Georgia State University
Digital Archive @ GSU

Philosophy Theses | Department of Philosophy

4-16-2008

Retribution Requires Rehabilitation

Joseph Q. Adams
josephquintonadams@yahoo.com

Follow this and additional works at: http://digitalarchive.gsu.edu/philosophy_theses

Recommended Citation
Adams, Joseph Q., "Retribution Requires Rehabilitation" (2008). *Philosophy Theses.* Paper 35.

This Thesis is brought to you for free and open access by the Department of Philosophy at Digital Archive @ GSU. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

RETRIBUTION REQUIRES REHABILITATION

by

JOSEPH Q. ADAMS

Under the Direction of Dr. Andrew J. Cohen and Dr. Christie Hartley

ABSTRACT

Herbert Morris argues in his influential retributivist paper, “Persons and Punishment,” that criminals deserve punishment because their actions represent an unfair distribution of benefits and burdens in society. The proper distribution of benefits and burdens is important, in part, to restore law abiding citizens’ confidence that others will follow the law. In this paper I show that Morris's argument for why criminals deserve punishment morally requires us to set up an institution of rehabilitation in addition to the institution of punishment. Such an institution is morally required because neither pure punishment systems nor punishment systems that incorporate quasi-rehabilitative aspects have ever worked to uphold the necessary confidence that Morris tells us law abiding citizens must have in order to protect the social order. Moreover, we cannot abandon Morris's appeal to the duty to maintain social order without also abandoning a plausibly Morrisian framework.

INDEX WORDS: Right to punish, Deterrence, Retributivism, Free rider, Justice, Justification of punishment, Master’s thesis, Morality of punishment, Hampton

RETRIBUTION REQUIRES REHABILITATION

by

JOSEPH Q. ADAMS

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2008

Copyright by
Joseph Quinton Adams
2008

RETRIBUTION REQUIRES REHABILITATION

by

JOSEPH Q. ADAMS

Committee Chairs:	Andrew J. Cohen Christie Hartley
Committee:	Andrew I. Cohen William A. Edmundson

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2008

DEDICATION

For Mom, Dad, Aimee, and Donald: your love and support make everything possible!

ACKNOWLEDGEMENTS

It would be very difficult to overstate the amount of help I received from my thesis advisors at Georgia State University, Dr. Andrew Jason Cohen and Dr. Christie Hartley. They held my academic hand as I learned to crawl and then walk from project planning to thesis defense. I tried to indicate where their ideas were especially crucial in forming my argument but I am sure I have failed to mention places where their questions, comments, and criticisms inspired in me new avenues of thought. Thanks to Dr. Andrew I. Cohen for an unforgettably meaningful and productive directed reading on the justification of punishment which grounded me in the debates.

My academic debt as it relates to this thesis extends well beyond the inner circle of those directly involved. The whole Department of Philosophy at Georgia State University allowed me to grow to a proficiency of writing and explore depths of philosophy that I could not have predicted. Thanks also to my undergraduate philosophy professors at Spring Hill College, without whose inspiration I would not be a philosopher. Dr. Carl D. Kobelja turned me on to the subject, Dr. K. Joanna Forstrom encouraged me to pursue it, and Dr. Melvin J. Brandon oversaw my first major philosophical project.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

CHAPTER

1 THE ETHICAL BACKGROUND 4

Consequentialism and Deontology 5

The Consequentialist Justification of Punishment 8

The Deontological Justification of Punishment 10

2 MORRIS'S RETRIBUTIVISM 13

3 THE ADVANTAGES OF MORRIS'S ARGUMENT 17

4 HAMPTON'S CRITIQUE AND A RESPONSE 21

5 WHY MORRIS'S ARGUMENT REQUIRES REHABILITATION 26

6 RESPONDING TO OBJECTIONS 35

Objection #1: Rehabilitation Does Not Work 36

Objection #2: Deny that Morris Needs to Restore the Social Order 42

7 CONCLUSION 44

8 WORKS CITED 45

Retribution Requires Rehabilitation

By Joseph Q. Adams

In the recent film, "The Brave One," Jodie Foster's character is brutalized and her fiancée beaten to death with a pipe during a late night walk in Central Park. She responds to the trauma by protecting herself with a gun. At first she finds herself in various situations where she kills career criminals to defend herself and others. Later she begins "walking the streets at night," ostensibly to diversify her poetic radio show, but, in fact, seeking out such situations. Finally she gets the nerve to hunt down her own attackers. The theater I was in cheered as she shot her attacker in the face.

By almost all philosophical accounts, Foster's revenge is not appropriate.[1] However, many thinkers identify a certain reaction to wrongdoing that, perhaps, underlies the cheers in the theater. The cheers were, at least partly, an indication of satisfaction that Foster's attackers got what they deserved. Philosophers carefully distinguish such satisfaction, what they call the retributive emotion, from revenge.[2] Revenge is the desire to lower a guilty person. The retributive emotion is the desire to restore proper moral order. In other words, revenge is satisfaction in another person's pain while retribution is satisfaction in justice, properly construed.[3]

---

[1] But see David Hershenov's "Restitution and Revenge." Hershenov believes punishment is justified partly because it satisfies the victim's desire for revenge. But even Hershenov stops short of endorsing vigilante revenge—under his account the victim still gets revenge through the state. Another writer endorses a model which rejects the role of the state and permits victims to individually exact not revenge, but restitution. See Randy Barnett's "A New Paradigm of Criminal Justice." So, even the theories that get us close to justifying Foster's revenge do not in fact do so upon close examination.

[2] I adopt Jean Hampton's approach to retributivism for the distinction here. See pages 137 and 138 of "The Retributive Idea" in *Forgiveness and Mercy* where she clearly distinguishes the two responses.

[3] This explanation of the cheers needs a little more analysis. Thanks to my 1010 class for challenging me on the limited explanation offered above. For example, if most people were really cheering just because Foster's attackers got what they deserved then presumably they would cheer just as much if, in the end, the police arrived to arrest them. I suspect, however, that such a turn of events would not have elicited the same reaction. Given this, some readers may conclude that revenge is not distinguishable from retribution. Many philosophers, including the early Hampton, considered the distinction dubious. See Hampton's "Moral Education Theory of Punishment." However,

Many thinkers defend the institution of punishment based on the thought that criminals get what they deserve through punishment, as opposed to the thought that punishment deters criminals. This paper critiques one such retributivist. Herbert Morris thinks that criminals deserve punishment because their action represents an unfair distribution of benefits and burdens in society (Morris, 477-478). To fail to redistribute benefits and burdens is to fail to respect persons, including the criminal, the victim, and society at large. Moreover, a just system of punishment is uniquely equipped to correct this maldistribution (488-490). *In this paper I show that Morris's argument for why criminals deserve punishment morally requires us to set up an institution of rehabilitation in addition to the institution of punishment*.[4] Such an institution is morally required because neither pure punishment systems nor punishment systems that incorporate quasi-rehabilitative aspects have ever worked to uphold the confidence that Morris tells us law abiding citizens must have in order to protect the social order. Moreover, we cannot abandon Morris's appeal to the duty to maintain social order without also abandoning a plausibly Morrisian framework. In short, either we must adopt a model that is not strictly punitive in order to meet the demand set out by Morris's theory or else we need to abandon the notion that

---

there is also a third kind of reaction, a "primitive defensive anger which any of us, as a species of animal, will feel toward our attackers" (Hampton and Murphy, 54). The cheers might have just been a satisfaction in expressing this "bite back" response (54). Further, the cheers could have been satisfaction in the expression of one (or more) of the three kinds of hatred outlined by Hampton: malice, spite, or moral hatred. This story is meant only to introduce readers to retribution using an experience with which all could identify: watching a movie. I note here though that an analysis of the class of "retributive emotions," which would involve a close reading and perhaps critique of Hampton's conceptual analysis on the topic, is the subject of future work. Perhaps experimental philosophy could be brought to bear on the subject.

[4] I am using the word "rehabilitation" to refer to a unique way of responding to wrongdoers, a response that is not deprivation-based and acts in addition to or perhaps instead of punishment. I explore the details of in Chapter Five. But I need to make clear here that I do not intend the meaning of "rehabilitation" as it is usually employed in the punishment literature. Usually, "rehabilitation" refers to the process of bringing about a change in criminal beliefs such that they now avoid crime because they think breaking the law is wrong (Ten, 7, note 1). But this usage is neutral with respect to the form of that process—even severely painful responses can theoretically bring about this change. I do not want to remain neutral with respect to form: rehabilitation, for the purposes of this thesis, should be distinguished from traditional modes the state uses to respond to law breaking: imprisonment, fines, execution, and the like. It should connote something like moral education including, as I outline in Chapter Five, subsidized access to vital resources upon release, such as housing and transportation.

punishment restores the fair distribution of benefits and burdens throughout society. If we abandon that notion, retributivism will need some other basis.[5]

In Chapter One I define punishment and survey both the deontological and consequentialist justifications. As a retributivist, Morris relies on a deontological justification. Thus I spend some time characterizing the brand of strong, Kantian retributivism which Morris endorses.[6] In Chapter Two I present Morris's specific argument. In Chapter Three I show the central advantage of his argument: it morally requires us to punish as a matter of respect. Chapter Four highlights the main drawback. Morris's characterization of law breakers as obtaining an unfair benefit of non-restraint implausibly implies that being restrained from violent crime is a rational burden (Murphy and Hampton, 114-117). After defending Morris against this charge, I show, in Chapter Five, why rehabilitation of a certain kind is morally required by

[5] This is a surprising outcome in light of the fact that redistribution of benefits and burdens has a powerful appeal to Morrisian retributivists who seem largely unconcerned with a non-punitive response to wrongdoers. This unconcern seems deeply flawed to me. See pages 26 through 29 of R. A. Duff's "Penal Communications: Recent Work in the Philosophy of Punishment." Duff cites seven philosophers who think "punishment functions to restore that fair balance of benefits and burdens that a crime disturbs" (26). Morris was the first to offer this kind of retributive argument. This surprising outcome might incline some retributivists, like Michael Moore, to deny that any such Morrisian arguments are appropriately retributivist. Retributivism properly understood by Moore appeals strictly to the notion that punishment is intrinsically valuable because it is deserved (Moore, 179). As Jean Hampton says, this kind of pure retributivism claims that "it is supposed to bedrock intuition that, at the very least, those who are not guilty ought not to suffer pain, and more positively, that those who are guilty deserve to suffer in proportion to the pain they have caused" (Murphy and Hampton, 113). However, this appeal alone does not seem to satisfy. Many readers will want an explanation of desert. Again, I quote Hampton on this point: "those suspicious of retribution are generally unpersuaded by the kind of defense" offered by pure retributivists. So Morris should be credited with defending retribution against the skeptics by offering an explanation of why criminals deserve punishment beyond "bedrock intuition." Perhaps Moore would respond that, if Morris's explanation of desert requires us to investigate punishment's ability to restore social order, then we will have to find another explanation. To fully respond to this worry I would need to show that (a) bedrock intuitions are not satisfactory and desert can be explicated, but that (b) any explication will require us to setup rehabilitation. If my argument in this thesis is cogent I will have shown why one highly-favored explication of desert requires rehabilitation.

