Replacing Punishment: The Ethics of Alternatives to Legal Punishment

Scott Gallagher

Submitted in fulfillment of the requirements for the degree of Master of Arts in Philosophy
David Benatar, Supervisor
Philosophy Department
University of Cape Town

April, 2013

"Ah, happiness courts the light so we deem the world is gay. But misery hides aloof so we deem that misery there is none."
-Herman Melville, from "Bartleby, the Scrivener"

Introduction

I began research for this dissertation while working as an investigator at the Public Defender Service for the District of Columbia (PDS). PDS is a government office that provides free legal representation in criminal law cases for indigent residents of DC who are charged with a crime. My job entailed meeting clients in the jail almost every day. Visitors to the DC Jail come to a room where they can speak with inmates through a plate glass wall, telephone receivers connecting the visitor side with the inmate side. It is a scene that Americans mostly know through movies, and I was shocked when I first saw the room of black women speaking across the glass to the black men in orange jump suits on the other side. It was my first confrontation with the vast racial injustices of the American criminal justice system. Since I visited the jail on legal business, I was able to sit alone in a small room with the client.

It was in one of these small rooms, with a square table separating us, that I met a nice client who was my age, and who was in prison for violating parole. He was working for a tow truck company, and in the course of following company policy, was pulled over and arrested in suspicion of stealing the car that he was towing. He was in the company's truck, had company paperwork. Because of the arrest he was revoked, meaning brought back to jail on a parole violation. He was never formally charged in court with stealing the car (the case was dropped immediately), and we at PDS thought that he would be freed from jail immediately. His boss at the company explained the situation to the parole officer, and told of how good of an employee he was. The parole officer, who could have released him from prison, decided instead to schedule a formal parole hearing three months hence, and also made our client stay in prison until the time of the hearing. The hearing was delayed twice, and our client ended spending over five months in the DC Jail, which, like many jails, is a cramped, dirty, dangerous place. Not until I was inside the system did I realize the kind of truly horrific injustice that a state could

inflict on its citizens in the name of criminal law.

As a contrast to my experience in DC, I also met a group of prisoners in Worcester Prison, about an hour away from Cape Town, South Africa, who are an inspiration. In response to the drugs and gangs that rule the South African prisons, they formed their own rehabilitation group, and named it the Group of Hope. They spend most waking hours making beads out of rolled paper to make into jewelry. Through a network of outside support, they sell the jewelry, and pay for clothes and birthday parties for orphans in the city of Worcester. Meeting these prisoners, I saw a different kind of expression on their faces: ones of warmth, curiosity and vitality. DC inmates are usually either angry or jaded, and if I were in the DC jail I would be one of the angry ones. The 16 Group of Hope members take their rehabilitation extremely seriously, and as a result, the prison staff has responded in kind. The staff has allowed the Group to give AIDS prevention conferences to the other inmates, grow vegetable gardens, and extraordinarily, let the orphans into the prison once a month so the Group members can give them small presents and play games with them. I attended one celebration in the prison courtyard in which the Group members had even rented a blow-up jumping castle.

When I read criminal justice theory, it is against the backdrop of these two lopsided experiences: the dehumanizing DC Jail, and the prisoners in Worcester making the most of their time. Criminal justice is an endlessly pockmarked and variable landscape, offering stories of outrage and despair, hope and redemption, sinister characters and normal men coping with prison hours that pass with the speed of syrup. Everyone has a theory of punishment, built over years by the accumulated details of stories that speak to what they believe. Each day yields a new crop of anecdotal evidence from news reports and popular media. Theorists aim to provide the tools to cut clear lines through the jumbled geometry of popular discourse. And drained of the headline-catching stories, theory makes for dryer reading, and this dissertation is no exception. Does theory make up in progress and consensus-building what it loses in vividness? I'm afraid that the tradeoff is not proportional. But it is with the sincere hope that theory may eventually lead us to more justifiable practices that that I offer this input

to the debate.

Below I will argue that a state should minimize legal punishment[1] and rely as much as possible on alternatives. In order to know to what extent we may minimize punishment and use alternatives, we need a clear idea of what those alternatives are, and when we can justifiably use them. The purpose of this dissertation is to analyze the morality of putative alternatives to punishment. I will explore what makes them non-punitive, define them, and analyze whether they can be justified.

The structure of the dissertation is as follows. The first chapter investigates the concept of punishment. I will defend a definition of punishment: authorized, retributive, intended harm. Then I will proceed to explain the need to justify punishment, and give an overview of how it is at least plausible to believe that no justification has yet succeeded. I will end the chapter with a brief discussion of the requirements of a criminal justice system. The second chapter is about money. I will scrutinize whether the theory of 'pure restitution' may completely replace punishment. I will argue that it cannot, and furthermore I will caution against the widespread use of mandatory monetary restitution. I will also provide a positive argument for the state's duty to provide compensation to victims of violent crime. The third chapter brings in the true heavyweights for non-punitive interventions: offender rehabilitation and offender incapacitation. After defining them, explaining why they are non-punitive, and defending justifications for them, I will conclude that they provide the most substantive opportunities for the state to shift its criminal justice burdens away from punishment. In the fourth chapter I will explore rituals: restorative justice conferences, trial and therapeutic jurisprudence, reentry ceremonies and apologies. My argument for a minimally punitive regime will come together in the last chapter. In doing so I will explain why a state must rely on punishment to a small but crucial extent, and that punishment can be minimized drastically in comparison to today's practices. I will also address concerns regarding security and deterrence.

1 Hereafter, note that when I refer to punishment I am referring to legal punishment.

Chapter One: Punishment

1.1 Defining Punishment

Since the 1950s scholars have defined legal punishment as intentional hard treatment administered by a legal authority and inflicted on a person for breaking a law. This is my reformulation of the definition that is regarded as canonical and known as the Flew-Benn-Hart definition. According to this 'standard' definition, the five conditions that an act must meet in order to be considered an act of punishment are: "(i) It must involve pain or other consequences normally considered unpleasant. (ii) It must be for an offence against legal rules. (iii) It must be of an actual or supposed offender for his offence. (iv) It must be intentionally administered by human beings other than the offender (v) It must be imposed and administered by an authority constituted by a legal system against which the offence is committed."[3]

A definition of punishment should remain neutral regarding the debate of whether punishment is justified. The Flew-Benn-Hart definition has been widely accepted and uncontroversial in this regard. An example of violating the neutrality requirement would be supplementing the Flew-Benn-Hart definition with terms that serve to justify punishment without further argument. A brazen example of this would be using the term 'deserved' within the definition, as in 'punishment is harsh treatment inflicted on an offender who deserved it by breaking the law.'

A definition of punishment should be sufficiently narrow to only delineate cases of punishment, but wide enough to capture all cases of punishment. While the Flew-Benn-Hart definition is useful as a historical standard, I think that a more precise and accurate one can be offered. Boonin defines

3 Hart, 1959-60, at 4. Michael Davis' useful discussion of this definition, 2008, at 75-76. Wood, 2010, is another example that discusses this definition, one of probably hundreds.

punishment as "authorized reprobative retributive intentional harm."[4] The definition I will be defending is a slightly amended version of Boonin's: **punishment is authorized retributive intended harm.** Granted, the difference between the Flew-Benn-Hart and Boonin definitions is minor, and the difference between Boonin's and mine is also small, but these differences matter most when talking about borderline cases, which is a primary task of this dissertation. One of Boonin's harshest[5] critics, Leo Zaibert, writes that the only differences between the Flew-Benn-Hart definition and Boonin's are that, "[f]irst, Boonin's definition incorporates an element of official disapprobation more conspicuously than the Flew-Benn-Hart definition. Second, Boonin avoids talking about punishment as ''pain,' 'mischief,' and similar terms; instead, he talks about punishment as 'harm.'"[6] Zaibert is right to point this out, but he misses another difference: Boonin also is much more precise and bold in describing what intentional harm means. In the classic definition, while Hart specifies that punishment is intentional, it is unclear what exactly that means. The amendment I make to Boonin's definition is to remove the reprobative requirement, to change a slight aspect of the retributive requirement, and to further specify what is meant by the intention requirement. I include the table below as a reference for the following sections in which I explain the constitutive parts of Boonin's definition and defend my amended version of it against objections.

	Flew-Benn-Hart	**Boonin**	**Amended Boonin**
Authorized	Yes	Yes	Yes
Retributive	Yes, Unspecified	Strong Reading	Weak Reading
Reprobative	No	Yes	No
Intention	Yes, Unspecified	'Intentional'	'Intended'
Harm	Yes	Yes	Yes

I will spend much more time discussing the intention requirement than the others, and for this reason I will save it for last. I focus on the intention requirement of punishment because it is the most

4 Boonin, 2008: 1.1.
5 Zaibert is harsh, but also unfair. In his review he mischaracterizes some of Boonin's arguments.
6 Zaibert, 2010, at 46.

important one for me to defend. Unlike the other requirements, I rely on it to frame what counts as alternatives to punishment. Boonin places similar weight on this requirement. However, he does not adequately defend his position against objections that argue the intention requirement artificially narrows the scope of punishment, so I will be defending it further.

