desirability and possibility of establishing and maintaining personal relations is in doubt. Moral reform seeks to provide offenders with the moral materials essential to the meaningful engagement with the social world that enriches their and our lives.

The good life for which reformees are prepared is not just moral but holistically satisfying in the sense that a person's own interests are found a place alongside moral constraints. Scheffler writes:

Moral norms do serve to regulate the conduct of human beings, and their content is constrained by their regulative role: they must be capable of being integrated in a coherent and attractive way into an individual human life.<sup>151</sup>

---

<sup>146</sup> For a more detailed account of excusing and exempting factors see Peter Strawson, 'Freedom and Resentment', 1962, pp. 59-80.

<sup>147</sup> Russ Shafer-Landau (1996), p. 295.

<sup>148</sup> It is noteworthy that many contemporary penal systems run education and skills development programmes.

<sup>149</sup> Samuel Scheffler (1992), p. 68.

<sup>150</sup> Ibid. p. 128, 129.

<sup>151</sup> Ibid. p. 4.

Reforming wrongdoers must be a practice guided by an understanding of morality's place in everyday life. Moral reform should help wrongdoers to, "integrate a respect for the worth of others with their naturally disproportionate concern to lead fulfilling lives themselves".<sup>152</sup> The value of a worldly-wise response to wrongdoing like moral reform lies in its ability to temper conflicts between moral constraints and the interests of subjects who live in particular social, political and economic circumstances. Reformees are people who are improved moral agents, better able to follow the letter of the law and less likely to inflict harm on others. The social benefits of moral reform are thus all those connected with having a citizenry who display an acceptable quality of will.

We can see moral reform as an extension of the other-directed use of moral knowledge we find in people concerned to see others become better moral agents and live better lives, where this altruism is motivated by the reformer's own virtuous character and his respect for the capacity for self-improvement possessed by all morally responsible wrongdoers. The humane 'paternalism' of moral reform displays much of the noble virtue that being concerned to strengthen commitment to or improve knowledge of the good has in the case of moral development. We recognise good character in people who like to see that someone learns from her mistakes or learns how better to integrate the demands of moral and non-moral values. To see someone who has wronged you come to display sincere contrition born of a meaningful cognisance of his responsibility for an immoral action bears all the satisfaction of seeing someone become who we think we all should be.

Reformers faithfully view offenders as people who can learn from their past because they retain a core sense of decency and appreciation of the good, but which does not neglect the victim. We ought not to ignore the goodness in offenders because to do so is to identify them with their wrongdoing. In this regard, Kant argued that we are never in a position to wholly identify a wrongdoer with the evil of his wrongdoing, either because of lack of evidence or the danger of biased judgement.<sup>153</sup> Jeffrie Murphy interprets this as a caution that we should not be blind to the potential for goodness that may exist in all but the most evil of wrongdoers.<sup>154</sup> It is to the credit of moral reform that

---

<sup>152</sup> Ibid. p. 129.

<sup>153</sup> Jeffrie G. Murphy, Jean Hampton (1988), p. 148.

<sup>154</sup> Ibid. p. 95.

it compassionately views offenders as persons still possessing a desire to be better, an inner virtue to which the practice appeals. Indeed, the practice depends upon there remaining a core sense of decency in its subjects since moral reform is ultimately an appeal to conscience.

The objection to the ‘benefit’ of moral reform reveals that the retributivist’s attitude toward wrongdoers’ moral improvement is misguided or what I have called “short-sighted”. Moral reform is not done primarily for the offender’s benefit, it is a practice that aims to satisfy the needs and desires of the victim, and accomplishes myriad goals of responses to wrongdoing in trying to achieve its goal. A reformee is aware of his obligations to the victim and society and this appreciation motivates him to offer the victim the sincere moral recognition that can only come from the wrongdoer himself. I have argued that it is the production of a “morally autonomous person freely attached to the good” that leads to the accomplishment of the moral, emotional, social and political goals that recommend punishment practices, including retributivism, to us. It is myopic to object to a practice that simply helps the offender properly respond to her own wrongdoing when so much else is accomplished thereby. All that I have described as the valuable consequences of moral reform come from ‘benefiting’ the offender in the way being objected to. An argument to the effect that moral reform accomplishes all the recognised ends of responses to wrongdoing and best satisfies the victim’s needs is also an argument that it is only by offering an offender the benefit of moral reform that a past wrongdoing receives as full and proper redress as is possible.