[6] Even though punishment under Morris's framework must protect the social order in the sense that it must restore law abiding citizens' confidence that others will continue to follow the law (through proper distribution of benefits and burdens), it is appropriate to classify Morris as a Kantian retributivist. The restoration of confidence is important insofar as it helps ensure proper respect for all participants through the fair distribution of benefits and burdens. Punishment is still a matter of respect—respecting the rights of law abiding citizens to be confident in the ability of our institutions to insure that state responses to criminal wrongdoing justly redistribute benefits and burdens. Thanks to Professor Christie Hartley for clarifying this for me.

Morris's argument. Finally, in Chapter Six I respond to the central objection that rehabilitation, like punishment, cannot meet the moral demands of Morris's theory.

Chapter One: The Ethical Background

Philosophically, the most interesting part of punishment is its justification. Because it involves harm, punishment needs justification. Hugo Adam Bedau summarizes the reasons that punishment needs justification:

> Since punishment involves intentionally inflicting deprivations on persons by someone with authority to do so, and since the deprivations themselves are typically not unlike the harms that crimes cause (fines are like theft, imprisonment like kidnapping, etc.), punishment has generally been thought to need justification (Bedau, 732).

We justify punishment by giving good reasons to believe that punishment is, under a certain set of conditions, either morally permissible or morally required. This chapter surveys some of the traditional reasons philosophers offer for the claim that punishment is sometimes an appropriate response. I focus on strong Kantian retributivism, which is the view that punishment is sometimes morally required.

Before we can examine different justifications for punishment, however, we need to get clear about what, exactly, we are justifying. We need a working definition of punishment. I adopt H. L. A. Hart's definition of punishment (Hart, 4-5). For my purposes, punishment will be understood as the deliberate state-sanctioned infliction of harm distributed by officials to law breakers who have been found guilty. The definition has five parts. First, punishment involves some sort of deprivation or loss, especially the removal of certain rights. Second, punishment is a response to someone who broke a law. Third, it is administered to people who have been

found guilty of breaking the law.[7] Fourth, it is a human institution. Finally, it is an action sanctioned by the state.

There are two kinds of ethical theories that we generally use when worried about the justification of an institution like punishment: consequentialism and deontology. We may think of ethics as that field of philosophy concerned with proper human action. Narrowed in this regard, ethics is the study of which human actions we should undertake in various sorts of situations.[8] Consequentialism and deontology provide two different frameworks from which to evaluate the rightness and wrongness of human action. These theories provide two different families of reasons to believe that a certain action is right or wrong. While most comprehensive theories of punishment incorporate elements of both kinds of reasons,[9] I explain each ethical theory separately.

Consequentialism and Deontology

I begin with consequentialism. Consequentialists rely on an important distinction between the right and the good. For consequentialists, goodness can be evaluated from an agent-neutral perspective (Darwall, Consequentialism, 2).[10] To understand agent-neutrality, imagine a planet like Earth in every respect except that it lacks humans. Sunsets, according to this ethical theory, could still be good. They are harmonious states of affairs even though no humans enjoy

[7] There is a distinction between "being guilty" and "being found guilty." The definition uses "being found guilty" in order to avoid begging the question against consequentialist theories. Because consequentialists can hold that we can acceptably punish innocent people (who would still be *found* guilty by the judiciary), we do not want to beg the question by limiting punishment to that class which is *actually* guilty.

[8] David Brink observes that ethical theories can apply to "virtually any object of moral assessment, including actions, motives, individual lives, institutions, and moral codes" (Brink, 381). This thesis focuses on the moral assessment of the legal institution of punishment. However, at this point, it is easier to approach the ethical theories as a whole from the perspective of individual action. There are complications when we apply a theory of action to institutions but the principles outlined herein should remain the same.

[9] H. L. A Hart, for example, is known for promoting dual theories. See the first chapter of his *Punishment and Responsibility*. Bedau characterizes the hybrid theory as the consensus among philosophers. See his entry "Punishment" in the *Stanford Encyclopedia of Philosophy*. By the end of this thesis we see why a traditionally consequentialist problem has the potential to undermine Morris's fundamentally deontological theory.

[10] I rely on Stephen Darwall's introduction to his book *Consequentialism* for my discussion in this paragraph.

them. Consequentialists believe that the right action (the ethical action) is that action which—on balance—brings about the most (real world) goodness, whether agent-neutral or agent-relative, as *evaluated from an agent-neutral perspective*. The point is that "there are values that are prior to morality," or prior to rightness (1). For consequentialists, rightness is a matter of promoting goodness.

One noteworthy feature of consequentialism understood in this broad way is that it remains neutral with respect to the nature of a given action (Brink, 381). Whatever action produces the most good is the right action. Actions are not in themselves intrinsically right or wrong. As a result, consequentialism has the theoretical ability to endorse actions that strike us as, at the very least, strange. In the case of punishment, consequentialists can, at least in theory, recommend that we should punish innocent people.[11] If, among the various alternatives, punishing an innocent person causes the most good, then punishing an innocent person is the right thing to do. The fact that punishing innocent people strikes many of us as morally wrong is one of the central objections to a pure consequentialist justification of punishment.

While they remain neutral with respect to the nature of an action which promotes the good, consequentialists disagree about "what constitutes the relevant consequences of an act" (Ten, 3). One consequentialist may call an action right where another calls it wrong (or remains indifferent) because the action causes a state of affairs the first considers good and the second does not.

A popular form of consequentialism is utilitarianism. According to some utilitarians, the right action is the one that, among the various alternatives, results in the highest level of total utility. Utilitarians divide themselves according to their understanding of utility. Call one who

[11] See C. L. Ten's excellent discussion of the issue of punishing the innocent in "Fantastic Examples and Moral Principles" on pages 18-36 in Chapter 2 of his *Crime, Guilt, and Punishment*.

defines utility in terms of happiness, where happiness is the presence of pleasure and the absence of pain, a classical utilitarian.[12] The right action is the one which maximizes pleasure over pain. One could also define utility in terms of preference satisfaction. In this case the right action is the one which satisfies as much and as many preferences as possible. There are other ways to define utility as well, but for the purposes of this chapter, I review the consequentialist justification of punishment from the perspective of the classical utilitarian.

Unlike the classical utilitarian and other consequentialists, deontologists deny that moral evaluations can be made from an agent-neutral perspective (Darwall, *Deontology*, 1). We cannot determine the right action from an agent-neutral perspective by reference to the agent-relative or agent-neutral goods that result. For deontologists some actions are morally required or prohibited–a right or wrong action to perform–regardless of whether they bring about good states of affairs. Some deontologists believe certain constraints on behavior are absolute—never to be overridden—while others believe these constraints can be overridden by, for example, duties of special care (McNaughton and Rawling, 425). Both of these camps agree, however, that at least some constraints cannot be overridden by the value of promoting the good. According to Darwall, we discover these constraints by reflecting on the nature of the action itself (Darwall, *Deontology*, 4-5).

Darwall uses the simple example of betrayal to explain deontology. "The moral duty not to betray *one's* friends, deontologists claim, is agent-relative [and is] irreducible to any agent-neutral values" (2). I should not betray my friend even if my betrayal causes many good consequences. Betrayal is just a bad thing to do.

[12] This is the utilitarianism of Jeremy Bentham. "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered" (Bentham, 12). While there are other classical utilitarians who have a different definition of happiness, Bentham's version suits my purposes.

The Consequentialist Justification of Punishment

Consequentialists usually justify punishment because it promotes the good by reducing crime (Ten, 7-8). Understood from the classical utilitarian perspective, reducing crime is good because crime is inversely correlated with the general happiness. All else being equal, rational people tend to experience more happiness when there is less crime.

Utilitarians identify two general ways punishment theoretically reduces crime: incapacitation and deterrence. Punishment straightforwardly reduces crimes committed by individuals through physically constraining or killing those individuals. The state ensures that, at least during the time of incapacitation (which is indefinite in the case of execution), the criminal herself cannot commit crime among the general population.[13]

Deterrence is a negative incentive (Blumstein, 19). A positive incentive increases the benefits of a behavior in order to encourage that behavior. A negative incentive increases the cost of a behavior in order to discourage that behavior. Punishment raises the cost of a behavior in terms of time, money, freedom, and humiliation in order to discourage law breaking.

Theoretically, punishment deters both the individual criminal and the general public. With "special deterrence," punishment discourages repeat individual criminal behavior (Nagin, 95). In this case, the punishment of a criminal reduces the chance that the same criminal will commit another crime. Punishment motivates criminals to stop breaking the law for two types of

[13] Of course, the criminal can still commit a crime in jail or prison. Depending on your view of criminal-on-criminal crime and (if you consider it legitimate crime) the volume of it, incapacitation will work more or less to reduce overall crime. I should note that the sometimes-cited "criminogenic effect" has been refuted by the empirical data (see the report of the panel cited in note 14 below). The criminogenic effect is the theoretical idea that criminals increase their crime skills while incapacitated with other, more experienced criminals. Prison thus increases the amount of crime committed by released criminals by making them more "hardened." If the criminogenic effect actually takes hold, utilitarians would have to factor it in when justifying punishment (it might offset benefits obtained through incapacitation and deterrence). However, the evidence suggests that neither criminogenic nor special deterrent effects hold.

reasons. They may be convinced that breaking the law is wrong,[14] or they may just be afraid of more punishment. Regardless of the motivating reasons behind a criminal's change of behavior we can say that the criminal is deterred in the sense that state has produced a stoppage in her criminal behavior.

With "general deterrence," punishment discourages criminal behavior in the general public. As Daniel Nagin puts it, "the imposition of sanctions on one person may demonstrate to the rest of the public the expected costs of a criminal act, and thereby discourage criminal behavior in the general population" (96). In this case, the punishment of a criminal reduces the chance that a potential future criminal will commit crime. [15]

I will not go into them here, but I should note that there are other ways, besides crime reduction, that punishment might promote the good. Punishment may satisfy the desire for revenge. If we are Machiavellian, punishment may appropriately cause fear in the population. Punishment may produce religious converts (at least seeming-converts). Any of these reasons to endorse punishment would be consequentialist, although, arguably, none of them would be part of good arguments.[16]

---

[14] The literature includes these types of criminals—the ones who stop for moral reasons as a result of punitive responses—as a part of the class of "rehabilitated" criminals. But this is not how I use the term rehabilitation so I just incorporate this set of criminals into the deterred set.

[15] Some notes on the effectiveness of these theoretical benefits. According to the "Report of the Panel," an extensive study commissioned by the government and conducted by academics during the 1970's, special deterrence is ineffective. General deterrence appears somewhat effective. Incapacitation appears somewhat effective. Deterrence resulting from hybrid punitive and non-punitive methods also appears ineffective. Ten, who compares the "Report of the Panel" with other studies basically sides with the report when he says that "We see that the evidence is perhaps more hospitable to the claim that punishment has some general deterrent effect and some incapacitative effect than it is to the claim that it has individual deterrent effect or that it rehabilitates offenders" (Ten, 12). One thing we can conclude from this evidence is that punishment works better as a symbolic mechanism than as an experienced reality. Can we find ways to create the symbolism without actually harming people? And would that be a morally better situation? These are questions for future research.

[16] Except perhaps Hershenov's argument from revenge. See note 1 above. However, Hershenov's argument is comprehensive so it involves both consequentialist and deontological justifications.