1.1.1 The Harm Requirement

By including harm as a part of his definition, Boonin means that "acts of punishment all, in some way, make the person who is punished worse off than she would otherwise be."[7] Some writers have argued against including harm as a necessary feature of punishment because the harm could become ultimately beneficial; an example is an offender who turns his[8] life around because of his time in prison. Definitionally denying that punishment could be ultimately good for the offender would violate the neutrality requirement because it would rule out the possibility that punishment could be justified by ultimately benefitting the offender. Boonin argues that this objection fails because including harm in the definition refers only to the immediate action of punishing but remains "neutral on the further question of whether or not being subject to such a harm might produce beneficial consequences in the future, including beneficial consequences that are great enough to outweigh (and perhaps even to justify) the immediate harmful ones."[9]

1.1.2 The Retributive Requirement

Retribution, in the context of legal punishment, is reprisal for breaking the law, and thus the retributive requirement of punishment holds that for an act to be punitive, it must be inflicted on an offender in response to his particular offense. I defend a slightly amended explanation of this

7 *Id*. at 7.
8 Because the majority of offenders and crime victims are male I will use this pronoun throughout the paper.
9 Boonin, 2008: 7. He goes on to argue against other objections to including harm as part of the definition, including objections that feature cases of offenders who want to be harmed and offenders who receive community service that they enjoy.

requirement than the one Boonin defends. Here is how Boonin explains this requirement: "to be a legal punishment, an act must involve intentionally harming someone because he previously did a legally prohibited act, which means that he is responsible for having done the act and that he had no valid legal excuse for doing so."[10] While at first this may seem unassailably straightforward, note that under this definition only offenders who are in fact guilty can be punished. One objection argues that innocent people can be punished too. If, this objection goes, an innocent person is mistakenly convicted of a heinous crime and he is sentenced to death, then surely the state's act of killing him is an act of punishment. This scenario actually occurred when Texas executed Cameron Todd Willingham in 2004 for the death of his children by fire in 1991. Arson forensic investigation technique had advanced so much by the time of his execution that scientists reviewing the evidence concluded with near certainty that Willingham was innocent.[11] I agree with this objection, and my amended version of the retributive requirement holds that in order for an act to be punitive it must involve the state intending to harm someone because the state has sufficient reason to believe that he committed a legal offense.

Boonin uses analogies with repayment and reward to argue that an innocent person cannot technically be punished. The concepts of repayment and reward are predicated on certain conditions. Inherent in the concept of repayment is prior payment and debt. If you are not owed money, then you cannot be repaid; and likewise if you are not guilty you cannot be punished. Or consider rewards. If you could somehow buy an honorary Nobel Prize by donating money to the foundation, then we would not say that you 'won' a Nobel Prize. The normal conditions of winning which are, *inter alia*, rising to the top of your field on your own merit, contributing original groundbreaking progress, and being chosen by careful deliberation regarding the merits of your work, have not been filled. You have not technically been rewarded the Nobel, rather it has been granted to you based on special conditions that

10 *Id.* at 17.

11 Grann, 2009. See also Santos, 2013, for a related story in which Arizona deemed a man innocent and released him after 42 years in prison for murder by arson.

differ from the normal conditions required to win a Nobel.[14] Likewise, an innocent person cannot be punished, because the concept of being punished is predicated on the condition of actually being guilty of the offense.

However, there is one major flaw with this argument: the concepts of repayment, reward and punishment are not predicated on the *actual state of affairs* that lead to a particular repayment, reward or punitive act, but rather on *the state of affairs the relevant party believes to be true, to the best of its knowledge*. To continue the Nobel analogy, consider the case of a scientist who rises to the top of his field, contributes groundbreaking progress, and is chosen by the Nobel committee to receive a prize, but unbeknownst to the committee and the rest of the scientific community the scientist has been cheating by plagiarizing and using the work of others without attribution. Under Boonin's strict conception, he has not be rewarded because he has not deserved the prize based on his own merit, i.e. he has not met the specific conditions required to qualify as a recipient of the reward. But this is not how the concept of reward works in practice. The relevant party deciding the merit of the recipient gives the reward, and even if their decision was based on information that turns out to be false, that does not change the fact that they did reward the recipient.

Consider a case of a counterfeiter lending someone money. The recipient spends the counterfeit money and later dutifully pays money back to the lender, but he does not realize that the original money was counterfeit. In this case, the recipient is a relevant person to be deciding whether he repaid the counterfeiter. If he told his friend that he repaid the money, it would be strange if we deemed that he was not properly using the concept of repayment. Rather, the state of affairs gave the recipient sufficient reason to believe that he did in fact repay the money, and so we should agree that he was properly using the concept of repayment. If at a later time the recipient learns of the counterfeit, then he might think that he was swindled into unfairly repaying. But this does not change the original condition of repayment in which the relevant party assessed that a repayment was made.

14 This example comes from Sandel, 2012, at 96.

Consider punishment: there are many cases in which we will never know the truth about the crime. The state does its very best to find out the truth, but it cannot be assured of absolute certainty. When Texas convicted Willingham, it had used the best means available of ascertaining the truth. After forensic technology advanced over the proceeding decade, those means were found to be faulty, and we now know with relative certainty that Willingham was innocent. He was punished, but we have the rare ability to look back and assess that the punishment was unfair. In most cases when the state punishes an innocent person, neither the state nor the public will ever know about the offender's innocence, and we still refer to these cases as punishment. The concept of punishment squares best with our intuitions when we use it from the point of view of the relative party assessing guilt. The relevant party is the state, and when it determines that an offender is guilty and inflicts harm to him for the offense, we say that he is punished without any reference to the certainty of some external 'Truth.'

One might object and insist that repayment requires actual debt, reward requires actual deservingness, and punishment requires actual guilt. I have discussed these three concepts because they are the ones that Boonin uses in his illustration, but there are many similar concepts that rely on certain preconditions. However there is a range of these concepts, spanning those whose preconditions are usually self-evident, to those whose preconditions are less certain and thus determined by a relevant party to the best of its ability. I argue that legal punishment is in the latter category. In the former, there are easy cases such as: if I recover from an injury, I must have been injured, or if I reacquaint myself with a person, I must first have been acquainted. In these such cases there is usually no reason to doubt the original condition, e.g. of injury or acquaintance, and we might simply call a person who claims to be reacquainting with someone they have never met a liar who is misusing the concept of re-acquaintance. However, even in these cases, there may be situations that would make us question whether the original state of affairs need actually be true in order to merit the use of the concepts of recovery and re-acquaintance. For instance if a hypochondriac takes a day to recover from an illness he perceived but did not actually have, or if an imposter lookalike goes around getting to know the old

acquaintances of the person he is impersonating. The uncertainty of preconditions intensifies as we move to concepts such as punishment. The very reason that there is a committee to determine the Nobel prize, or a trial to determine guilt, is because conditions of deservingness and guilt are difficult to ascertain. There is less certainty, and we therefore use these concepts based on the eye of the relevant deciding party. For these reasons I conclude that the term 'retributive' should be defined in terms of a state's knowledge, so that a state's actions against an offender are retributive if it has sufficient reason to believe that the offender is guilty.

Finally, one might claim that the term 'sufficient' is too ambiguous, and claim that it is better to rely on the actual state of affairs to define the concept of punishment. However, I see no problem with leaving the degree of certainty to be decided by the relevant party. In the U.S., guilt is often determined by jury trials, in which twelve citizens need to make this determination of what is sufficient reason. Various systems have various methods of determination, and while some may be more successful than others, I see no reason that this ambiguity should challenge my defense of the retributive requirement. Furthermore, there are usually varying determinations of guilt throughout the process of a trial. If someone is arrested and held before trial based on strong suspicion of guilt, then this action is retributive, even if later at trial he is determined to be innocent. At the time of arrest, based on the best available knowledge, the state determines that there is sufficient reason to believe that the suspect is guilty in order to detain him. This fulfills the retributive requirement, and thus we may call the action retributive. Later when all the evidence is weighed in a more systematic way, he is determined innocent, but this determination does not retroactively change the status of his pretrial detention. It was a retributive action all along.

Boonin labels the two readings of the retributive requirement the 'strong' and 'weak' readings: the strong one claiming that only guilty people may be punished and the weak one claiming that innocent people may be punished. I have argued for the weak reading, and have focused on defending this reading because it will influence some of my later discussions, specifically the places I use pre-trial

detention as an example.

1.1.3 The Reprobative Requirement

Boonin and others argue that punishment involves not only retribution, but also reprobation. Harm is inflicted because a law is broken, but also because the state condemns the offender's actions. According to Boonin, "[t]he reprobative requirement maintains that part of what makes an act a punishment is that it expresses official disapproval of the offender's behavior."[15] R.A. Duff in particular has focused on the expressive aspect of punishment as a way to justify punishment. But including reprobation within the definition of punishment does not violate the neutrality requirement because it remains neutral on the further question of whether the condemnatory aspect of punishment can justify it.

The most immediately puzzling question about this requirement is whether it is necessary. I will argue that adding reprobation to the definition is redundant because the retributive requirement of punishment (regardless of strong or weak reading) makes punishment necessarily reprobative by condemning the actions of offenders who break the law. For the reprobative requirement to be justified, there would need to be cases of authorized, retributive, intended harm that do not condemn the breaking of the law. Boonin attempts, but fails to show that there are such cases, and he uses the following example. Imagine a gang whose initiation rites include making its initiates "break the law – say, by stealing a car – in order to be eligible for initiation."[16] Imagine further that only after the initiates break the law do they undergo the next phase of the initiation, which includes the gang members intentionally harming the initiates, for example branding them with a hot iron. In this case, an authority (the gang leadership), is intentionally harming someone as a direct consequence for breaking the law, yet the action carries no condemnation with it. Rather, praise and approval of breaking the law

15 Boonin, 2008, at 23.
16 *Id.* at 22-23.

are implicit in the rite.