Besides, the retributivist might be too hasty in believing that moral reform is a practice that benefits the offender in a way that makes life easier. These objectors have not taken account of the suffering and sacrifice an offender experiences in moral reform. Consider also that the reformee must commit herself to “forbearance in the future”,<sup>155</sup> a commitment that all moral agents know from experience may often be more burden than benefit. Many reformees will emerge from moral reform with only further battles to fight, changing their social circles, their habits, their lives. The work of reformers and the commitment of reformees will be put to the test when the subjects of moral reform return to that social, economic and political context in which their wrongdoing was

---

<sup>155</sup> Herbert Morris (1981), p. 265

perpetrated. Temptations to re-offend will arise for many reformees, and reformers hope that at least they will have made the decision to offend that much more difficult.

## §5. Conclusion

The objection that the offender benefits from moral reform casts the objectors in a bad light as persons cynically blind to the goodness within wrongdoers and myopically blind to the various desirable effects of successful moral reform. Attempting to ‘benefit’ the offender by exposing her to reformatory treatments is a virtuous project that the victim, her relations and society will find satisfying because the production of a “morally autonomous person freely attached to the good” offers them so much of what they want. The worry that the reformable and the morally responsible offender are sometimes not the same person is dispelled by the remarkable fact that a person lacking the capacity to appreciate and thus be moved by the morally good is neither reformable nor morally responsible. The correspondence between reformability and moral responsibility is exact. I have offered a *tu quoque* response to the indeterminate sentencing problem, arguing that punishment practices have equal difficulties with internal sentencing principles. Both sorts of practice appeal to external moral constraints to ensure that offenders suffer as much as our deep intuitions deem just. Since all practices involving suffering must abide by external moral constraints, it remains my contention that moral reform satisfies the demand that an offender suffer an appropriate amount as a consequence of his wrongdoing, whether moral reform is successful or not.

## Conclusion

The moral reform of a wrongdoer is commonly recognised as a good result of a response to wrongdoing, but my task here has been to strengthen this claim so that moral reform is identified as the overriding aim of our practice. The “General Justifying Aim”<sup>156</sup> of responses to wrongdoing ought to be the production of a “morally autonomous person freely attached to the good”, a constructive goal of value to all parties. What is truly attractive about moral reform is that it achieves so many goals of responses to wrongdoing: moral recognition of the victim, the expression and relief of the reactive attitudes, deterrence, and even the satisfaction of the basic demand for the offender to suffer. Punishment practices are less adept at obtaining the most important desideratum of responses to wrongdoing (moral reform) and unable to better moral reform in the accomplishment of any other end.

The moral improvement of a wrongdoer is done using any reformatory technique that is respectful of the Kantian injunction not to treat moral agents as mere means but as ends in themselves. The practice appeals to the cognitive and affective capacities of morally responsible wrongdoers and is rightly described as being an *attempt* at persuading wrongdoers to take responsibility for their conduct. The production of a “morally autonomous person freely attached to the good” cannot be achieved other than through a wrongdoer choosing to focus his capacities as a moral being. Wrongdoers must be left free to work towards their own atonement and moral improvement or merely tolerate the reformatory techniques to which they are exposed.

Moral reform supersedes punishment practices by employing a variety of reformatory techniques, especially treatments designed to assist an offender confront the truth with regard to his culpability and the quality of his will. The mechanisms of casuistry respond to the negative feedback an offender is bound to face, skewing an offender’s interpretation of the meaning of his suffering, and are thus a genuine obstacle to lasting moral improvement. Inflicted suffering alone leaves the offender free to make mistakes or generate rationalisations that excuse his conduct in his own mind. Effective moral reform requires more than punishment.

---

<sup>156</sup> H. L. A. Hart (1968), p. 3.