The Deontological Justification of Punishment

I now turn to the deontological justification of punishment. Deontologists are associated with the retributive justification of punishment. I present the retributive justification of punishment as traditionally found in the literature, with a focus on strong retributivism. Then I explicitly draw the connection between retributivism and deontology.

Retributivists usually argue that punishment is justified because it is deserved. Michael S. Moore characterizes retributivism generally as follows:

> We are justified in punishing because and only because offenders deserve it. Moral culpability ("desert") is in such a view both a sufficient as well as necessary condition of liability to punitive sanctions. For a retributivist, the moral culpability of an offender also gives society the *duty* to punish. Retributivism, in other words, is truly a theory of justice such that, if it is true, we have an obligation to set up institutions so that retribution is achieved (Moore, 179).

Call this form of retributivism the "strong form." Under strong retributivism we *must* punish criminals. By contrast, weak retributivism, which I will not deal with, merely requires that we have the *right* to punish. Under strong retributivism, the criminal's past action creates a demand that society is obligated to meet. Fundamental to retributivism is the idea of "moral culpability" or desert. Someone deserves something when that person's past action places an obligation for further action on the part of some other party. For example, under normal conditions, my winning a professional tennis tournament places an obligation on the tournament committee to award me a trophy and prize money. In this case, the desert is either explicitly outlined by the tournament committee prior to the tournament or implicitly demanded by social convention. On account of my winning, I deserve a trophy and prize money. Similarly, with punishment, my committing a crime places an obligation on the state to punish me. Like the tennis tournament, the state's obligation is laid out beforehand in the law. On account of my crime, I deserve punishment.

Strong retributivism, as understood by Moore, can be defended in a variety of ways. Authors differ in their account of criminal desert. I focus here on strong Kantian retributivism whereby criminals deserve punishment because we must respect their choices as moral agents.[17] Other arguments include intuitionism,[18] arguments from social cohesion,[19] and annulment theories.[20]

Whatever version of retributivism we adopt, we need an understanding of desert or moral culpability. Contained in the idea of desert is the idea that criminals choose to commit a crime. I

---

[17] The brand of *Kantian* retributivism we have in mind here should be distinguished from *Kant's* retributivism. Deontological justifications that appeal to the equal inherent worth of all as a reason to believe that criminals deserve punishment may be called "Kantian," without adhering to Kant's own theory of punishment. Kant's own theory of punishment has drawbacks which I cannot cover in detail. Moreover we should distinguish both of these from interpretations of Kant's retributivism. One such interpretation involves Kant's advocacy of strict proportionality between crime and punishment, the doctrine of *lex talionis*. *Lex talionis* demands that criminals suffer a form of punishment which exactly matches the harm inflicted by the criminal. "But what is the mode and measure of Punishment which Public Justice takes as its Principle and Standard? It is just the Principle of Equality [which] may be rendered by saying that the undeserved evil which any one commits on another, is to be regarded as perpetrated on himself. Hence it may be said: 'If you slander another, you slander yourself; if you steal from another you steal from yourself; if you strike another you strike yourself; if you kill another, you kill yourself'" (Kant, 196). Jean Hampton points out that one way to interpret this doctrine is to demand, for example, that if a criminal throws acid on a victim's face, the criminal should have acid thrown on his face (Murphy and Hampton, 136). Hampton believes, and I agree with her, however, that a "sophisticated retributivist" will deny strict proportionality while maintaining the Kantian doctrine that no human being, including a criminal, should receive inhumane treatment (136). Kant himself backs off the claim of strict proportionality. While he says that "there is no Equality between the crime of Murder and the retaliation of it but what is judicially accomplished by the execution of the Criminal," which seems to imply that an exact match is needed, he qualifies this statement saying, "his death, however, must be kept free from all maltreatment that would make the humanity suffering in his Person loathsome or abominable" (Kant, 198). In other words, we cannot endorse a Kantian theory which allows the state to treat persons in ways which violate their personhood. "A good retributivist generally wants limits placed on punishment sentences" (Murphy and Hampton, 135). How, then, does a sophisticated retributivist still give the criminal what he deserves? "There are methods of [punishment] that may be severe if the criminal has gravely mistreated his victim, but which are nonetheless not so harsh that they are identical either to the treatments which are considered permissible only for entities far less valuable than human beings or to treatments which actually make them less valuable" (136). In other words, we can insure proportionality (which is demanded by a Kantian retributivism) without committing ourselves to strict proportionality. A commitment to strict proportionality can only by called "Kantian" by inspiration.

[18] See J. L. Mackie's "Morality and the Retributive Emotions."

[19] Emily Durkheim and Patrick Devlin, for example, believe that criminals must be punished because they have threatened society's moral fabric, a fabric upon which law and order vitally depend. Morris's theory is partly of this type. See brief summaries of Durkheim and Devlin in Chapter Two of Kadish and Schulhofer's *Criminal Law and its Processes*.

[20] See Chapter Four of *Forgiveness and Mercy* where Hampton argues that punishment is deserved because it annuls the crime's status as evidence that the criminal has more worth than the victim. See also Hegel's *Philosophy of Right*, Sections 95 to 103. He says that punishment cancels or negates the crime itself, where the crime itself is a null negation of the victim's right (Hegel, 123). Hampton takes herself to be clarifying Hegel, but the differences between the two are worth exploring in future projects. I believe this kind of basis for retributivism, while still flawed, is the most promising retributive argument.

only deserve something if the action which is thought to give rise to that desert is done voluntarily.[21*] Desert does not arise if the action is non-voluntary. A mother does not punish a child who strikes another while having a seizure because the child with the seizure does not choose to cause harm.[22] Hence the child does not deserve punishment.

Kantian retributivists believe we must punish criminals because we must respect their choices as moral agents. We start with the idea that a criminal's action is morally culpable and sufficiently bad so as to merit punishment by the state (as opposed to mere disapproval by others). This culpability creates an obligation on the part of the state to punish—we say the criminal deserves punishment. The source of this culpability is the criminal's voluntary choice—she chooses to commit a crime. As Kantians, we must respect the choice of the criminal because all human beings have the same moral worth and it is their ability to choose which differentiates them from, for example, animals and inanimate objects (Morris, 490).[23] As Morris says, "When we 'look upon' a person as less than a person or not a person, we consider the person as incapable of rational choice" (490). This is the Kantian idea: all rational beings, in virtue of their

---

[21] * A special note on intentionality. I have chosen to define a culpable action as one which is done "voluntarily" or "by choice." I want to stay neutral with respect to what constitutes "intentional." I assume that whatever (you think) intentionality *is*, it is required for culpability. So the only people to whom this discussion might be untenable are those who think intentionality does not exist. But perhaps even those readers will allow that voluntary actions exist.

In legal terminology an individual must have the appropriate mens rea (in addition to the actus rea) in order to be punishable. One must be in the appropriate frame of mind (mens rea) in addition to actually doing the crime (actus rea). By defining culpability as I have done I subscribe to what Kadish and Schulhofer call the "general" or "all-encompassing" sense of mens rea. According to them, "one way the requirement of mens rea may be rationalized is on the common sense view of justice that blame and punishment are inappropriate and unjust in the absence of choice" (Kadish and Schulhofer, 203). They endorse, however, the "special" sense of mens rea where "mens rea refers only to the mental state required by the definition of the offense to accompany the act" (203).

[22] I do not mean to take a position with respect to the debate about metaphysical freedom. By "voluntary" action I do not mean to imply a commitment to any position in that debate. It is a placeholder for the conditions for moral responsibility. My purpose here is just to acknowledge that the debate on metaphysical freedom itself and its intersection with moral responsibility is wide open. While I find it fascinating and think it vital for metaphysicians to more freely dialogue with moral theorists, I do not have time to engage this debate here.

[23] Human *worth* should be distinguished from human *value*. The worth of a human being is that which determines the level of respect due in virtue of personhood, defined by Kant in terms of a human's capacity to reason. The value of a human being is that which determines the level of respect due in virtue of some other characteristic, for example, one's painting skill. A Kantian must hold that all persons have equal worth but is not committed to the idea that all persons have equal value. I can be better than you *at tennis*, but I cannot be better than you *as a person*. A non-Kantian, on the other hand, might believe that persons can have various levels of worth.

personhood or rationality, have equal moral worth. To treat them as not responsible for their rational choices is to treat them as non-persons. There is an absolute deontological constraint on our actions: we can never treat persons in such a way that disrespects their rationality. To *not* punish the criminal is to fail to respect the criminal as a rational chooser. Respecting rational beings requires respecting their autonomy. This means holding them responsible for actions in which they voluntarily engage. Thus, under strong Kantian retributivism we must punish the criminal because to do otherwise disrespects the personhood of the criminal (violates the absolute constraint) by failing to respect the criminal's choice.

Chapter Two: Morris's Retributivism

In this chapter I focus on Herbert Morris's retributivist justification for punishment. Morris's argument is a brand of Kantian retributivism. I present Morris's argument in this chapter. In Chapter Three I show its main advantage, while Chapter Four covers the central disadvantage.

Here is Morris's argument in a one-liner: criminals must be punished because criminals must "pay" for their free ride in a mutually advantageous, cooperative system of self-restraint (Morris, 477). Social order is maintained by rules which restrict persons from engaging in a certain set of actions. These restrictions are just because they "provide advantages for all" (492). People who break the law get the benefit of not restraining themselves and thus violate the mutual advantage. They cheat. It is unfair for criminals to receive this benefit when others have properly restrained themselves. Thus, criminals must be burdened with punishment in order to

re-balance the distribution. Morris seems to think a system of just punishment is the only way to restore the balance of benefits and burdens without disrespect (488-490).[24]

Morris starts with a certain idea about the nature of society and law. The law defines a sphere of non-interference for individuals. In exchange for assuming the burden of following the law, individuals receive the benefit of non-interference by others. Society is, thus, mutually advantageous.[25] We limit our own freedom in exchange for a certain level of security.

> Applying to the conduct of [people] are...rules that prohibit violence and deception and compliance with which provides benefits for all persons. These benefits consist in noninterference by others with what each person values, such matters as a continuance of life and bodily security. The rules define a sphere for each person, then, which is immune from interference by others. Making possible this mutual benefit is the assumption by individuals of a burden. The burden consists in the exercise of self-restraint by individuals over inclinations that would, if satisfied, directly interfere or create substantial risk of interference with others in proscribed ways (477).

The first thing to see here is that Morris presents the purpose of law. Laws, or rules, exist essentially to protect individuals from "violence and deception." We burden ourselves by restraining from certain actions (we follow the law) and receive, in exchange, the benefit of "life and bodily security" resulting from a corresponding restraint on the part of our fellow citizens. Such a system of cooperation or "mutual benefit" can be maintained only if individuals take on a

---

[24] There are, however, a variety of ways in which we might restore the balance. Morris entertains therapy which he rejects outright (480-488), forgiveness (i.e. pardoning) which he endorses only if used rarely (478,488), shaming systems which he rejects after brief discussion (488), and of course punishment which he endorses (477-480). I find his discussion curious because he never mentions rehabilitation, a fifth option. In Chapter Five I show why Morris's commitment to balancing out benefits and burdens requires him to endorse rehabilitation. One way to reconcile this curiosity is to read Morris as assuming that his system of just punishment incorporates rehabilitation (he never fleshes out the details of the form of punishment). I think it is pretty clear, however, that like most retributivists he conceives of his system of just punishment as mainly deprivation-based, concerned only with harming, rather than rehabilitating, offenders: "the deprivations visited upon any person are justified by that person's having violated the rules" (480). See Chapter Five for a full discussion of these points.