This example, while clever and original,[17] does not succeed in allaying worries of the redundancy of including reprobation within the definition. One could argue that the authority in the case of gang leadership is importantly different from the state's authority in legal punishment. By using laws set by an entirely different authority from the one meting out the harmful treatment, the example may provide room for reprobation and retribution to disjoin. What is needed is an example of legally authorized, retributive, intended harm that is not reprobative. I do not believe there is such a case. In making a criminal law, the state is drawing a line, and in doing so, it inherently condemns any instances when offenders cross that line. The act of retributively harming the offender for crossing that line is an expression of the state's condemnatory attitude toward the offender. There is no way that the two could become separated. As von Hirsch puts it, criminal law is "morally loaded…The message conveyed through prohibition is that the behavior is reprehensible, and that the person should consider its wrongfulness (and not just the unpleasant consequences) as a reason to desist. The criminal law expresses blame in its very design."[18] Or as Duff writes, criminal law is "an institution that defines, and by implication condemns, a range of public wrongs."[19] Punishment does in fact include reprobation,[20] but there is no use including it as a separate requirement because it is redundant given the retributive requirement.

1.1.4 The Authorized Requirement

While there are non-legal kinds of punishment, they are not the focus of this paper, so when I use the term 'punishment,' I am referring to legal punishment. Legal punishment must be "carried out

17 Zimmerman, 2011, uses the same example, at 17, to illustrate the same point but does not attribute it to Boonin.
18 Von Hirsch, 2011, at 271.
19 Duff, 2011b, at 72.
20 This is an uncontroversial claim: see e.g. Feinberg, 1970, von Hirsh, 1993, and Duff, 2001, among others.

by an authorized agent of the state acting in his or her official capacity."[23] The requirement that an action needs to be authorized to qualify as legal punishment is the least controversial of the constitutive parts of punishment, so I will not address it further.

1.1.5 The Intention Requirement

Punishment is not only harmful, but intended to be harmful. When we punish we "invoke our authority in order to inflict pain."[24] There are many cases when two harmful treatments are identical, but one is clearly punishment and the other is not. Boonin gives the following illustration: the state could harm someone by making him pay $5,000. However, if the state is making him pay it for taxes, then he is not being punished. But if the state is making him pay because it is harming him due to his guilt, then it is clearly punishment, because "[w]hen the state punishes someone... it inflicts various harmful treatments on him *in order* to harm him."[25] It is important to distinguish the difference between merely foreseeing that one will cause someone harm and intentionally causing him harm.[26] It may seem objectionable to distinguish whether an act counts as punishment based on the intentions of the punisher, but this provides an accurate account of what we consider punishment. Consider another example Boonin gives: "[i]f you see a uniformed official forcing a laborer to lift a heavy rock, you cannot know whether what you see is a prisoner being punished or a slave being exploited" unless you know the intention of the officials.[27]

The glaring problem with examples like the laborer and the $5,000 is that they don't isolate intention as the only difference between the harmful treatments. It is possible that the reason we view the slave and the prisoner differently has to do with other properties of punishment (besides intended harm) that don't exist in the slave labor case. Perhaps what accounts for the difference we see between

21 Boonin, 2008: 24.
21 Tunick, 1992, at 1.
25 Boonin, 2008, at 13, original emphasis.
26 Hanna, 2009, also makes this argument.
27 Boonin, 2008, at 14.

the two is that the slave hasn't broken a law. In order to isolate whether intention is essential to punishment, we should choose examples that differ only with regard to the intention of the harm.

In the following discussion I cover seven such examples. The first three I list briefly as they have been considered examples of problem cases that have not been able to be explained by the standard definition of punishment. One of the benefits of the definition I defend is that it neatly handles these problem cases that have plagued the standard definition. Three years after Boonin's book was published, Dolinko wrote in *The Oxford Handbook of Philosophy of Criminal Law* that the standard definition needs modification because cases such as deportation, which is clearly not punishment, fall under its scope. Dolinko goes on to note that there are also problem cases such as impeachment of a president, which is non-punitive because, quoting Fletcher, "removal of the president is constitutionally authorized to protect the public, not to inflict 'unpleasant consequences' on the offending president."[28] Disbarment of attorneys is another practice that "satisfies all of Hart's criteria and clearly carries a condemnatory message, but whose status as 'punishment' is debatable."[29] These three problem cases (deportation, impeachment, and disbarment), however, provide examples of acts that fulfill all the requirements of punishment except one: intended harm. Dolinko concludes that "[f]or present purposes, however, there is no need to arrive at a precise set of necessary and sufficient conditions for punishment."[30] I contend that there is a pressing need for such a definition in the literature to avoid further misunderstandings, and it was very nearly filled with Boonin's definition in 2008. Below I give four further examples which demonstrate that when the intention to harm is not present, but the other elements of punishment are, then the act is non-punitive, and therefore intended harm is a necessary requirement of punishment.

While Boonin's discussion of definition makes the mistake of not isolating intention in order to defend its necessity, he also provides a compelling example in a footnote that does isolate intention as

28 Dolinko, 2011, at 405, quoting Fletcher, 1978, at 410.
29 *Id.* See also Dolinko's note 15 which explains this example further.
30 *Id.*

the only relevant feature differentiating a non-punitive action from punishment.

In the U.S. there are laws "which authorize states to post photographs of convicted sex offenders and other information about them on the internet after they have been released from prison."[32] The constitutionality of these laws was challenged on the basis of punishing offenders above and beyond the punishment they already received while serving their sentences. In 2003 the U.S. Supreme Court ruled that while these so-called Megan's law provisions harm offenders, they do not punish offenders. By making this distinction, the Court

> "relied precisely on the distinction between foreseeable and intentional harm. As Justice Kennedy wrote for the majority in the 6-3 decision, 'The publicity may cause adverse consequences for the convicted defendants, running from mild personal embarrassment to social ostracism,' but the intention of the laws is 'to inform the public for its own safety, not to humiliate the offender.'" [33]

Megan's law provisions fit all the requirements of punishment except for the intention to harm. When the state posts a photo of a sex offender online, it is authorized because it is administered by the state; it is retributive because it is a direct consequence of an offender being convicted for breaking a law; and it is harmful. If we agree with the Supreme Court's interpretation that posting the photo is not punishment, then it must be because the intention to harm the offender is missing, which aligns with the Court's stated reason for its interpretation.

An objection to the argument from the Megan's law provisions is that they are not truly retributive because they are not directed at the offender. But this is a misunderstanding, and perhaps a conflation of retribution with intention. Retribution is backward looking: harmful treatment of an offender (like publishing the photo) that is a direct result of the offender's conviction qualifies as retributive. The fact that the harm isn't directed at the offender is precisely the reason that the act does not qualify as punitive.

Here is another example of a non-punitive act that is retributive, authorized harm, but not

32 Boonin, 2008, at 15, n. 18.
33 *Id*., citing Holland 2003: 7A.

intended to be harmful. Imagine a virus that is very dangerous for a large portion of the population but for the rest of the population it has no medical effects. Furthermore, this virus is extremely contagious, and can be easily caught if someone comes into contact with a mosquito in a particular region of forest. The government has made it illegal for citizens to enter that region. A man enters the region, contracts the virus but is not affected by it, and when he exits the region he is detained by a border patrol agent. He is tested for the disease and quarantined, but his quarantine is in a comfortable environment, say a hospital, without any effort to harm him further than keeping him from spreading the virus. Once he is determined to be safe to reenter the public, he is released from the hospital. We would not say that this is punishment, and furthermore this is an example of authorized, retributive harm without intending to harm the offender. He is guilty an offense, which makes the state intervention retributive because it is in direct response to the offense. The government has the law in place to protect its citizens, and it detains any and all offenders to test them for the virus. Imagine further that if the offender has the virus, he is quarantined, and if he doesn't then he is set free with a fine to help pay for the upkeep and improvement of the border patrol. Neither the detention for testing, nor the quarantine, nor the fine is intended to harm the offender. All three harmful treatments are explicitly intended to prevent a deadly outbreak of the virus, and furthermore not intended to harm the offender.

Another example is the practice of pre-trial detention,[34] which resembles the virus example in many ways: in both cases a government detains an offender first to determine whether he is potentially dangerous, and if he is found to be so, it detains him not to punish him but out of concern for the safety of the community. Interpretation of the US Constitution has held that pretrial detention is not punishment. *United States v. Salerno*[35] is a US Supreme Court case that is perhaps most useful[36] for understanding the history of the discussion of whether pretrial detention is punishment. The case came

34 Boonin brings up pre-trial detention, 2008: at 15, n. 18.

35 *United States v. Salerno* 481 U.S. 739 (1987).

36 Boonin and Tunick have used the case of *Bell v. Wolfish*, but I think it is less illustrative. *Bell* deals with a very specific question of whether a particular detainment practice is inhumane, while *US v. Salerno* includes discussion of the history of the debate surrounding pretrial detention. *Salerno* deals broadly with the constitutionality of pretrial detention in general, citing *Bell* as well as other precedents.

to the Court because Anthony Salerno was arrested in relation with mafia activity, and detained before his federal trial because he was found to be potentially dangerous to the community. Salerno's side argued that his detention constitutes punishment before trial and is therefore unconstitutional. The Court concluded that pretrial detention is not punishment, and the Court's opinion is very helpful to understand the intention requirement, so I will quote it at length. The opinion explains that pretrial detention is not punitive but regulatory, and that the intent of the legislators is the key to determining whether a law is punitive. It also includes a discussion of the limits and restraints placed on regulatory detention.