A reformee offers his own value-laden response to his wrongdoing that includes a measure of suffering sufficient for us to say, "He has suffered as much as deserved". Reformees suffer their just deserts, pay their debts to society and surrender their ill-gotten profits as the consequences of their offences. All this suffering makes it easy to aver that moral reform is a response to wrongdoing that is unpleasant enough to deter potential wrongdoers from committing offences. Moreover, the lack of a rational justification of retributivism and the violation of offender's dignity by deterrence punishment are problems overcome in moral reform. The practice does not aim at these ends, but lends itself to description under the argot of retribution and deters simply because moral reform must be unpleasant to be effective.

Reformative punishment theorists have omitted treatments other than inflicted suffering at least partly because of their loyalty to the basic demand that the offender must suffer as a consequence of his wrongdoing, and punishments ensure that offenders suffer. This fundamental intuition is the root of all punishment practices, epitomised by retributive punishment, and reformative techniques such as rational debate and psychological counseling would seem to do nothing to satisfy it. However, my contention has been that, though these treatments do not inflict suffering, they contribute to the production of a reformee who feels uncomfortable emotions (guilt, shame and remorse) and who is thereby motivated to voluntarily suffer acts of atonement. The modes of suffering brought about by a full programme of reformative treatments satisfy the indelible demand for the offender's suffering.

Atonement and the reformative emotions are particularly meaningful sorts of suffering for their connection to an understanding of the immorality of an offence and the moral status of the victim. Victims, their relations and society derive an impoverished satisfaction from responses to wrongdoing that do not offer the meaningful consolations of an offender's remorse, sincere apology and atonement. It is only through moral reform that a victim receives the moral recognition so valuable for its coming from the person responsible for the original moral affront. Punishments cannot match the moral worth offered by moral reform since what is *taken* from the punished offender is not as valuable as what is *given* by the reformee.

Victims, their relations and society will feel certain appropriate emotions (resentment, indignation, moral hatred) which they desire to express by inflicting suffering on the wrongdoer. Moral reform includes inflicted suffering for exactly the purpose of expressing the moral condemnation manifested by these emotions, since it is only inflicted suffering that carries the message. Furthermore, this inflicted suffering respects Strawson's point that meaningful moral discourse consists in the exchange of emotional responses.

To Strawson's insistence that the emotions of the victim be expressed I have added that responses to wrongdoing ought to relieve these same reactive attitudes. Moral reform is best at providing victims, their relations and society with what their emotions indicate they want, because the practice offers the most persuasive evidence for the beliefs that relieve reactive attitudes, even the so-called retributive emotions. An offender comes to feel the reformatory emotions (especially remorse) and undertakes atoning acts in recognition of the victim, thereby affirming the victim's self-respect and moral status, and affirming a value that has seen violation. In this way, the emotional exchange between victim and offender that Strawson identifies as at the core of meaningful moral discourse is completed. A reformer taking responsibility for his actions is thus the most meaningful, constructive and morally valuable response to wrongdoing.

## Bibliography

- 1) Christopher Bennett, 'The Varieties of Retributive Experience', in *The Philosophical Quarterly*, Vol. 52, No. 207, April 2002: 145-163.
- 2) Jonathan Bennett, 'Accountability', in *Philosophical Subjects: Essays Presented to P. F. Strawson*, Z. Van Staaten (ed.), Oxford: Clarendon Press, 1980: 14-47.
- 3) Hilary Bok, *Freedom and Responsibility*, Princeton University Press, 1998.
- 4) Jennifer Church, 'Taking it to heart: What choice do we have?' in *The Monist*, Vol. 85, No. 3, 2002: 361-380.
- 5) John Cottingham, 'Varieties of Retribution', in *Philosophical Quarterly*, Vol. 29, July 1979: 238-246.
- 6) John Cottingham, 'Punishment and Respect for Persons', in *Law, Morality and Rights*, M. A. Stewart (ed.), 1983: 423-431.
- 7) John Cottingham, 'The Philosophy of Punishment', in *An Encyclopaedia of Philosophy*, G. H. R. Parkinson (ed.), London: Routledge, 1988: 762-783.
- 8) Fyodor Dostoyevsky, *Crime and Punishment*, trans. Constance Garnett, New York: Random House, (19--) (no date available).
- 9) Gena M. Gerard, 'Community-Based Restorative Justice: a Capacity-Building Tool for Confronting Crime', at <http://tcfreenet.org/org/ssco/rj/rjpaper.htm>, 10/09/2003.
- 10) Peter Goldie, 'Introduction', in *Understanding Emotions: Mind and Morals*, Peter Goldie (ed.), Aldershot: Ashgate, 2002.