[25] I do not mean to suggest that society is contractarian or contractualist. We may still have a duty to obey the law in light merely of reciprocity and this seems the best way to read Morris. So while laws are "mutually advantageous" they are not, on Morris's account, necessarily "agreed to" in the sense that social contract theorists assert. I am friendly to social contract theory but I wish to remain agnostic about it here since Morris himself seems to be agnostic. In short both social contract theorists and their opponents may endorse Morris's justification of punishment, albeit for different reasons.

burden of self-restraint. Law abiding citizens take on the burden of *not* committing crime. They pay the cost of denying themselves a certain set of behaviors. The benefit is that law abiding citizens are assured that others also restrain themselves in a similar way. Thus citizens maintain the health of society by sharing equally the burdens of self-restraint and the benefits of non-interference. When we break the law we create an imbalance by throwing off our burden of self-restraint.

The most straightforward examples of Morris's cooperative system of self-restraint come from property laws.[26] Suppose that my farmland is situated such that it would be significantly more convenient for you to drive across my land to reach your golf course. But your farmland is situated such that it would be significantly more convenient for me to drive across your land to reach my yoga class. Neither of us want our crops to suffer car damage, even though both of us would prefer (and have no qualms about) driving across the other's land. We pass a law against trespassing. In exchange for the burden of restraining myself from driving across your land I receive the benefit of your not driving across my land. Both of our crops prosper and we each find the arrangement advantageous.

Criminals are people who throw off their burden of self-restraint. Suppose I am in a big rush one morning. I am late for my yoga class, and we are learning an important new move. I decide to break the mutually advantageous trespassing law and drive across your farmland. In this case I throw off my burden of self-restraint. I break the law.

We must punish people who break the law for two reasons. First, we must safeguard the confidence law abiding citizens have that others will continue to restrain themselves (477-478).

[26] Hampton points out that Morris's argument works best with property crimes in her critique of Morris on pages 114-117 of *Forgiveness and Mercy*.

Second, if we focus just on the maldistribution between the victim and the criminal, the maldistribution just demands to be rebalanced because it is unequal (478).

The farmland example readily shows the first reason. When you inspect your crops the next day and notice my tire marks, you have two possible responses. You may continue to follow the law, hoping that my actions are a one-time offense. In this case I have the benefit of your self-restraint without the burden of my self-restraint. On the other hand, you may decide to throw off your own burden of self-restraint. As Morris puts it, law abiding citizens' "disposition to comply voluntarily will diminish as they learn that others are with impunity renouncing burdens they are assuming" (477). My breaking the law reduces your confidence in my ability to comply and thus tempts you to break the law yourself. Before long, we destroy each others' crops. Punishment is thus partly justified as a "mechanism designed to prevent a maldistribution" which would otherwise undermine the social order (477). Sanctions work to "induce compliance with [the law] among those who may be disinclined to obey" (477-478). In short, the deterrence of criminals through punishment works to uphold the social order by insuring equal distribution of benefits and burdens.

The unequal distribution of benefits and burdens also appears to require punishment purely on grounds of fairness. Even if your tendency to obey the law voluntarily is not diminished by my trespassing on your farmland, punishment is still required. By breaking the law, the criminal obtains the benefit of not restraining herself. Punishment is an act of the state which burdens criminals by depriving them of important goods. Punishment is thus seen as a way of balancing the scales.[27] With punishment the criminal receives a burden which counters the benefit of non-restraint. "Matters are not even," Morris tells us, until the criminal's unequal benefit is counter balanced through the burden of punishment (478). In our farmland case, my

[27] Thanks to Professor Andrew J. Cohen for putting it this way.

driving across your farmland will trigger a punishment (a burden) that counter balances the benefit I gained from not restraining myself in order to get to yoga class on time.

Chapter Three: The Advantages of Morris's Argument

In this chapter I highlight the central advantage of Morris's theory. I show why Morris's theory is consistent with an egalitarian theory of human worth. Such a commitment in Morris's case not only makes punishment morally permissible, but, like Kant's legal philosophy, makes punishment morally required.

What was a controversial claim in Morris's day has become the central advantage of his theory: the idea that punishment actually respects the criminal. "There is a tradition...not notable for its present vitality... [whose] adherents have urged that justice requires a person to be punished if he is guilty. Sometimes—though rarely—these philosophers have expressed themselves in terms of the criminal's *right to be punished*" (475-476). Here Morris clearly places himself in the strong retributivist camp. We see that "justice requires" punishment because we have a duty to respect the rights of others—in the criminal's case, we have a duty to respect her right to be punished.[28] So a Kantian theory of human worth appears to be at the heart of Morris's argument: respecting people as persons entails punishing them (assuming they are guilty).

Why is Morris's commitment to a Kantian theory of human worth advantageous? Most importantly, such a commitment gives us a strong justification for the institution of punishment.

[28] Depending on our view of the nature of rights, characterizing punishment as a right of the criminal's may imply that this right can be waived. Thanks to Professor William A. Edmundson for pointing this difficulty out to me. I cannot get into the debates about the nature of rights here. However, I think it is clear that Morris means that this right cannot be waived. So, it may have just been better for Morris to leave out talk of the criminal's right to be punished. His thought that the criminal has a right to be punished is clearly intended to support the idea that justice requires us to punish culpable criminal wrongdoing, no matter what (except for the rare use of pardon).

If punishment respects the criminal, then we have a powerful reason to consider it morally appropriate. We can "rest easy" in the confidence that we have given good reasons to do things to criminals that the law would otherwise forbid: take their money and forcibly confine them, for example.

Such a commitment also comes with certain safeguards that we find morally appealing. These include prohibitions against punishing innocent people and the ability for lawbreakers to defend themselves. These safeguards flow naturally out of a justice system that protects reason and autonomy. I address each of these safeguards in turn.

If punishment respects criminals, it must distinguish guilty lawbreakers from non-guilty lawbreakers. I cannot debate here the various ways in which a justice system might go about doing this. I just note that, respecting criminals means we must find some way to distinguish the guilty from the non-guilty. Furthermore, most philosophers tend to agree that a proper justice system allows for excuse, justification, and mitigation. Morris's theory demands such safeguards. "A person has not derived an unfair advantage if he could not have restrained himself or if it is unreasonable to expect him to behave otherwise than he did" (478). So Morris's theory has the advantage of requiring these proper safeguards, but I leave it an open question as to how to ground excuse, justification, and mitigation.

Besides prohibiting the punishment of innocent persons, Morris's system also endorses the attractive principle of innocent until proven guilty.[29] Punishment of the truly guilty respects

[29] This is a further advantage that should be distinguished from, although it arises out of, the prohibition on punishing the innocent. In theory, one could prohibit punishing the innocent and consider individuals guilty until proven innocent, or just "neutral" until proven guilty, or even neutral until proven innocent. Because the practice of determining guilt or innocence combines the many imperfections of human beings (bad memory, poor cognition, over- and under-zealous prosecution and defense, just to name a few) and the blunt instrument of the state (ambiguous and unclear laws, limited police and judicial resources, etc.) a theory that recognizes these limitations has a distinct advantage. Because these epistemic limitations as to the truth of guilt exist, we should endorse an innocent until proven guilty standard as long as we consider the moral cost of punishing innocents greater than the moral cost of freeing guilty persons.

all parties involved (victim, society, criminal) while punishment of the innocent respects none of these. Morris recognizes that in the pursuit of giving criminals what they deserve we might accidentally punish an innocent (479). Such a situation is unjust under Morris's theory because innocent people have not gained an unfair benefit. As such, Morris endorses a theory which emphasizes that citizens are innocent until proven guilty. Such an emphasis is proper given that the point of law in the first place is to provide a certain protected sphere of non-interference (479). Punishment of the innocent would pierce this sphere. Of course it is still possible, even under the principle of innocent until proven guilty, that we may pierce the sphere by punishing the innocent but falsely proven guilty. Whenever we punish the innocent, we transgress the very principle which gives rise to law in the first place. An innocent until proven guilty principle is thus the lesser of a variety of evils, and to that extent, a virtue. The other options (guilty until proven innocent) make such transgressions more likely.

In a punishment system that treats persons as ends in themselves, lawbreakers are also given the space to defend themselves. Criminals are entitled to hold on a variety of beliefs about both their behavior and the law which they broke. They can maintain their innocence both before and after being found guilty. Alternatively, they may admit to their deed, but maintain that it is not blameworthy or should not be illegal. Criminals are free to express their opinions about the law which they broke, which makes room for civil disobedience. All of this is in line with a commitment to free speech, an important principle in a liberal democracy.

I should note here that to remain consistent Morris will have to admit that non-criminals, as well, are free to express their opinions as to the criminal's behavior and the law which they broke. Obviously, Morris must permit individuals to advocate for or against the lawbreakers and for or against the morality of the law which they broke. I focus, in Chapter Five, on the further

question of what the state may or must do, in addition to punishment, to rehabilitate the criminal. Let me briefly anticipate the discussion here. The relatively uncontroversial (but still helpful) point is that nothing in Morris's theory prohibits us from setting up an additional institution parallel to the institution of punishment that seeks to change criminals through reasoned and respectful argumentation. The more interesting point is that, on Morris's justification, we must set up such an institution. So it seems strange that he omits a discussion of rehabilitation, especially in light of the fact that he acknowledges other alternatives to punishment and therapy.[30]

I structure the rest of this paper as follows. In Chapter Four I present Jean Hampton's critique of Morris's argument which shows the main disadvantage of his theory: he seems to say that people find abstaining from violent crimes to be a rational burden. I respond to Hampton's critique on behalf of Morris before exposing, in Chapter Five, my own critique. Accepting Morris's characterization of criminals as throwing off a burden (contra Hampton) requires us to implement an institution which tries to rehabilitate the criminal. I think it is important to see that retribution on Morris's terms demands that we rehabilitate criminals, at least in addition to (but perhaps instead of), harming them.

Chapter Four: Hampton's Critique and a Response

It will be useful at this point to standardize Morris's argument in order to see clearly how Hampton's critique differs from my own. Morris's argument is actually composed of two arguments. The first establishes that criminals obtain a benefit of non-restraint (this argument is contested by Hampton). The second establishes that punishment removes (or balances) this benefit (and is thus justified). My contention is that the evidence developed by Morris's

[30] See note 24 above.

Argument 2 requires us to set up an institution of rehabilitation—that premise 5 must include rehabilitation as well as punishment. We can standardize the arguments as follows:

Argument 1

1. Society sets out mutually advantageous rules (laws) which prohibit persons from engaging in certain acts (defining a "sphere" of non-interference).
2. Criminals are people who (culpably) break the law; they are people who pierce the sphere of non-interference.

Therefore,

3. Criminals get an unfair benefit: they throw off their own burden of self-restraint while enjoying others' restraint.

Argument 2

3. Criminals get an unfair benefit.
4. We must remove[31] the criminal's benefit because he takes advantage of law abiders.
5. Punishment removes the benefit that criminals get from lawbreaking by forcing them to incur an additional burden.