> [...]The Court of Appeals assumed that pretrial detention under the Bail Reform Act is regulatory, not penal, and we agree that it is.
>
> As an initial matter, the mere fact that a person is detained does not inexorably lead to the conclusion that the government has imposed punishment. *Bell v. Wolfish, supra,* at 537. To determine whether a restriction on liberty constitutes impermissible punishment or permissible regulation, we first look to legislative intent. *Schall v. Martin,* 467 U.S. at 269. Unless Congress expressly intended to impose punitive restrictions, the punitive/regulatory distinction turns on
>
> "whether an alternative purpose to which [the restriction] may rationally be connected is assignable for it, and whether it appears excessive in relation to the alternative purpose assigned [to it]."
>
> *Ibid.,* quoting *Kennedy v. Mendoza-Martinez,* 372 U.S. 144, 168-169 (1963).
>
> We conclude that the detention imposed by the Act falls on the regulatory side of the dichotomy... Congress instead perceived pretrial detention as a potential solution to a pressing societal problem. *Id.* at 4-7. There is no doubt that preventing danger to the community is a legitimate regulatory goal. *Schall v. Martin, supra.*[37]

I include this quotation to demonstrate three points. First, there is a long history of courts determining that pretrial detention is non-punitive precisely because it lacks the intent to harm the offender. Secondly, in order to determine that the intention is non-punitive, it is important to ask what the intention is, and where it originates. In the case of pretrial detention, the intention to maintain community safety comes from the legislators who decide what circumstances would potentially justify

37 *United States v. Salerno* 481 U.S. 739 (1987).

pretrial detention. The legislators' regulatory power[39] in this case is restricted by the detainee's due process rights (to a speedy, fair trial, to legal representation, etc.). Third and finally, it is up to the courts to determine whether the legislative intent is in fact non-punitive; and whether the restriction of liberty in pretrial detention is proportional given the further legislative goal (e.g. of community safety for pretrial detention); and whether the "procedural safeguards"[40] have been adequately upheld. The Court's opinion describes the test for whether an action qualifies as harmful but nonpunitive regulation as a dialogue between legislation and courts: legislators declare intent, and the courts interpret and judge its appropriateness and proportionality.

A quick and final example is confiscation of illegally obtained profits. When the state seizes assets obtained through criminal acts, its action is clearly not intended to harm the offender, but rather to recover stolen property. The action is retributive, harmful to the offender, but not punitive. This shows again that the lack of intention to harm the offender results in a non-punitive action. The Dutch government classifies such confiscation is a "measure," explicitly defined as non-punitive because of the lack of the intention to harm.[41]

One objection to the intention requirement comes from Jan De Keijser, who criticizes the Dutch government's categorization of "measures" as non-punitive, calling this a "definitional stop."[42] De Keijser's accusation of a definitional stop means that he is claiming that the intention requirement I have defended is guilty of violating the neutrality requirement for defining punishment, because it "prevents us from examining the rational and moral status of deliberately inflicting pain."[43] But why can we not examine the moral status, or justification, of harmful but non-punitive interventions? De Keijser gives no reason. I take him to mean that the distinction unfairly justifies (or puts a stop to moral scrutiny without further discussion), harmful state interventions that are not intended to harm. But there

39 Tunick, 1992, at 126 cites another Supreme Court case "where regulation is distinguished from punishment": *Schall v Martin*, 467 U.S. 253 (1984), at 269-270.
40 *United States v. Salerno 481 U.S. 739 (1987)*. at 742.
41 De Keijser, 2011, at 200.
42 *Id*.
43 *Id*.

is nothing about calling these interventions 'non-punitive' that works to justify them, or violates the neutrality requirement. Taxes are one such intervention, and there is no reason that we are prevented from examining their "rational and moral status" simply because they are non-punitive.

Due to current penal practices, which do not sharply distinguish sanctions based on intended harm, punishment is widely seen as any harmful sanction for breaking the law. The lay public will have differing intuitions on whether practices such as pretrial detention and Megan's law provisions count as punishment. And this is in part because the public usage of the term 'punishment' is broad, with indistinct boundaries. But the mere fact that current public intuition might blanche at the idea of a nonpunitive speeding fine should not exclude the possibility that under the most precise definition of legal punishment, a nonpunitive speeding fine could exist. Especially considering the legal precedents I have highlighted, a nonpunitive speeding fine is plausible with sufficient legislative and judicial dialogue regarding the regulatory intent. I have provided a defense of a precise definition of punishment, one that sharpens the standard philosophical definition with the specification of the intentional requirement. Unless this definition can be shown to be faulty, there is no good reason to dismiss it based on current popular usage of the term.

It is likely that, despite the foregoing discussion and arguments, some will still take this definition to be mistaken, and insist that intended harm is not a necessary feature of punishment. Yet even if that is the case, the reader could take this dissertation to be a conditional argument: if one understands punishment as I have defined it, as authorized, retributive, intended harm, then I will explore the alternatives to punishment and show that we should minimize punishment by relying on these alternatives.

It is still useful to keep in mind that most scholars writing about punishment construe it as including intentional harm.[49] To take two examples out of dozens, Kolber recently noted, "Criminal

49 Boonin, 2008, at 13 n. 14, lists 15 examples of writers endorsing it, which he calls an "extremely small sample of those who explicitly endorse it." I agree with this assessment, and in addition to his list of 15, would add Feinberg, 2011, at

law theorists overwhelmingly agree that for some conduct to constitute punishment, it must be imposed intentionally;"[50] and Sayre-McCord writes "[p]unishment has at its core the *intentional infliction of pain or harm*."[51] It may seem strange that so many scholars use a definition that might now strike the reader as controversial, but the full implications of the intended harm requirement have generally gone unnoticed.[52] Also, it may have struck my reader that I have been careful to use the term 'intended' instead of 'intentional,' which is the much more popular phrasing among the literature. This is because there is an ambiguity that arises in many scholars' definitions of punishment by the use of 'intentional.' One may construe the phrase 'intentional hard treatment' to mean that an act was taken deliberately that resulted in hard treatment of an offender. That is to say, the action itself was taken deliberately, but may or may not have as the intended result the harm of the offender. In this way, taxation is intentional harm that the state imposes on its citizens. However this is importantly different than the concept I have been defending. I use the adjective 'intended' to modify the harm requirement because it leaves no such ambiguity: the aim of the action must be to inflict harm.[53]

The reason I have taken such care to defend the intention requirement of punishment is because in order to determine what counts as an alternative to punishment the primary difference between an alternative and punishment in many cases will be the state's intention behind the act. In some cases, the intentions of the state interventions that I take to be alternatives to punishment are clear enough to be uncontroversial, but in many others the intentions might be debatable. Objectors to Boonin's definition of punishment have accused him of directing the intention based on what is convenient to label an alternative to punishment.[54] This objection doesn't actually make a case against Boonin's definition, but

---

67, Duff, 2001, at xiv, McDermott, 2001, at 403, Zimmerman, 2011, at 20, Hanna, 2009, at 238, Corlett, 2003, at 285, Sayre-McCord, 2001, at 504, De Keijser, 2011, at 200, as well as dozens of others that I won't list here.

50 Kolber, 2012, at 1.

51 Sayre-McCord, 2001, at 504, his emphasis.

52 Even when scholars use Boonin's definition, they are apt to misuse the intention requirement. See Ward, 2009, at 290, where he uses the intention requirement to apply to the intended therapeutic outcomes of clinicians, rather than harm. Also see von Hirsch, 2011, at 263 for an explicit rejection of using the intention requirement to determine an alternative.

53 Husak, 2011, at 222 makes the intentional/intended mistake, which is very common.

54 See especially Zaibert, 2010.

signals a potential pitfall when discussing alternatives. Leo Zaibert points out, and I agree with him, that someone's reported intentions may not be his true intentions, and morally categorizing an act based on such a fickle mental state presents serious problems. But I disagree with him that categorizing a state intervention following crime as either a punishment or an alternative faces this problem. As the long quotation I used above from *US v Salerno* shows, the intentions behind such interventions are not hidden mental states within one person's mind. They are explicit and publicly debated, arising from the nexus between the legislators and the judiciary: the legislators declaring the intention, the judiciary interpreting whether the implementation or means of achieving the intended goal are appropriate or excessive. It is true that there are criminal justice decisions that come down to the intent of one person, say a police officer or a judge. But even in those cases there is (or should be) an extensive system of regulations to prevent abuse, such as sentencing guidelines, and the language of the regulations is where we may determine the intent of the legislators. Zaibert is correct to point out that we should be vigilant when discussing the intention behind alternatives to punishment. While many alternatives are straightforwardly non-punitive, like apologies or monetary victim restitution, others are less obvious. When discussing an alternative that I think might be controversial, in that some may view it as a punishment, I will address it and attempt to use legislative and judicial precedent where possible to illustrate the state's intention

1.2 Justifications of Punishment and Their Objections

In this section I review the major justifications of punishment, and some of their leading objections. The major justifications, consequentialist, retributivist, expressivist, and others such as educational, have produced a vast literature, hence I don't intend to give a detailed account of any of the justifications and objections, but rather a brief, general overview.

First it is necessary to understand why there is a need to justify punishment. But the need is not

simply a theoretical dilemma, detached and awaiting for an Arthurian sage to finally pull the justification from the anvil; rather, meeting the need is long past overdue. We would be guilty of tolerating a very serious wrong if our current punitive practices turned out to be unjustified. The need to justify our practices presses upon us with the weight of the intense suffering of the millions of people being punished around the world daily. This suffering is inflicted purposefully, by the state, with the moral smugness of unflinching certainty. But scholars have tried to meet this need to justify punishment for many decades, and there is good reason to believe that they have failed.