- 11) Jean Hampton, Jeffrie G. Murphy, *Forgiveness and Mercy*, Cambridge University Press, 1988.
- 12) Jean Hampton, 'The Moral Education Theory of Punishment', in *Philosophy and Public Affairs*, Vol. 13, No. 3, 1984: 208-238.
- 13) H. L. A. Hart, *Punishment and Responsibility*, Oxford University Press, 1968.
- 14) David B. Hershenov, 'Restitution and Revenge', in *Journal of Philosophy*, Vol. 96, No. 2, 1999: 79-94.
- 15) Ted Honderich, 'Punishment, the New Retributivism, and Political Philosophy', in *Philosophy and Practice*, A. Phillips Griffiths (ed.), Cambridge University Press, 1985: 117-147.
- 16) Dennis Klimchuck, 'Retribution, Restitution and Revenge', in *Law and Philosophy*, Vol. 20, 2001: 81-101.
- 17) Thomas Mautner (ed.), *The Penguin Dictionary of Philosophy*, Penguin Books, 1999.
- 18) M. S. Moore, 'The Moral Worth of Retribution', in *Responsibility, Character and the Emotions*, F. Schoeman (ed.), Cambridge University Press, 1987: 179-219.
- 19) Herbert Morris, 'A Paternalistic Theory of Punishment', in *American Philosophical Quarterly*, Vol. 18, No. 4, October 1981: 263-271.
- 20) Herbert Morris, 'Persons and Punishment', in *The Monist*, Vol. 24, 1968: 472-501.
- 21) Iris Murdoch, *The Sovereignty of Good*, London: Routledge, 1970.

- 22) Jean-Paul Sartre, *Being and Nothingness: A Phenomenological Essay on Ontology*, Hazel E. Barnes (trans.), Washington Square, 1993.
- 23) Samuel Scheffler, *Human Morality*, Oxford University Press: New York, 1992.
- 24) Moritz Schlick, *Problems of Ethics*, David Rynin (trans.), Prentice-Hall: New York, 1939.
- 25) Russ Shafer-Landau, 'Can Punishment Morally Educate?' in *Law and Philosophy*, Vol. 10, No. 2, 1991: 189-219.
- 26) Russ Shafer-Landau, 'The Failure of Retributivism', in *Philosophical Studies*, Vol. 82, 1996: 289-316.
- 27) Nancy Sherman, *The Fabric of Character*, Oxford: Clarendon Press, 1999.
- 28) Peter Strawson, 'Freedom and Resentment', in *Proceedings of the British Academy*, Vol. xlviii, 1962: 59-80.
- 29) Roger J. Sullivan, *Immanuel Kant's Moral Theory*, Cambridge University Press, 1989.
- 30) Shelley E. Taylor, Jonathan D. Brown, 'Illusion and Well-Being: A Social Psychological Perspective on Mental Health', in *Psychological Bulletin*, Vol. 103, No. 2, 1988: 193-210.
- 31) Steven Tudor, 'Accepting One's Punishment as Meaningful Suffering', in *Law and Philosophy*, Vol. 20, 2001: 581-604.
- 32) Bernard Williams, *Shame and Necessity*, University of California Press: California, 1993.

33) Bernard Williams, *Problems of the Self*, Oxford university Press, 1993.

34) 'The US Penal System: Restorative and/or Retributive Justice?' in *Woodstock Report*, No. 61, March 2000, at <http://www.georgetown.edu/centers/woodstock/report/r-fea61a.htm>, 20/10/2003.