Therefore,

6. We must punish the criminal.

The characterization of the unfair benefit of non-restraint under Argument 1 has been criticized by Hampton and others (Murphy and Hampton, 114-117).[32] While it is plausible to say that rational people are burdened by restraining themselves from property crimes, it is not plausible to assume that rational people are burdened by restraining themselves from violent

[31] "Remove" is a little misleading. Morris does not suggest that punishment makes the unfair benefit disappear, but just that it adds a burden which then equalizes the distribution. The reader may think of punishment as removing the unfairness resulting from the lawbreaking.

[32] See R.A. Duff's "Penal Communications: Recent Work in the Philosophy of Punishment," pages 26-27, for a review of those authors who agree with Hampton.

crimes like murder and rape. Hampton points out that Morris's system assumes that crime is something we all wish we could do. To characterize law abiding citizens as assuming a burden of self-restraint is to say that they want to commit the crime. In order for my following the laws against violent crime to be burdensome, I must first be motivated to break that law. Rational citizens are not so motivated. In the farmland example, I want to drive over your land. With property crimes it is easier to see why not committing the crime may be a burden—an action that I want to do but choose not to do. With violent crime, the act does not tempt the normal citizen such that not doing it imposes a cost: "If I refrain from murdering you, am I really imposing a cost on myself for the sake of creating a collectively advantageous social order?" (Murphy and Hampton, 116). Hampton thinks that Argument 1 fails: the murderer does not deserve punishment because she threw off a burden of self-restraint. It is unreasonable to say that rational people find it burdensome to avoid murder. There must be some other way to link punishment to the wrongdoing of the criminal (117).

I am not so sure, however, that Hampton's critique of Argument 1 is decisive. I think Morris is right to see law breaking as a benefit. I also think Hampton is right to say that rational citizens do not consider restraining themselves from violent crime a burden. Perhaps we can have it both ways and respond to Hampton as follows. Rex murders my sister, a mentor of mine of whom I am very fond. Rex's crime still creates an unfair burden on me, an additional victim, which needs to be counter-balanced through the punishment of Rex. I suffer an unfair burden, although it is not the burden of restraining me from killing my sister (or anyone). Rather the unfair burden is the post-crime burden of living without my sister whom I cherish and living under the fear of murderers. Rex creates a burden which he unfairly forces upon me. I do not

have my sister and am now concerned that others may murder me. I am also concerned about murder in general. Society appears to me "unfair."

Under this interpretation of Morris's argument, punishment still, at least partially, restores the equilibrium of benefits and burdens without forcing us to admit that avoiding murder is a rational burden. We see punishment not so much as balancing out the additional unfair benefit the criminal gains but rather *compensating* the unfair burden which Rex forces on his victims, including me and others who knew my sister or learned of her murder through the media or otherwise. Punishment removes this burden, in part, by restoring our confidence in the system of cooperation. We do not fear murderers (as much) anymore.[33]

It may be objected that this response to Hampton does not stay true to Morris's argument. After all, Morris explicitly cites the "unfair advantage" gained by "renouncing what others have assumed, the burden of self-restraint" (478). Thus, to truly respond to Hampton we will have to identify what unfair benefit the criminal gains without falling into the trap of saying that restraining one's self from murder (and other heinous crimes) is a rational burden.

Consider the following.[34] In a diverse democratic society, individuals are at liberty to pursue their own conception of the good within certain constraints. In the case of Rex we say that part of his conception of the good entails murdering my sister. However, Rex's full pursuit of his good is outlawed by a legal prohibition of murder. Such a law is consistent with the principles free and equal citizens in a democracy would collectively will.

---

[33] The qualification "as much" here shows that this account needs more detail than I have space to address. In particular, it is not clear how this account satisfies the demand for proportionality. Ideally, punishment would make me whole by restoring me to my pre-crime state of not fearing murderers. In my own experience, however, human psychology does not permit such a wholesale restoration. I will likely still fear murderers even after Rex is punished because of the trauma imposed by Rex which may never fully subside. For example, my experience of being mugged produced in me a fear, which I still have, of two or more youths gathering together in urban spaces at night. The punishment of the individuals who mugged me would not, it seems to me, eliminate my fear (to my knowledge, the punishment of my assailants never came to pass).

[34] Thanks to Professor Christie Hartley for bringing this line of argumentation to my attention.

Like Rex, the rest of us are constrained in some way in the pursuit of our own conceptions of the good. While my own conception of the good may not involve murder, it may involve other things which are prohibited. In continuing with the farmland example, my conception of the good includes driving across your farmland when I am late for yoga. Such an act is prohibited because it violates the principles of justice to which I am bound.[35] Thus, like Rex, I am constrained in my full pursuit of my conception of the good.

Most, if not all, individuals are at least partially constrained in their pursuit of their conceptions of the good. Such restraint takes on a variety of forms according to the variety of conceptions of the good (whereas I like driving across farmlands, someone else likes petty vandalism). Inasmuch as law abiding citizens restrain themselves from full pursuit of their conceptions of the good, they are burdened. Rex did not restrain himself from an act which violates the principles of justice. Law abiding citizens, on the other hand, do restrain themselves. The reason we must punish Rex is not because he did not carry the burden of restraining himself from murder in particular, but because unlike the rest of us, Rex did not restrain himself from that pursuit of his conception of the good which is in violation of the principles of justice.

Under this interpretation of Morris, punishment still restores the equilibrium of benefits and burdens without forcing us to admit that avoiding murder is a rational burden. The only rational burden is avoiding that pursuit of our own conception of the good which violates the principle of justice. Additionally, we stay true to Morris's theory by identifying the additional benefit gained by the criminal. Punishment eliminates the unfair benefit of Rex's unadulterated pursuit of his conception of the good. We must punish Rex because the rest of us still restrain

[35] I cannot comment here on how or why I am so bound, although Morris seems to suggest the Rawlsian view. Because one reason we must punish according to Morris is to insure fairness, those who follow the law seem to be upholding fairness. Fairness thus seems to demand that I sometimes stop short of fully pursuing my conception of the good.

ourselves from certain actions which fall within the purview of *our* conception of the good, even though the prohibited actions we would like to do are not identical to Rex's prohibited actions.[36]

I suppose there are objections to this response. We could say, for example, that not all violations of the principles of justice are equal. In other words, the response on behalf of Morris assumes that everyone finds it equally difficult to restrain themselves from full pursuit of their conceptions of the good—Rex finds it as difficult to restrain from murder as I do from driving over your farmland. This seems questionable. Still, I suspect that there might be other ways to respond to Hampton's claim. We could, for example, adopt George Sher's approach and say that the unfair benefit is the additional freedom obtained (Sher, 82). Yet, there are objection's to Sher's approach as well (Dolinko, 546-548). Thus it might be instructive to put aside Argument 1 for the time being and see if there are any problems with Argument 2, especially in light of the fact that many find Morris's account compelling. For if Argument 2 fails, any arguments that criminals obtain an unfair advantage become irrelevant no matter how decisive. Alternatively, if Argument 2 has a non-obvious implication (such as requiring rehabilitation) then those who find Morris's account compelling will still have something to think about (but some previous fans of Morris might begin to question Morris's account on the basis of this implication).

[36] This account, while plausibly Morrisian, may still be objectionable to Morris. Morris focuses on the unfair benefit gained by the criminal and appeals to the notion that this benefit should be counter balanced by adding the obvious burden of pain and suffering attendant on punishment. Any burdens incurred by the criminal's victim or society are largely ignored. I think, however, that we need some kind of solution to the lack of attention Morris pays to victims and I think we can stay within the Morrisian account to do so (my rehabilitation system also has this advantage). The lack of concern for the victim is part of the reason Hershenov embraces the idea that victims' satisfaction of revenge helps justify punishment. He says that Morris "fails to locate the main creditors or the proper currency for the payment" (Hershenov, 80). In other words, Morris tells us who deserves to pay for the crime (the criminal, obviously) but does not tell us who gets restored by the criminal's payment. So I think Hershenov is right about his critique but wrong about his suggestion. It seems that the victim needs some restoration and if punishment does that through the satisfaction of revenge or through the account I outlined above then we still justify punishment as restoring fair distribution of benefits and burdens.

Chapter Five: Why Morris's Argument Requires Rehabilitation

I believe that Morris's theory requires us to set up an institution of rehabilitation because legal punishment fails to uphold the social order. If I am right, then this is a critique of Morris in that he failed to address something that he should have addressed.

Before we begin, I need to clear up a potential confusion around the meaning of "rehabilitation." The standard usage of "rehabilitation" is to "refer to cases in which the offender, after serving a sentence, no longer commits crimes because he believes that criminal behavior is wrong and not because he fears punishment" (Ten, 7, note 1). So authors generally distinguish rehabilitation from special deterrence on the basis of the motivating reasons for the criminals' subsequent law abiding behavior. Rehabilitated criminals stop their crimes because they think breaking the law is wrong. Deterred criminals stop because they want to avoid more punishment. But both stop committing crimes, so they are both "deterred" in a broader sense.

On this standard usage, we could use "rehabilitation to theoretically refer to any method which gets the criminal to stop for moral reasons. So traditional modes of pain and suffering associated with punishment—even cruel and unusual punishment—could be called rehabilitation if they convince the criminal to stop breaking the law because she thinks breaking the law is immoral. This is not the usage of rehabilitation that I intend for this thesis. I use "rehabilitation" to refer to softer responses to wrongdoing, responses that are not straightforwardly painful in the sense that punishment is straightforwardly painful.[37] Whereas the traditional literature uses "rehabilitation" to distinguish the reasons criminals stop committing crime, I use "rehabilitation" to mark off a different kind of response to wrongdoing.

[37] My rehabilitation may still be painful in a sense. As I say below I have in mind a kind of moral education where we enter into a dialogue with the criminals and try to expose any false belief systems that may be underlying their criminal behavior. There may be some pain associated with the dissolution of long-held, false beliefs. But this is not the same kind of pain associated with the traditional forms of punishment: incarceration, execution, and fines.

So when I say that Morris's account requires us to set up an institution of rehabilitation I am not saying that we have to do whatever it takes to get the criminal to stop for moral reasons. Rather, I am saying that, since punishment does not do the job that Morris's theory needs it to do, we must setup a different kind of institution. A softer, rehabilitative institution seems an obvious choice. I will flesh out the details as my argument requires. Briefly, I envision a place of moral transformation or moral education. It is not therapy, where we assume the criminal is sick and force him to change. Rather, it is akin to interpersonal debate where the criminal is given the space (in a controlled environment) to share his views about why he chose to commit the crime. His views would then be contrasted with the point of view which believes that following the law (especially as it relates to violent crime) is right or proper. The hope is that the criminal would come to see the error of his false beliefs and choose to follow the law voluntarily in the future. His beliefs would thus be rehabilitated or transformed to reflect the truth—he would become a better moral agent.

Recall that part of the reason we must restore the benefits and burdens through punishment, according to Morris, is that it restores law abiding citizens' confidence, or trust, in the system of non-interference.[38] As I will show shortly, the consensus of empirical data suggests that a punishment system fails to restore this trust. Furthermore, our current system is not designed to facilitate that goal. Since we are still required to restore this trust (without completely abandoning a Morrisian account) we must explore other options and rehabilitation seems a good candidate.