This need for justification is what philosophers refer to as the 'problem of punishment.' Generally, books or articles that attempt to justify punishment start by acknowledging this problem, so there are many different wordings.[56] Generally it can be stated that by punishing someone the state treats him in a way that would otherwise be wrong, so the resulting problem is the moral need to justify this practice. Ted Honderich writes, "The problem of punishment arises mainly but not only for the reason that the practice involves...a deliberate and avoidable infliction of suffering."[57] Hart writes of the "mounting perplexities" with punishment and of the need for a "morally tolerable account of this institution."[58] Duff writes that "when we reflect on the punishments inflicted (in our name) on so many of our fellow citizens and on the effects of those punishments on those who suffer them, we cannot but raise the question of legitimacy—of what can justify any practice of criminal punishment."[59] Boonin adds precision to these general calls for justification. He writes that since it is harder to "justify intentionally harming someone than it is to merely foreseeably harm [him]," then the problem becomes more acute than it first seems: "we must explain not only why the line between offenders and nonoffenders is morally relevant at all but, in particular, how it can be important enough to justify not

56 See Dolinko, 2011, at 403 for a description of how the question of the justification of punishment goes back to "at least the time of Plato." McDermott, 2001, asks succinctly "How could evil acts have the effect of making morally permissible what would otherwise be further evil acts?" at 404.

57 Honderich, 1969, at 11. The question of whether it is in fact avoidable suffering is the subject of my later chapters.

58 Hart, 1968, at 1.

59 Duff, 2001, at xii.

merely harming those on one side of the line but intentionally harming them."[60]

While philosophers are in unanimous agreement that there is a need to justify the practice of punishment, few note how pressing the need is. Because the need is so urgent, due to the currently enormous and ever growing[62] amount of suffering that is intentionally inflicted via punishment, it is dire enough to compel us to desist from our current practices of punishment as much as possible in favor for alternatives, only punishing as much as required by necessity. Adam Gopnik recently wrote in *The New Yorker*, "The scale and the brutality of our prisons are the moral scandal of American life." He goes further by saying that mass incarceration is the fundamental fact of the USA today, "as slavery was the fundamental fact of 1850."[63] I will not evaluate here whether Gopnik, and other writers like him[64] who decry current penal practices, are justified in making such sweeping claims. I quote him merely to suggest that punishment could be analogous to other, once-widespread practices that inflicted tremendous suffering, but have failed to be justified. Slavery is one such practice. Gross notes other practices like this, including torture, which used to be respectable but is now "widely regarded as morally deplorable."[65] To further illustrate the link between torture and punishment, I point to Stephen Kershnar's article arguing for the permissibility of punitive torture. His argument in part relies on the acceptance of forfeiture-based retributivism as a justification of punishment.[66] To this list I can confidently add the practice of legally discriminating against people based on their race or sexual orientation. Since I will ultimately argue that we should minimize punishment, slavery is not the most apt comparison. There is no acceptable 'minimal slavery.' This difference points to the fact that the moral questions surrounding punishment are more nuanced than those surrounding slavery, torture and state-sanctioned discrimination. A better analogy is killing non-combatants in war. History is littered

60 Boonin, 2008, at 28. Boonin also considers two objections to the claim that punishment is in need of moral justification.
62 In 1980 there were 220 prisoners per 100,000 people in the USA; by 2010 there were 731 per 100,000. See Gopnik, 2012.
63 *Id.*
64 See especially Davis, 2003, Alexander, 2010, and Stuntz, 2011.
65 Gross, 2012, at 162.
66 Kershnar, 2010.

with wars in which opponents practiced this without reservation. It is only relatively recently that we have arrived at consensus that this practice must be minimized as much as possible. Over the course of this dissertation I argue that we should treat punishment the same way: using it only when necessary, and immediately reform our practices to minimize it as much as possible.

To add one final reason that the need to justify our practices is more urgent than widely believed, Kolber has made a compelling case that unintended but reasonably foreseeable harms should also need justification.[69] This means that foreseeable harms such as loss of jobs, restriction of sexual activities, infliction of heightened risk of assault, etc. also need justification. The massive suffering that goes on in crowded prisons that is not intended by the state but is foreseeable also needs to be justified, and this adds to the burden of justification.

The goal of this section is to convince the reader that it is at least plausible that no justification for punishment has succeeded. It should be noted that many scholars who are not abolitionists agree that all justifications up to now have failed. Stephen Garvey, who is critical of abolitionists concedes that "[a]ll efforts to come up with a justification have so far failed, and are perhaps are bound to fail."[70] Dolinko writes, "The search for an acceptable justification of punishment continues" and that "the quest for the justification of punishment may never come to an end."[71]

There are two reasons I take the time to explain the justifications of punishment and their objections, given my goal of assessing alternatives. First, I want the reader to have at least a taste of how forceful the objections are without simply taking other writers' word for it. Alternatives to punishment become more important once one admits the possibility that we may not be able to justify punishment. And secondly, if one is familiar with the rigorous objections that have been fired at punishment, it provides a ready arsenal of possible arguments against the alternatives I consider later. Throughout this brief summary I rely heavily on Boonin's book *The Problem of Punishment*. The reason for this is that it is by far the most comprehensive and detailed book-length treatment of the

69 Kolber, 2012.
70 Garvey, 2011, in Deigh and Dolinko eds., at 498.
71 Dolinko, 2011, in Deigh and Dolinko eds., at 427 and 428. Interestingly, Dolinko only briefly considers that this search may be futile, in a footnote mentioning Boonin's book, but never fully considers the implications of the weight of the objections to the putative justifications. See note Dolinko's note 225.

justifications for punishment and their objections.[72] Because this is merely a glancing overview, there are many forceful objections that I do not mention, and furthermore, I do not have the space to fully explain the power of the objections that I do mention. Readers should not take anything in this section to be a full argument against a particular justification; but rather it should be seen as an abbreviated sketch of some of the reasons why scholars are pessimistic about justifications of punishment.

1.2.1 Consequentialist Justifications and Objections

All consequentialist justifications argue that punishment is justified because of its good consequences. The most common good consequence that these justifications appeal to is a decrease of law-breaking, which means that the community will be safer, and people will be happier and better off because of punishment. Consequentialist justifications cite various ways in which punishment is able to bring about this increased well-being, and the most prominent are: (1) general deterrence, meaning that potential offenders choose not to commit a crime because of the potential punishment; (2) incapacitation, meaning that offenders being punished are restricted from committing further crimes because of the constraints of their punishment; and (3) specific deterrence, or desistance, meaning that offenders, once finished serving the terms of their punishment, choose to not commit further crimes because of their prior punishment.

There are two main consequentialist justifications. The act-consequentialist version maintains that happiness and well-being are increased by punishing offenders, and moreover, that each act of punishment should be evaluated based on its usefulness of increasing well-being. The rule-consequentialist justification also maintains that happiness and well-being are the goals of punishment, but rather than evaluating each act of punishment based on its usefulness in achieving that goal, it argues that the overall good consequences are what justify the institution of punishment. When

72 Honderich, 1969, provides a good starting point, but old enough to miss many subsequent important developments in the debate. Compare also with Tunick, 1992, Golash, 2005, Zimmerman, 2011, and Gross, 2012, each useful but not nearly as comprehensive.

evaluating individual cases, according to the rule-consequentialist, if we determine that an offender is in fact guilty, then we can justify punishing him according to the general rules which justify the institution as a whole. This allows the rule-consequentialist to speak of constraints that can be placed on specific acts of punishment, while the justification is based on the system as a whole. Hart provided a famous version of this, arguing that "it is perfectly consistent to assert *both* that the General Justifying Aim of the practice of punishment is its beneficial consequences *and* that the pursuit of this General Aim should be qualified or restricted out of deference to principles of Distribution which require that punishment should be only of an offender for an offence."[73] This broad kind of consequentialism, which includes rule-based, side-constrained and Rawlsian versions, among others, arose in part because of serious objections against the act-consequentialist justification, but it also faces serious objections.

Here I will explain the two main problems with the act-consequentialist justification, and I will then show how the rule-consequentialist justification tries but fails to solve these problems. A major objection to the act-consequentialist justification is the 'punishing the innocent objection.' Because the main goal of punishment is to increase overall well-being on a case by case basis, "then in at least some circumstances, deliberately punishing an innocent person is also morally justified because it produces more overall utility than would any available alternative."[76] There are many examples in which punishing an innocent person might be the best way to prevent more crime. For instance, to prevent a riot the state could punish an innocent person whom an angry crowd wants punished. Or the state could make an example of someone innocent in order to deter future criminals. There are many similar examples in the literature.[77] Because this justification allows for punishing innocent people, this justification is unacceptable.

Another objection besides punishing the innocent is the 'not punishing the guilty objection.' There are many cases in which consequentialist reasoning will lead to the unacceptable result of not punishing a guilty person. For instance imagine a famous movie star that is so widely beloved, that

73 Hart, 1968, at 9. His stated reason for why he took up his mixed theory is that there are too many problems plaguing the act-consequentialist justification. See at 21.