While empirical data on the effectiveness of punishment is messy in many areas, there seems to be one area of consensus: criminals do not stop committing crime as a result of punishment of any type that we have employed in the modern era (except capital punishment).

[38] See Chapter 2 pages 15-16 above.

Empirical research suggests this as an uncontroversial assumption.[39] Recidivism rates are unequivocally high. In note 14 above I cited C.L. Ten who discussed a report by the National Academy of Sciences which concluded that the evidence suggests that individuals are not deterred by punishment. Punishment does not appear to stop the individual from repeating his behavior (although it might deter the general population). In one of the papers commissioned by that report, Daniel Nagin says that

> On balance, recent evidence tends to suggest that special deterrence, which observationally is difficult to distinguish from other forms of 'rehabilitation,' is not operating. This tentative conclusion is suggested by the apparent invariance of recidivism to any type of special rehabilitative program. The figures suggest that recidivism rates cannot be affected by varying the severity of punishment, at least within acceptable limits (Nagin, 95-96).

Here Nagin uses "rehabilitation" in the standard usage—to simply refer to criminals who stop committing crime as a result of their incarceration, whether for moral reasons or out of fear. The "special rehabilitative programs" which he mentions are correctional treatment strategies, not the kind of moral education response that I have in mind. In short, recidivism rates are high. Criminals do not stop committing crime as a result of being punished.

Punishment does not restore the law abiding citizens' confidence in the system of self-restraint. Punishment does not (as it is supposed to according to Morris) work to support law abiding citizens' "disposition to comply voluntarily"; rather, citizens' disposition to comply

[39] One could respond by saying that while real world punishment schemes have failed to stop criminals, the empirical research does not address a hypothetically improved punishment scheme. Thus punishment without rehabilitation might get criminals to stop committing crime and thus fulfill the demands of Morris's social order premise. Here is my response: punishment-only schemes have been tried in a variety of forms, from grotesque (quartering) to anesthetic (lethal injection), for a much longer period of time than rehabilitative schemes. Admittedly it is always open for us to explore more effective forms of punishment. However, in light of the fact that many variations of one family of state responses (punishment) have been tried in the real world in a variety of ways over the course of centuries to little or no avail, it would seem at least prudent to investigate more fully another family of state responses, rehabilitation. The rehabilitative responses do not have quite as miserable a track record since we have only begun using them in minor forms during the last half of the twentieth century.

“diminishes” because “others are with impunity renouncing burdens they are assuming” (477). In other words, citizens begin to lose confidence in the mutually advantageous system because punishment does not ensure that criminals will, in the future, use self-restraint. Punishment does not keep individuals from renouncing their burdens. Thus punishment fails to maintain the social order by failing to provide “some assurance” that we “will not be assuming burdens which others are unprepared to assume” (477).

Part of maintaining the system of non-interference means helping to ensure that criminals stop violating the system of self restraint. It is reasonable to think that law abiding citizens want to ensure as best they can that criminals do not violate the system. A “punishment only” system conflicts with Morris's account because it does not respect the rights of law abiding citizens who are entitled to some assurance that criminals *stop* violating the system of self restraint, as opposed to the mere assurance that the benefit of non-restraint will continually be counter-balanced by the burden of deprivations. In short, we cannot simply respond to the criminal act by depriving the criminal, unless that deprivation helps insure that the criminal stops committing crime. Our response to criminal wrongdoing, if it is to maintain the system of non-interference, must do more than merely keep “matters even” (478).

Because punishment fails to meet the duty to restore confidence in the system, we have a duty to explore other kinds of state responses that differ significantly from the way we have been responding to criminal behavior. In other words, we must explore other options that appear to us to have a better chance of stopping criminal behavior. Rehabilitation seems a good candidate.

I am unclear whether rehabilitation is required in addition to, or instead of, punishment. I am open to mixed model here where rehabilitation works in concert with punishment according to the two reasons that Morris offers to punish. The first reason—the duty to maintain social

order—turns out not to be a reason to punish but instead requires us to setup rehabilitation. The second reason—the duty to restore the equilibrium of benefits and burdens—may still require us to punish. Punishment-only may make sense for some offenders where the punishment will ensure that they cannot offend again, where keeping matters even does in fact work to protect the social order as Morris thought was the case for all crimes. For example, revoking a doctor's license for bad practice may be sufficient to ensure he does not repeat. For other criminals, rehabilitation-only make sense, as may be the case for young offenders. For still other offenses, such as violent crimes that do not call for total incapacitation (i.e., life in prison or execution), a combination of punishment and rehabilitation may be required. In cases of negligent homicide, for example, we may believe strongly in burdening the individual with punishment. Also, however, we will want to insure that this kind of thing does not happen again upon release and will need to educate the criminal about his false moral beliefs. The negligent killer will need to see that he inappropriately values, say, wealth, at the expense of harming innocent people.

Rehabilitation, as I conceive of it, is permitted by Morris because it maintains respect for criminals as demanded by the Kantian framework. Rehabilitation thought of as moral education is permitted by the Kantian framework because we are allowed to try to convince those who break the law that their behavior is contrary to society's rules and likely immoral. Nothing in Morris's theory seems to deny the permissibility of a state institution of rehabilitation, as long as that institution does not take to treating criminals as patients, rather than agents. As we saw, Morris is committed to respecting the role of reason and must allow criminals and others to express their beliefs about criminal behavior and law. This commitment seems to extend to us the moral permissibility of setting up an institution of rehabilitation.

We take it for granted that there are ways of attempting to convince people of the immorality of their behavior, and implore them to change, without treating them as animals to be conditioned or as sick individuals needing to be cured. Indeed, higher education is founded upon the belief that we can and should try to disabuse each other of false beliefs. There should be no special prohibitions against arguing about moral beliefs, including the criminal's generally false moral belief that breaking the law is justified in some way. Thus Morris must be thought to consider it morally required to establish an institution dedicated to helping criminals see the moral error of their ways.[40]

If rehabilitation is to be conducted like interpersonal dialogue with attendant parameters ensuring respect, there is a question about whether eligible criminals would participate. They may simply refuse to talk. Alternatively, they may be convinced that their behavior is unjustified but refuse to change because they see no reason to be moral. They may be amoral. If the criminal refuses to talk, deontological constraints block us from using certain modes of force (we cannot torture them or give them a truth serum). If they are amoral the situation seems even worse—no amount of convincing will work to change their behavior.

I can assuage the worry that criminals will not talk, but I may not be able to eliminate it. We should keep in mind that rehabilitation as I conceive of it is inherently respectful. It is

[40] We might be tempted to change the definition of punishment so as to require the state to tell the criminal that he is punished out of respect for his choices. I would prefer this revision, but such a move, I am afraid, would be vulnerable to the "definitional stop" objection outlined by H. L. A. Hart. Hart cautions us against loading our definition of punishment with, for example, prohibitions on punishing the innocent. Prohibiting the punishment of innocents taps a strong intuition and so appears to be an innocent addendum. Who would object to defining punishment as an act limited to those who are guilty? But as Hart shows us, to do so is to beg the question against consequentialist justifications to whom it is a theoretically open question as to whether we should punish the innocent. "[The definitional stop] would prevent us from investigating the very thing which modern skepticism [about the justification of punishment] most calls in question: namely the rational and moral status of our preference for a system of punishment under which measures painful to individuals are to be taken against them only when they have committed an offence" (Hart, 6). Requiring the criminal to understand that punishment respects his choices would beg the question in favor of my thesis since imparting that knowledge would require some measure of education (and abstract knowledge of it is not enough). I *do* think that we have to impart this knowledge, but not because the definition should include such a requirement. Rather, I think we must impart this knowledge for the reasons outlined above.

inviting, in a way. Some wrongdoers may never have even had an experience of being respected by being asked to share their views. There is a certain incentive to participate in rehabilitation. This suggests that there will be more willingness to participate than we might initially think. I admit, however, that this would be a problem, especially in light of long-standing suspicion of the criminal justice system in general.

The other thing to keep in mind, though, is that we are still limiting the criminal's liberty through rehabilitation. In this way, we are still accounting for the unfair benefit the criminal received by throwing off his burden of self-restraint.[41] The difference is that we are just choosing to do things differently during the time in which the criminal's liberty is restricted. Instead of punishing him and throwing off our duty to uphold law abiding citizens' confidence, we instead engage in rehabilitation.

My response to the question of amorality is stronger. If a criminal believes that he has no reason to be moral, then he is denying that morality can make a claim on him. He may be completely unable to understand how he could be under the dictates of morality. In this case we have a sociopath on our hands and rehabilitation probably does not make sense.[42] This person appears to have lost his status as a moral agent. Our deontological constraint based on respect for persons would be lifted. On a practical level, however, we should be careful with leveling this charge in light of our limited capacities to mark out such a class. Antisocial behavior, even grossly antisocial behavior like serial killings, does not obviously merit inclusion of the

[41] Thanks to Professor William A. Edmundson for pointing this out to me.

[42] See Jean Hampton's account of immorality in "Mens Rea," where she says that defiance of moral authority is the ground upon which we declare an action culpable. To be in defiance, however, one must know that authority makes a claim on her. But, "someone who cannot understand the authority of morality cannot be expected to conform her actions in accord with it when the moral dictates oppose her desires" (Hampton, "Mens Rea," 94). These are people that completely lack consciences. I do not like allowing for this theoretical possibility, but I admit I have no argumentative resources at this point. I am more inclined to deny that anyone could ever be completely amoral. But I cannot defend this point here.

individual in the class of amoral people (since behavior is a notoriously difficult indicator of mental states).

If, on the other hand, the criminal understands that morality is authoritative over him—that morality forbids his criminal act—but just refuses to obey the authority of morality, then I think this person is in a position of self-hatred and rehabilitation *is* the appropriate response. How can a non-sociopath come to believe that he has little or no reason to obey morality? One obvious answer is when the person sees himself as inherently evil. Hampton gives us reason to believe that when one begins to worry that he is inherently evil one is, to that extent, engaged in self-hatred: "If the wrongdoer fears that the victim is right to see him as cloaked in evil, or as infected with moral rot, these fears can engender moral hatred of himself" (Murphy and Hampton, 86). The correct response to a self-hater is not more punishment which will reinforce the worrisome message that the criminal is inherently evil, but just the kind of rehabilitative approach I have outlined. Getting the person to stop hating himself may be the crucial element in getting him to stop committing crime.

Besides the problems of participation stemming both from refusal to talk and refusal to understand or obey morality's authority, we may object to rehabilitation on other grounds. Because rehabilitation or education is often confused with therapy one might read Morris as prohibiting rehabilitation on the grounds that it disrespects wrongdoers. Morris cautions us against any therapeutic "tendencies of thought" (482). So we might read him as saying that any system which tries to engage in the business of changing the criminal (even if done in an acceptably respectful environment) sets the stage for treating criminals as patients rather than agents.

> Therapy and punishment differ widely in their implications. In bringing out some of these differences I want again to draw attention to the important fact that while

> the distinctions [between voluntary and non-voluntary actions] we now draw are erased in the therapy world, they may, in fact, be reintroduced but under different descriptions. To the extent they are, we really have a punishment system combined with a therapy system (482).

Perhaps the term "rehabilitation" is just the type of "different description" of therapy of which Morris is afraid. The idea is that if we try to convince the criminal of the immorality of her behavior we may come to see her behavior as a symptom of some disease and thus disrespect her by failing to look upon her as a rational chooser. Rehabilitation cannot do this within the Kantian framework.

At best this objection merely reminds us to be careful in how we define and implement rehabilitation. But there is nothing in moral education—trying to convince criminals of the immorality of their behavior through reasoned discussion—that should raise the red flag of therapy.

Another objection to the claim that rehabilitation is permissible under Morris's framework might say that rehabilitation gives an unfair, additional *benefit* to the criminal. The state's response is supposed to burden the criminal in order to balance things out. If rehabilitation makes the criminal better, then it seems we are giving a real benefit to the criminal that he does not deserve. We need to think about what justifies the extra state expenditure on criminals for moral education.

We need to distinguish, however, between broader social benefits (such as subsidized healthcare, education, and social security) and the benefit of breaking the law which requires a retributive response.[43] Wrongdoers who undergo rehabilitation are not incurring any additional benefit of not complying with the system of self-restraint. They have not gained another unfair benefit in the relevant sense.

[43] Thanks to Professor Christie Hartley for this line of thought.

The benefit that *is* in play here is the restoration of law abiding citizens' confidence—a consideration about which Morris cares deeply. If it works to convince the criminal of his erroneous beliefs, then rehabilitation restores our confidence that criminals will be law abiding, in a way that punishment does not. We are interested in rehabilitating the criminal because it ensures social order. Furthermore, the victim herself is at least partially restored along the lines I suggested in Chapter Four. The victim is relieved of her fear. This relief is more permanent in the case of rehabilitation than traditional punishment. Limited state resources mean that violent criminals who are not executed are often released from prison. This is a fantastic let down for victims, especially in the case of violent crime. If the criminal is rehabilitated, however, the victim is given greater assurance that her assailant will not repeat his behavior. She is not merely given a reprieve during the time of the criminal's incarceration.

Let me summarize. We justify punishment according to Morris because it ensures equal distribution of benefits and burdens throughout society. The state has a duty to ensure this kind of equality in order to protect the social order. Part of protecting the social order is ensuring that criminals stop committing crime. Punishment fails in this regard. Thus we have a duty to endorse suitable alternatives like rehabilitation.

Chapter Six: Responding to Objections

While I have tried to clear away initial objections in Chapter Five by showing that there is nothing in Morris's theory which prohibits us from endorsing rehabilitation, two crucial objections to my argument remain. The first claims that rehabilitation, like punishment, fails to uphold the social order. The second position revises Morris's account, claiming that we can

jettison the appeal to social order and justify punishment solely on the grounds that we have a duty to keep the scales balanced.

Objection #1: Rehabilitation Does Not Work

The central objection to my position points out that empirical research suggests uncontroversially that rehabilitation and punishment alike both fail to keep criminals from violating the system of non-restraint. While there is some research which suggests that the quasi-rehabilitative methods employed in our criminal justice system may, in fact, work, I will proceed during this objection on the assumption that research suggests unequivocally that the rehabilitative methods we have so far employed fail.[44] If I can rebut this objection it has the effect of making my argument stronger by forcing me to overcome a greater hurdle than is, perhaps, the case.

The motivation for this objection stems from the same report cited in Chapter Five. It appears to make rehabilitation a non-starter. Such research, if truly uncontroversial, suggests that I am wrong to morally require rehabilitation since it, like punishment, cannot do the job. While Morris's system of just punishment may fail to meet the social order demands of his theory, my rehabilitation suggestion seems to fail just as clearly:

> The available research on the impact of various treatment strategies both in and out of prison seems to indicate that, after controlling for initial selection differences, there are generally no statistically significant differences between the subsequent recidivism of offenders, regardless of the form of 'treatment'. This suggests that neither rehabilitative nor criminogenic effects operate very strongly (Blumstein et. al., 66).[45]

---

[44] See Robert Martinson's "New Findings, New Views: A Note of Caution Regarding Sentencing Reform," 7 *Hofstra Law Review* 243 (1979) and Michael Vitiello's "Reconsidering Rehabilitation," 65 *Tulane Law Review* 1011 (1991).

[45] The panel here uses "rehabilitative" in the standard usage outlined above. In the previous paragraph they say "the net long-term benefits from incapacitation might also be affected by any possible rehabilitative or 'criminogenic' *effects of prison* [emphasis added]" (66). So the panel uses the word ambiguously among punitive and non-punitive measures. However, the "treatment" mentioned here by the panel does seem to include non-punitive methods. This

“Criminogenic effects” result when going to prison actually makes the criminal commit more crime upon her release, presumably because she learned more “tricks of the trade” while surrounded by other inmates. The relevant point of the quotation is that most scholars agree that state responses of various types do not stop criminals from committing crime. If this skepticism is inclusive of rehabilitation, then I cannot require rehabilitation on the grounds that it might, unlike punishment, eliminate recidivism. Because responses to wrong-doing “both in and out of prison” showed “no statistically significant differences between the subsequent recidivism of wrongdoers regardless of the form of 'treatment'” we seem entitled to conclude that rehabilitation likewise fails. Neither punitive (in prison) nor non-punitive (out of prison) responses work to stop criminal behavior. In short, it appears that no kind of response to wrongdoing reviewed by the panel seems sufficient to meet the demand laid out by Morris. No attempts to change or convince the criminal that her behavior is wrong (either punitively or non-punitively), and thus restore the system of non-interference, serve us any better than regular punishment. There seems to be no sense in wasting time investigating the moral requirement to setup an institution of rehabilitation of the type I am suggesting.

First, this quotation may not be as inclusive as this objection suggests. The authors may put the word “treatment” in quotation marks in order to distinguish it from the kind of rehabilitative response that I am suggesting. The studies that are referenced by this quotation have titles which investigate the “effectiveness of correctional treatment” and “correctional programs.” Moreover, out-of-prison responses may be equally as punitive as in-prison responses. For example, a criminal could be doing forced labor.

---

suggests that even rehabilitation (as I conceive of it) will not do the job that Morris requires. This constitutes the central objection discussed here.

I do not have time, however, to review the actual studies to see if "treatment" and "correctional" activities include rehabilitation of the type I am suggesting. Thus I am prepared to admit that they may include some quasi-rehabilitative efforts mainly because our actual justice system has incorporated some rehabilitative aspects from time to time. For example, some prisons employ pet therapy programs, where prisoners who demonstrate good behavior are allowed to enjoy the responsibility of caring for pets. Such programs may help criminals become better moral agents insofar as they increase their capacity to empathize. In short, I am not prepared to dismiss this empirical evidence from the panel (while holding on to other empirical evidence) on the ground that, despite appearances, it only covers purely deprivation-based models.

To respond to the objection I must, at this point, further distinguish the program which I think Morris's argument requires. It must be sufficiently different from any previous state responses to wrong-doing so that I can claim that the scope of the empirical studies clearly fails to include rehabilitation as I conceive it.

To get a better idea of my suggestion, consider the story of David.[46] Several weeks before David's release from a long prison term, Sarah visits our center. As David's advocate, Sarah investigates resources in preparation for David's transition from prison knowing that many ex-offenders move straight from jail to homelessness. A child of incest, David began a radical twelve step program in prison. Criminals and Gangbangers Anonymous thinks of crime as, in some cases, an addiction over which some have no control.[47] David comes to me a happy,

[46] In 2005 and 2006 I worked at a drop-in center for homeless men. This is an actual story but I changed names to preserve anonymity.

[47] This part of David's history is related only to highlight that David was trying to make himself a better person. Some may suppose though that if David were truly addicted to crime then he appears to lose his autonomy as a moral agent, freeing even Kantians to condition him as they please. But for Kantians, to lose his capacity for rational agency, he would have to altogether lose that capacity. Addiction does not strip an individual of that

healthy, funny, and gentle human being. I have no idea what crime(s) he committed. All at once, David needs to secure permanent housing, employment, identification, transportation, clothing, and more. David's halfway house is in a crime-infested part of town. The demands of living independently in our society—difficult enough for those of us with privilege—begin to wear on David. A few weeks after his release David smokes marijuana, fails his mandatory drug test and violates his parole. David goes back to prison.

David's story draws us to the kind of softer, rehabilitative response to wrongdoing that I think Morris's argument requires in order to safeguard citizens' confidence. More punitive measures in David's case are clearly not the answer. David needs the kind of moral education combined with post-release subsidies of the kind that Sarah tried to secure, all on an institutional level.[48]

David needs to be confronted about his mistaken moral judgments. He needs to investigate the rationale behind his decision to smoke marijuana and his rationale behind his

---

capacity. Furthermore the twelve step program of recovery is designed to relieve David of the obsession to commit crime. In other words, the program restores him to full autonomy, even with respect to criminal activities.

[48] It is because of stories like David's that I began to become suspicious of punishment in general. The "David Problem," however, is not the purview of traditional philosophy. At the most, the David Problem is one deep in applied ethics. We could include among the class of David Problems all of those injustices of our current penal system that do not reflect problematically on whatever *theory* of punishment a philosopher may endorse. Putting innocent people on death row, for example, is a problem of our imperfect system, but not necessarily a problem for the moral justification of the death penalty. Philosophers are more interested in whether punishment—given a certain definition and under ideal applicative circumstances—can be justified. We entertain the ideal case because, if the justification fails in the ideal, it fails in the real. We may reach a consensus, as philosophers, that the death penalty is morally acceptable but refuse, as a society, to implement the death penalty given the system's imperfections.

The purpose of David's story is thus not to highlight the problems of implementing the institution but to paint for the reader in general outline a response to wrongdoing that we have inadequately explored—a softer, rehabilitative response I believe to be morally required.

That said, I also want to note that I think David Problems are too easily brushed aside. A "feasible" defense of punishment may turn out *not* to be "morally accessible." According to Allen Buchanan, a feasible system is one that is "compatible with human psychology, human capacities generally, the laws of nature, and the natural resources available to human beings." A merely "accessible" system is one that is feasible but also has "a practical route from where we are now to at least a reasonable approximation of the state of affairs that satisfies its principles." Finally a "morally accessible" system is one that "should be achievable without unacceptable moral costs." See Allen Buchanan's *Justice, Legitimacy, and Self-Determination* (NY: Oxford University Press, 2004, 61). It is important to note when an argument that is feasible turns out not to be morally accessible because we do not want to go around endorsing systems that can only be implemented at "unacceptable moral costs." Thanks to Professor A. J. Cohen for pointing out the distinctions here.

original crime(s). This will require inquiry with someone else who challenges his beliefs. I believe part of this process will also be opening David's mind to the possibilities of a flourishing life. He may be ignorant of what life can be like out of prison and outside of the drug-infested part of town. Part of his moral education may involve, for example, bearing witness to the qualities of a virtuous life. As part of our mission in Oakland we took the men to high-end theater, museums, and national parks. We may even establish programs where David goes to live with, or at least visit, functional citizens. It will be best if he can witness functional citizenry among formerly incarcerated individuals.