76 Boonin, 2008, at 41.

77 *Id*., see note 4 at 41 for a partial list.

after committing a significant crime, the state calculates that it would be overall more painful for everyone to see him suffer than to punish him. Clearly such reasoning is unacceptable because it treats the movie star unfairly based on his popularity, which should not be taken into account for how to respond to his crime. There are at least three further convincing objections to the act-utilitarian version which I will not discuss here.[78]

The rule-consequentialist family of justifications faces some of the same problems as the act-consequentialist version, including the 'punishing the innocent objection.' In overcoming this objection, the rule-consequentialist version "must ultimately rest on two claims: that legislators reasoning on utilitarian grounds would endorse a system of rules on which innocents are never deliberately punished and that this would suffice to justify adhering to such rules even in cases where more good would be done by breaking them."[79] Rawls attempts to argue just this.[80] However, this is a chimerical task, because such a system of rules cannot be fully adhered to when using consequentialist reasoning. There would simply be too many cases that would arise in which punishing innocents would lead to more foreseeable good, and if the state were using consequentialist reasoning it could not justify declining to punish the innocents in these cases. Consider a murderously angry mob that wishes to see a certain man jailed. The state suspects, or even knows that he is innocent. But it also knows that the mob will likely kill ten or more people in riots if the state does not placate it by sentencing the man to jail. Even if it chose a hefty sentence of five years for the innocent man, it would be avoiding the much worse outcome. In this case, using consequentialist grounds, it would not make sense to stick to rules forbidding punishing the innocent. Rawls' formulation artificially cuts out plausible candidates for punishing innocents based on consequentialist grounds, such as this case and many others like it.

This demonstrates a problem plaguing rule-consequentialism known as the "rule worship problem." Suppose, despite the foregoing arguments, there is a way for a utilitarian legislature to craft a rule against punishing innocents,[81] there remains the problem: "why, on utilitarian grounds, should

78 Boonin explains these in detail. See 2008, 2.1.
79 *Id*. at 66.
80 *Id*. at 65-7, quoting Rawls, 1955, at 8-12.
81 And furthermore assume that this rule accounts for our strong intuition against punishing innocents rather than simply avoiding practical problems. See the preceding note.

judges always follow the rules that the legislature sets down for them?"[82] This is a general problem that rule-utilitarians face. If there were a clear case when the best overall utility can be gained by breaking the rule, why, on utilitarian grounds, would someone be constrained by the rule? Rawls gives an answer specific to the 'punishing the innocent' objection, which is that a judge can rightly claim that his authority, *qua* judge, does not allow him to break rules in service to greater utility. But this is simply evading the question by fixing on practice guidelines for judges. "The practice conception of rules simply replaces the rule worship problem with the practice worship problem: why should a person go along with a practice in those cases where abandoning it would do more good?"[83] Thus, Rawls' answer to the "rule-worship" objection fails. These same reasons that prevent the rule-consequentialist from solving the problem of punishing the innocent also obstruct it from solving the problem of not punishing the guilty.

Besides the problems of punishing the innocent and not punishing the guilty, there are other objections that plague the rule- consequentialist justification. I will not mention most of them, but to give a brief example of how forceful they are, consider this one further objection: the rule-consequentialist justification may result in wildly disproportionate punishments. If a justification entailed a practice of executing motorists for speeding, then we would have to abandon the justification. And there is no reason to believe that rule-consequentialist reasoning may not lead to such an outcome, especially given the enormous death toll of accidents on highways, and how many fewer people die when motorists adhere to lower speeds. "If a single act of executing someone for speeding would produce enough deterrence to compensate for the harm done to the speeder himself, for example, then establishing a system of rules in which anyone caught speeding is executed would do so even more effectively."[84]

1.2.2 Retributivist Justifications and Objections

There are many versions of retributivist justifications for punishment, including: desert-based,

82 Boonin, 2008, at 70.
83 *Id.* at 73.
84 *Id.* at 76. See 2.3 and 2.4 for Boonin's objections to other, less powerful, versions of consequentialism.

forfeiture-based (and the similar debt-based), fairness-based, trust-based, and revenge-based.[85] These versions focus on the offender, and claim that punishing him is justified because he deserves it, or because he forfeited his right to not be punished by offending, or because it is only fair to punish him, or because he violated trust, or simply because society wants revenge. Together these justifications share the quality of being backward looking. That is to say, rather than concerns for future benefits which consequentialist justifications aim for, retributivist justifications focus their reasons on features of the offender based on their past offense. I will explain three of these justifications (desert-based, forfeiture-based and fairness-based), and summarize some of the objections against them.

Perhaps the most intuitive version is desert-based retributivism, which maintains that the state is justified in punishing an offender because the offender deserves the punishment. There are two versions of claims that desert-based retributivists make. The more expansive one claims that we are obliged to punish an offender because he deserves it; and the narrower version claims that we are merely justified in punishing him.[86] I will only be addressing the versions supporting the narrower claim because this is the easier claim to justify. Many desert-based justifications argue from particular cases in which we feel strong intuitions that someone should be punished, and these are usually cases of horrific offenses. Because of our strong intuition that these offenses are punishable, the argument goes, then the moral permissibility of punishment in general "best accounts for" our strong moral beliefs, and can then be applied to all law breaking.[87] Another way to understand this argument from particular cases of desert avoids the messiness and question-begging of extrapolating from particular intuitions to justification of punishment in general.[88] This second construal is what Boonin claims is the strongest reading of Moore, Kershnar and other desert-based retributivists. "To say that a person deserves to be punished, on this account, is not simply to say that she should be punished. Rather, it is

85 These names come from Boonin, 2008, at chapter 3.
86 Kant advocated for the broader reading in *Philosophy of Law*, tr. 1887, quoted in Honderich, 1969, at ch. 2.2.
87 Boonin, 2008 at 87 quoting Moore, 1987, at 98.
88 See Boonin, 2008, at 87-92 for an explanation of these problems.

to say that the world will be intrinsically a better place by one morally relevant measure if she is punished than it will be if she is not."[89] In other words, the strongest, most charitable reading of desert-based retributivism takes our intuition that particular offenses deserve punishment, and reasons that the world is inherently a better place when offenders are punished. Many philosophers argue that this reasoning is invalid,[90] but assuming for the sake of argument that it is valid, the question remains whether this can support a justification of punishment. Here I will summarize two objections which argue that this is not possible: the 'not punishing the guilty objection,' and the 'punishing the innocent objection.'

A justification of punishment would need to draw the line between innocent and guilty such that the state would retain the right to punish anyone who is guilty. The 'not punishing the guilty objection' shows that the desert-based retributivist cannot draw such a line. The driver who drives without a license and speeds to get a pregnant woman to the hospital does not deserve to suffer based on moral blameworthiness.[91] Therefore, under the desert-based retributivist justification, the state would not maintain the right to punish the driver because he does not deserve it. There are many cases in which an offender is morally blameless, and hence not deserving of retribution, but still legally guilty. The state should still be able to preserve the right to punish the legally guilty but morally blameless, and the desert-based retributivist cannot justify this. A response to this objection could say that breaking the law is still overall immoral, but that would entail that the driver would be morally prohibited from driving the woman to the hospital and would have to search for his license before leaving, which is clearly not the case.[92]

However, even if this weren't the case, and if, for the sake of argument, every instance of breaking the law could be counted as overall immoral, there are still many cases of law breaking in which the offender does not deserve to suffer. A person could act out of ignorance or for a good motive. A poignant example of acting from ignorance comes from a Pulitzer prize winning piece of journalism about parents who forget that their infant is in the back seat and leave the child to die in the

89 Id. at 92.
90 See *id.* at 93 note 10 for a list.
91 There is a limitless number of examples like this that show that someone has broken a law but remained morally blameless.
92 Boonin, 2008, at 93-96.

hot car.[93] This happens with surprising frequency, and it is clear from the cases that any parent is susceptible to this tragedy. It usually happens when the parent is busy, stressed, not following normal routine, and the child is in a car seat in the back seat facing the rear of the car, as is mandated by new child safety laws in the USA. The parent returns to find the child dead, and begins in that moment what is surely a lifetime of intense suffering because of his action. Some judges have conceded that the parent does not deserve to suffer any more than he clearly is and convicts the parent of criminal negligence but is lenient with his sentence. The desert-based retributivist, acknowledging that the parent neglected the child due to external factors and with no intention to harm, would agree with the lenient judge that the parent does not deserve to suffer any more than he already is. It seems uncontroversial, given our current practices, that the state should preserve the right to punish such parents. Therefore in order for this retributivist justification to succeed it would need to provide reason for the state's right to punish the parent, but it cannot.

A less dramatic example explains why desert-based retributivism cannot maintain the right to punish (*nor can it actually punish*) the legally guilty but morally blameless. A man is in a rush to park his car, and he parks it in an illegal spot that he does not realize is off limits. The state punishes him regardless of whether he morally deserves it, and the desert-based retributivist, whose only criterion for punishment is moral blameworthiness, cannot justify the state's action. There are many strict liability offenses that states punish even if the offenders are blamelessly ignorant of committing a crime.[94]

Another objection to desert-based retributivism is the 'punishing the innocent objection.' There are many cases of people who do not break the law but who clearly deserve to suffer. The desert-based retributivist must hold that the state can punish such a person, which is clearly a case of punishing the (legally) innocent (despite being morally culpable). Boonin gives the example of a man who has familiarized himself with spousal abuse laws "and does everything he can to make his wife miserable without crossing that [legal] line. If he is legally allowed to scream at her, he screams at her. If he is allowed to cheat on her, he cheats on her. In any way he is allowed to embarrass, belittle, degrade, and insult her, he does, and with relish."[95] However, given that desert-based retributivism originates in

93 Weingarten, 2009.
94 Zimmerman, 2011, at 155.
95 Boonin, 2008, at 99. There are many similar examples of this kind of morally reprehensible but legally permissible

cases of morally reprehensible behavior and reasons that the culprits deserve to suffer, then clearly this man, according to this reasoning, deserves to suffer. And "since such desert is the basis for the right to punish on the desert-based retributivist account," then it follows that under such an account the state should punish the man even though he "violated no just and reasonable laws."[96]

One response to both of the above objections is to say that desert may come from both legal culpability and moral culpability. This would explain why the state may punish the morally blameless but legally guilty, as well as not punish the morally blameworthy but legally innocent. However, this is an ad-hoc response that begs the question: the problem the justification is supposed to solve is why the state is allowed to punish, and only allowed to punish, the legally guilty. The justification cannot accomplish this without adding two caveats to address the above objections: the morally blameworthy deserve punishment, but only if they are legally guilty *and* the legally guilty but morally blameless also deserve punishment. Adding these qualifications vitiates the explanatory power of the justification.