The process of moral education will involve restoring David to a positive belief about his own dignity. Remember, we punished him in the first place because we respect and thus belief in his equal worth as a person. David likely does not understand this—neither that we punish him out of respect nor, more fundamentally, that he actually does have worth equal to that of the wealthiest, or most powerful, or most famous, or most righteous members of our society. The importance of restoring dignity should not be underestimated in the cases where criminals internalize the message (perhaps reinforced by prison sentences) that they are inherently evil. Nothing insures so much that a criminal will continue to *do* evil as the belief that he *is* inherently evil.

Part of restoring dignity is showing trust in David as a responsible person. For example, at our daytime facility for homeless men we employed on-the-street men as managers of our community, entrusting them—some former hard-core criminals—with keys to our facility and access to our alarm code. We believed in nurturing dignity even if that meant taking risks with our property. Of course, the risk was a calculated one. For example, while the managers had

access to the small “cash register” which held the coffee money, they did not have access to the safe.

People like David cannot engage in this kind of internal transformation without some significant resources which alleviate the stress of re-entering life as a productive citizen. David will need some space to think, uncluttered by the demands of mere survival. We could invoke Aristotelian arguments that some moderate amount of resources is necessary for a virtuous life (or at least a life that does not involve the violation of the law). “It seems clear that happiness needs the addition of external goods...for it is difficult if not impossible to do fine deeds without any resources” (Aristotle, 1099a30). Recall that we are concerned with morally educating the criminal and that for Aristotle, happiness consists in “an activity of the soul in accordance with virtue” (1098a15). Aristotle's position in ethics further supports our endorsement of moral education since he believed that “moral virtues, like crafts, are acquired by practice and habituation” (1103a11). David will need resources in order to support his change in habit.[49]

Substantive access to two crucial resources comes to mind: housing and transportation. First he needs housing outside the drug-infested part of town. Since David has neither income upon release nor any immediate way to generate that income, he will likely need subsidized housing. Perhaps the housing could be part of a work-release program where, on site, David could work and live, generating an income and gaining experience. Encouraging accountability is an important part of moral education so he should pay some portion of his rent, but it will have to be accountability in the context of the extreme stress under which these people live. Behaviors do not change overnight. For example, my supervisor in Oakland ran an organization where he offered housing at a rate determined not by market forces, but by what a man could

[49] To develop this fully, I also need to argue against other parts of Aristotle's theory, especially where he claims that people cannot be morally educated after the mid-twenties. In general I am attracted to a eudaemonistic conception of justice and would like to explore what kind of state response is justified under this view.

typically earn as a day laborer in the area. Residents who failed to pay on a given week were not immediately evicted but instead joined the Saturday morning cleanup crew.

Because secure long-term employment is particularly hard to get for the formerly incarcerated (for obvious reasons), the subsidized housing will also have to be long-term (at least one year).[50] The housing will also have to be away from the crime-ridden part of town, i.e. our current halfway house system will not do. This serves to reduce temptation (an admittedly thin reason since he could seek things out on his own) but also to restore dignity.

Secondly, David needs secure transportation. David needs secure transportation so that he can search for a long-term job and, in his case, attend twelve step meetings. Housing and transportation are the two crucial background resources that will allow David to engage in moral transformation because these resources allow for eventual functionality after the subsidized release program terminates.

This general outline is the rehabilitative program which I think has not been shown ineffective empirically and is the kind of thing that might get criminals to stop their criminal behavior. Since Morris's argument requires us to set up a system which gets criminals to stop breaking the law in order to protect the social order, we need to look into these kinds of suggestions if we are to remain plausibly Morrisian.

Objection #2: Deny that Morris Needs to Restore the Social Order

Another solution to the supposed uncontroversial skepticism about the effectiveness of punishment is to deny that Morris needs to insure that criminals stop breaking the law. We

[50] Perhaps the reasons are not so obvious to my readers. Employers run background checks and routinely refuse candidates on the basis of criminal misdemeanors, especially if the job is for an entry-level position such as telemarketing jobs. Getting these misdemeanor charges dropped was the main benefit received by participants in the Homeless Court program. For felons it may be impossible to get secure employment within a year. There was no program to drop felony charges. Even so-called "expungement" programs did not stop employers from *seeing that a felony was expunged.* I saw three main options for felons in the Oakland area: some unions (mainly the Longshoremen), day labor, and sole proprietorship. Sole proprietorship is closed to most (but not all) felons due to lack of skills leaving only the union which had a monstrous waiting list and day labor which was not secure.

would be left then with only the duty to balance the benefits and burdens. Since the restoration of equal distribution seems to be the fundamental reason that criminals deserve punishment, simply rejecting the social order premise seems like a good idea. This approach has the advantage of avoiding the potentially messy empirical data surrounding any new rehabilitative suggestions: we have no need to ensure that criminals stop committing crime because Morris does not seem to rely too heavily on the idea that we should restore citizens' confidence in the system of non-interference. Rather, we have done enough to simply counter-balance the unfair benefit by taking away the criminal's time and/or money.

I argue, however, that Morris's appeal to the social order is more important than it initially appears. The appeal to the social order gives us a powerful explanation of the motivation behind law abiding citizenry. Restoring the moral order is inexorably linked to Morris's conception of society as a mutually advantageous system of self-restraint.

We can characterize law abiding citizenry in two general ways. (1) Citizens follow the law because they understand that restraining themselves in this way is necessary to enjoy the benefits of a mutually advantageous society. (2) Citizens follow the law because they do not want to be punished. These represent the extremes of a continuum—citizens are often motivated by both and in different degrees. However, to completely jettison the social order premise we would have to embrace (2). This will not work for Morris because (1) is a perfectly plausible explanation of law abiding behavior for some citizens. Furthermore, it seems implausible to suggest that citizens are only motivated under (2). Thus we should recognize that (1) is an integral part of the account of law abiding behavior, that to embrace (2) is to leave out a substantial class of individuals for which we need to account.

Many criminals, then, are individuals who fail to recognize that breaking the law undermines the social order. They fail to see that breaking the law is bad for everybody, including themselves. Thus they need to be rehabilitated—they need to be educated about why they should follow the law.

If we are to remain plausibly Morrisian it seems we must retain the appeal to social order. Without it, we endorse an implausibly limited explanatory account of law abiding citizenry. If we reject the duty to maintain social order we thus undermine Morris's conception of society as a cooperative system of self-restraint and we fail to justify punishment on the basis of the maldistribution of benefits and burdens. There may, of course, be other ways to justify punishment retributively by explicating a criminal's desert without reference to any issues of distributive justice. Such an account may have no need to restore the social order. Given, however, that Morris's account gained wide popularity (see note 5 above), we have good reason to think that an influential retributivist argument requires us to setup rehabilitation, a surprising outcome of a notoriously retaliatory theory.

Chapter Seven: Conclusion

I have approached the strong retributive justification of punishment—the idea that not punishing criminals constitutes an immoral act—with some suspicion throughout this inquiry. We saw, though, how retributive justice gains its appeal by grounding punishment in our duty to ensure fair distribution of benefits and burdens. Retribution seems to retain its appeal despite Hampton's observation that Morris implausibly suggests that law abiding citizens find abstaining from violent crime to be a burden. So, being unable to endorse Hampton's refutation of Morrisian retribution, I have argued instead for the softer side of retribution.

We have just seen why rehabilitation is morally required under a Morrisian retributive framework. We must morally educate criminals, and provide for certain goods after their release, in order to safeguard citizens' confidence in the cooperative system of self-restraint. We turn to rehabilitation because punishment does not acceptably restore the social order and we cannot ignore the role that maintaining social order has in justifying punishment. Law abiding citizens are entitled to some assurance that criminals stop committing crime. Since moral education has not been ruled out empirically like punishment, we have a duty to endorse this alternative state response. We must confront the unexpected: retribution requires rehabilitation.

Chapter Eight: Works Cited

Aristotle. *The Nicomachean Ethics*. Trans. J.A.K. Thompson. London: Penguin Books, 2004.

Barnett, Randy E. "A New Paradigm of Criminal Justice." *Ethics* 87.4 (1977): 279-301.

Bedau, Hugo Adam. "Punishment." *The Oxford Companion to Philosophy*. Ed. Ted Honderich. Oxford University Press, 1995. 732-733.

Bedau, Hugo Adam. "Punishment." *The Stanford Encyclopedia of Philosophy* (Fall 2005 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2005/entries/punishment/.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Ed.s J. H. Burns and H. L. A. Hart, University of London: The Athlone Press, 1970.

Blumstein, Alfred, et al. "Report of the Panel." *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Ed.s Alfred Blumenstein, Jacqueline Cohen, and Daniel Nagin. Washington, D. C.: National Academy of Sciences, 1978. 1-90.

Brink, David O. "Some Forms and Limits of Consequentialism." *Ethical Theory*. Ed. David Copp. Oxford University Press, 2006. 380-423.

Darwall, Stephen. *Consequentialism*. Malden, MA: Blackwell Publishing, 2003.

---------- *Deontology*. Malden, MA: Blackwell Publishing, 2003.

Dolinko, David. "Some Thoughts about Retributivism." *Ethics* 101.3 (1991), 537-559.

Duff, R. A. "Penal Communications: Recent Work in the Philosophy of Punishment." *Crime and Justice* 20 (1996): 1-97.

Feinberg, Joel *Doing and Deserving*. Princeton, NJ: Princeton University Press, 1970.

Hart, H. L. A. *Punishment and Responsibility*. Oxford: Oxford University Press, 1968.

Hegel, G. W. F. *Elements of the Philosophy of Right.* Trans. H. B. Nisbet. Ed. Allen W. Wood. Cambridge University Press, 1991

Hampton, Jean. "Mens Rea." *The Intrinsic Worth of Persons: Contractarianism in Moral and Political Philosophy*. Ed. Daniel Farnham. Cambridge University Press, 2006. 72-107.

---------- "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13 (1984): 208-38.

Hershenov, David B. "Restitution and Revenge." *Journal of Philosophy* 96.2 (1999): 79-94.

Kadish, Sanford H. and Stephen J. Schulhofer eds. *Criminal Law and Its Processes*. Gaithersburg, MD: Aspen Law & Business, 2000.

Kant, Immanuel. *The Philosophy of Law: An Exposition of the Fundamental Principles of Jurisprudence as the Science of Right*. Trans. W. Hastie. Clifton, NJ: A. M. Kelley, 1974.

Mackie, J. L. "Morality and the Retributive Emotions." *Criminal Justice Ethics. 1.1* (Winter/Spring, 1982): 3-10.

McNaughton, David and Piers Rawling. "Deontology." *Ethical Theory*. Ed. David Copp. Oxford University Press, 2006. 424-456.

Moore, Michael S. "The Moral Worth of Retribution." *Responsibility, Character, and the Emotions*. Ed. Ferdinand Shoeman. Cambridge: Cambridge University Press, 1987.

Morris, Herbert. "Persons and Punishment." *The Monist* 52 (1968): 475-501.

Murphy, Jeffrie G. and Jean Hampton. *Forgiveness and Mercy*. Cambridge: Cambridge University Press, 1988.

Nagin, Daniel. "General Deterrence: A Review of the Empirical Evidence." *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Ed.s Alfred Blumenstein, Jacqueline Cohen, and Daniel Nagin. Washington, D. C.: National Academy of Sciences, 1978. 95-139.

Sher, George. *Desert*. Princeton, NJ: Princeton University Press, 1987.

Ten, C. L. *Crime, Guilt, and Punishment*. Oxford: Clarendon Press, 1987.