The justification, to review, tracks our impulse that certain offenders deserve to suffer because of moral culpability. Boonin cites Kershnar who gives the example of a woman who helps out the Nazis by pointing out overlooked Jews; our "intuition is that she deserves a rather severe punishment" even though she acts to help the legal authorities.[97] This goes to show that we track moral culpability rather than legal culpability for the kinds of intuitions founding this account, which provides additional reason that the above response fails. One might claim that a desert-based retributivist could simply draw the line at legal culpability and claim that the state should only be able to inflict suffering when an offender commits a moral violation that also breaches a just law. But this response cannot explain why the state is justified in giving the parking ticket to the time-crunched and morally blameless driver who parked in the off-limit spot. And once again, this is an ad-hoc adjustment: if the justification results in undesirable consequences such as punishing the innocent and not punishing the guilty, it is the justification that should be changed or dismissed rather than tweaking the justification's logical outcome according to preference.

In contrast to the desert-based version, the forfeiture-based retributivist justification holds that

---

behavior.

96 *Id*. at 99-100.

97 *Id.* at 101, citing Kershnar, 2001, at 33.

an offender forfeits a right not to be punished by committing an offense. Goldman writes that "by violating the rights of others in their criminal activities, [offenders] have lost or forfeited their legitimate demands that others honor all their formerly held rights."[99] To defend the claim about forfeiting rights, a forfeiture-based retributivist may argue that there is an essential link between moral rights and moral duties. This link provides the basis to contend that "negating one's moral duties would involve negating one's moral rights."[100]

Boonin details seven forceful objections to the forfeiture-based retributivist justification of punishment, and Lippke notes that many others have also subjected it to "withering criticism."[101] I only explain two from Boonin: the 'unforfeited rights objection' and the 'disproportionate punishment objection.' These two are no more or less compelling than the five others Boonin presents, but are good representatives.

The 'unforfeited rights objection' claims that it is extremely unlikely that infringing certain duties would entail forfeiting the corresponding rights. For example, the forfeiture claim "entails that a rapist forfeits the right not to be raped and a torturer forfeits the right not to be tortured." Other examples are someone who prevented a religious gathering forfeiting his religious freedom, or a judge who takes a bribe forfeiting his right to a fair trial. These examples are deeply counterintuitive, and if we reject them, "then we must be unwilling to accept the forfeiture claim" and must reject the justification.[102] One response is, as Goldman puts it, "if we ask which rights are forfeited in violating rights of others, it is plausible to answer just those rights that one violates (or an equivalent set)."[103] The most common equivalent set that people have in mind is of course time in prison. But if it is truly equivalent, then we should be willing to choose freely between the prison time and the corresponding forfeited right. But since we are unwilling to accept that the state is justified in sentencing the rapist to being raped, then we wouldn't accept the equivalent loss of freedom in prison. In other words, "if a

99 *Id*. at 105, quoting Goldman, 1979 at 31 and 33. See also McDermott, 2001, at 405-406, who gives a debt-based account of the same concept: "when wrongdoers violate their victims' rights, they incur a debt for the value of the treatment they withhold from their victims – the treatment owed to the victims as right-holders – and that as a result of incurring this debt, wrongdoers forfeit their rights to certain goods that may be denied to them as punishment."

100 *Id.* at 107.

101 Lippke, 2002, at 127.

102 Boonin, 2008, at 110.

103 *Id.* at 111, quoting Goldman, 1979, at 33.

given right is really equivalent to another, then if it is unacceptable to deprive someone of one right, it must be equally unacceptable to deprive him of the other."[104] This is in addition to the insurmountably daunting task of trying to give reasons why the severity of a particular punishment equates to the lost rights ensuing from a particular offense.

The 'disproportionate punishment objection' arises from the following question: Does the offender forfeit his right temporarily or permanently? It would seem unreasonable to think that a petty thief caught stealing $20 would forfeit any right for the rest of his life simply for this one transgression. If it is a temporary forfeiture, how do we know what the length of time is, and how do we ensure that it is not disproportionate? "[T]here seems to be no nonarbitrary means of determining how long that would be without simply abandoning the foundations of the forfeiture-based position."[105] Most crimes result in a temporary period of harm for victims, which means that most crimes could potentially be punished too much or too little because the forfeiture claim leaves no way to safeguard against this.

Fairness-based retributivist justifications, like forfeiture-based versions, frame the justification of punishment within larger concepts. The forfeiture-based version is couched in rights, while the fairness-based version is framed "in terms of distributive justice."[107] Specifically, this version contends that "offenders must be viewed as enjoying an unfair distribution of benefits as a result of having committed an offense, so that imposing on them a punitive harm will restore the overall distribution of benefits and burdens to its previous and presumptively fair level."[108] The unfairness comes from the offender free riding on nonoffenders' cooperative obedience of the law that results in "peace and stability." The free riding specifically comes from the offender's evasion of "the burden of self-restraint assumed by those who obey the law."[109] Punishment therefore takes the unfair advantage away from the offender.

There is one objection here that I will focus on, and five others that I will not discuss.[110] The 'not punishing the guilty objection' stems from the founding claims of the fairness-based account: first that

104 *Id.* at 111.
105 *Id.* at 113.
107 Boonin, 2008, at 120. Boonin also claims that based on its popularity "the fairness-based approach is arguably the preeminent form of retributivism in the current literature."
108 *Id.*
109 *Id.* at 121.
110 However my omission is due solely to concern for concision. These are also forceful objections; see 3.3.

offenders can be understood "as free-riders who take unfair advantage of their law-abiding fellow citizens."[111] And secondly that punishment is justified in leveling the fairness. There is a clear problem with the first claim: not only does it not account for the wrongness in the harm to the victim, but it also fails to account for most people's lack of desire to commit crimes. It is perverse to think that what is wrong with a murder is that it takes advantage of people who refrain from murder rather than the harm to the victim. Most people have no desire to murder, and hence "there is no cost to them refraining." And if there is really no cost, then there is no unfairness due to free-riding, which means that "there are many cases of lawbreaking for which the fairness-based retributivist cannot justify punishment."

1.2.3: Expressivist Justifications and Objections

While consequentialist and retributivist justifications have been the most common in the literature, there are other important justifications. The most prominent is expressivism, also known as the reprobative or communicative justification. Other justifications include the consent justification, the moral-education justification, the self-defense justification and hybrids of any of two or more of any of the aforementioned justifications. For the purposes of brevity I, summarize expressivism and some of its objections, while leaving other justifications out of this overview.

In a sentence, expressivism makes two claims, first that the state should condemn crime, and secondly that only "penal hard treatment" is adequate to fully express the censure, making it a necessary component of the state's condemnation.[115] Expressivism had an early and notable treatment in an influential paper by Joel Feinberg called, "The Expressive Function of Punishment." But it retains its force in the contemporary debate largely due to R.A. Duff, and his book *Punishment, Communication and Community*.[116]

Feinberg writes that the expressive function of punishment serves many important purposes, namely that: it allows the state to disavow the offenders' actions; it serves to "speak in the name of the people" the state's symbolic nonacquiescence; it vindicates the law by "emphatically reaffirming" it

111 Boonin, 2008, at 122.
115 Duff, 2001, 29.
116 Feinberg, 2011, and Duff, 2001. See also Boonin, 2008, at 171, n. 10 for a thorough list of other expressivists. In addition to his list, see also Tasioulas, 2006, for a nuanced treatment.

when it is broken; it validates the rights of victims that were transgressed; and finally it absolves innocent suspects who were mistakenly accused of the crime. Feinberg asks, "Could not the state do this job without punishment? Perhaps, *but when it speaks by punishing, its message is loud and sure of getting across*."[117] Thus Feinberg claims that punishment is necessary because other means of official reprobation are insufficiently reprobative.[118] He does not give reason why other means are deficient beyond simply making the claim, so I turn now to R.A. Duff.

Duff is also emphatic in explaining the state's need to condemn crimes. The state should "mean what it says" by outlawing certain acts, and thus "it is committed to censuring those who nonetheless engage in such conduct. To remain silent in the face of their crimes would be to undermine—by implication go back on—its declaration that such conduct is wrong."[119] In answering why punishment is necessary over other methods of censure, Duff writes that one route that some theorists, most notably Andrew von Hirsch, have taken is to argue that "penal hard treatment adds a deterrent incentive to the law's moral appeal."[120] Note that this type of move means that the expressivist is sliding into a consequentialist justification. Duff prefers to claim that punishment is a kind of secular penance, and is thus a kind of "moral reparation," or "reformative enterprise...[for] his future conduct."[121] Duff includes a grab bag mix of many different punitive purposes within his notion of penance. He succinctly restates them in a later article, claiming that the intentional harm of punishment is enforced penance and "should serve both to assist the process of repentance and reform, by focusing his attention on his crime and its implications, and as a way of making the apologetic reparation that he owes."[122]

Here I will summarize three of the main objections against expressivism: the 'nonpunitive censure objection,' the 'not punishing the guilty objection' and the 'punishing the innocent objection.'[123]

---

117 Feinberg, 2011, at 117, original emphasis.

118 Tasioulas, 2006, rests on a similar assertion to overcome the question of why intentional harm is needed above other forms of censorship: "And here the answer must be that only punishment adequately conveys the blame the wrong-doer deserves. This captures a widespread and deeply ingrained judgment, viz. that the seriousness of the wrong committed warrants a blaming response that operates *through* the infliction of hard treatment, since only such a response adequately reflects the gravity of the wrong that has been committed," at 296, his emphasis.

119 Duff, 2001, 28.

120 *Id*. at 29. See von Hirsch, 1999.

121 Duff, 2011a, at 179.

122 Duff, 2008, §6. He also acknowledges that this faces serious objections, "in particular that it cannot show penal hard treatment to be a necessary aspect of a communicative enterprise." See also his 2011 article and his explicit discussion of how there could be alternatives that achieve many of the goals of expressivism.

123 Boonin, 2008, at 177-181.

First, the 'nonpunitive censure objection' claims that the expressivist has failed to show why other means of condemnation cannot be used to adequately censure the offender. Very little work has been done to explain this, and most scholars resort to a method similar to Feinberg's: by simply claiming that punishment will be "loud and clear" enough to express whatever it is that the state is expressing. This is inadequate. Even Duff and von Hirsch fail to give compelling reasons why alternatives may not, instead of punishment, perform the functions of expression and inducing penance (Duff) or expression and deterrence (von Hirsch).[124] There are many means of nonpunitive censure, and the foremost of them is the criminal trial. As I will discuss later, the state's intentions behind having a criminal trial include ascertaining guilt, showing symbolic non-acquiescence with the crime, and expressing sentiments such as remorse, condolence for victims and condemnation of the offender. But harming the offender is not a plausible intention. If we grant that there are many nonpunitive ways that a state may condemn an offender, including trial, then the expressivist needs to show how these methods are so inadequate as to necessitate the recourse to punishment. This is a daunting task. We can easily imagine scenarios in which there are two offenders guilty of the same offense, one receiving punishment and the other receiving a nonpunitive condemnation, and the state's condemnation of the punished offender falls much shorter of its condemnation of the other offender. Consider if in the punitive case the offender was ushered through a trial in which the judge made a quick decision, the offender was whisked away to prison for a couple of months, and then released, all without press or fanfare. In the other case, the state held a well-publicized trial, and the judge issued a lengthy opinion with his sentence which was published prominently in newspapers. Perhaps the mayor issued an official denunciation as well. However, the sentence that was issued was non-punitive. It is clear that the state's condemnation is louder and clearer in the non-punitive case, despite being non-punitive.

The 'not punishing the guilty objection' holds that "there are many cases in which a person violates a just and reasonable law but does not seem to merit moral criticism." A successful justification of punishment would be able to "justify the right to punish in such cases, but the reprobative solution clearly cannot do this."[125] Consider a person who violates a strict liability law, meaning that he can be punished even if he has an excuse. The expressivist justification cannot account

124 See Duff, 1999 and 2001, and von Hirsch 1999. I will address the deterrence question in Chapter 5.
125 Boonin, 2008, at 180.

for punishing these kinds of cases.

The 'punishing the innocent objection' claims that the state may condemn certain actions without making them illegal. The expressivist justification would then require the state to punish those legally innocent people who commit such actions. Consider Boonin's example of the state that "declared a certain day to be Holocaust Memorial Day. In doing so, the state would clearly be acknowledging that those who deny the fact of the Holocaust merit disapproval." But because most states would not make it illegal to express such denial, we have a clear case in which someone is guilty of an action that the state morally condemns, but does not legally proscribe.

1.3: Requirements of Criminal Justice

Even if punishment cannot be justified, it may be the case that there are no viable alternatives. In this case punishment may be a 'necessary evil.' Boonin calls this the 'argument from necessity:' justifying punishment by demonstrating the insufficiency of alternatives. In order to determine whether punishment may in fact be justified on the basis of necessity, and if so to what extent it may be used, it is important to be able distinguish various non-punitive alternatives. Throughout this dissertation I will use the term 'post-arrest regime' to refer to the collective institutions of the criminal justice system that operate after an offender's arrest. A post-arrest regime is in charge of pre-trial detention, trial, sentencing, parole, etc.; in other words, it is the part of the state that is in charge of implementing punishment, and any alternatives to punishment.

Before assessing alternatives to punishment on their own merits, I will pause to give some basic requirements that any post-arrest regime must meet. I will not consider these again until the final chapter, in which I will argue for a minimally punitive post-arrest regime, but these requirements will be useful to keep in mind while considering alternatives. What are the minimum thresholds that any post-arrest regime would need to fulfill in order to be considered tolerable and practicable? Here I briefly propose and discuss three minimum requirements. I merely intend to give a workable standard

by which to measure the various alternatives in the final chapter.

In order to determine requirements for a post-arrest regime, the first step would seem to be to ask: what are the goals and purposes of a state's response to crime? That is, setting aside questions about the purpose of a police force, what should a state hope to accomplish once a crime has been committed? Without wading into theoretical territory that is beyond the scope of this essay, there seem to be three obvious candidates:

1. Maintaining public safety and order by pressuring adherence to the law
2. Communicating condemnation of offenses
3. Redressing harm done to victims and society.

Pressuring citizens to adhere to the law and preventing recidivism both aim for the same consequence: to maintain safety and order. Of course, every state allows for risk to its citizens, and the same is true when it comes to crime. No state could be expected to maintain completely safe conditions for its citizens, and any attempt would likely involve severe abuses of individual liberty. However, every state is also expected to maintain enough order and safety to enable people to live productive and fulfilling lives without undue fear of crime. Kershnar calls this the "the minimum conditions for just mutual relations to exist,"[130] which echoes Kant, and is similar to the often-quoted Hobbesian concept of 'right to sufficient security.'[131] A state is also in the business of communicating moral condemnation. As Feinberg and Duff argue persuasively, the state should denounce crime for various reasons. The state should condemn crime, among many other reasons, to show its moral non-acquiescence, communicating a stance of opposition to the crime, and giving appropriate moral weight responding to the crime. Finally, the state should also make amends for the harms of crimes. A state's response to crime should acknowledge the harm to the victim, and make efforts to redress the harm.

This chapter has hopefully set the stage for exploring alternatives to punishment. I've defended

130 Kershnar, 2012, at 74.
131 See e.g. Ramsay, 2011, at 146 and Matravers, 2000, at 240.

a definition of punishment, which allows us to be precise about what precisely counts as punishment, and what stands outside of its ambit as alternative candidates for dealing with offenders. I've given the reasons why theorists should scrutinize alternatives and their justifications. If we can enjoy the safety of crime prevention without inflicting so much intended suffering, then we should make haste to rely more on such alternatives. The idea of avoiding punishment except what is necessary is known as 'parsimony' in the literature.[132] Before we can tell just how parsimonious we can be with punishment, we need a better idea of what we would rely on in its stead. There has been relatively little theoretical work done on these alternatives, which was the inspiration for this dissertation.

132 See e.g. Lippke, 2007, at 13-14.

Chapter Two: Restitution and Compensation

Boonin concludes *The Problem of Punishment* by arguing for an alternative to punishment called pure restitution, and contends that punishment cannot be justified out of necessity because pure restitution is a viable alternative. Pure restitution is, as of yet, the only theory that by itself aims to fully replace punishment and has received decent treatment among theorists.[136] In this chapter I explain the justifications for and objections to pure restitution, and conclude that pure restitution, even in its most defensible form, cannot be a standalone alternative to punishment. However, I also discuss the possibility that restitution can be one part of a minimally punitive crime response regime. I argue that restitution is best treated as a voluntary option for offenders rather than a compulsory method of redressing harm to victims.

In the second part of this chapter I discuss a concept closely related to restitution: compensation. When scholars use the term 'restitution' they typically refer to money paid directly from the offender to his victim, and that is how I use the term. 'Compensation' is a broader term, used loosely to refer either to money paid to a crime victim or to money paid to the state from the offender. When I use the term 'compensation' I refer specifically to money paid by the state or other agency to the victim, but excluding money from the offender (that is, excluding restitution). I will argue that victim compensation for violent crimes is an essential, obligatory component of any post-arrest regime, and that it is currently neglected.

2.1 Pure Restitution

One of the first serious proposals to replace punishment altogether was put forth by Randy

136 The restorative justice movement has produced a vast literature, but it cannot be considered a single theory, let alone one that attempts to fully replace punishment. And while some writers, e.g. Menninger, 1966, argue that punishment could be replaced with rehabilitation and treatment, this theory has not received serious attention.

Barnett in 1977 when he proposed to replace punishment with 'pure restitution' (which I will define below). He argued that all punishment could be replaced with monetary restitution paid from the offender directly to the victim. Barnett's article inspired a vigorous debate, and drew many criticisms. Scholars responding to the objections (but not necessarily embracing pure restitution or the wholesale replacement of punishment) paved the way for amended, more defensible versions of pure restitution; they include Mane Hajdin (1987), Stephen Wilkinson (1996), Joseph Ellin (2000), Geoffrey Sayre-McCord (2001, 2002), David Boonin (2008), and Nathan Hanna (2008, 2009).[137] Sayre-McCord and Boonin offer the most analysis, and Boonin's is by far the most comprehensive. I want to acknowledge the risk of seeming parasitic of a single author; but I believe that a fair treatment of the subject must address Boonin's detailed and thorough arguments, especially because I am rejecting his conclusion. In this section I describe how defendants of pure restitution have responded to objections, and why these objections have not been sufficiently overcome.

To avoid confusion, when Barnett uses the term 'pure' to describe his theory of restitution, he means that under his theory the state can only force an offender to make monetary payments to his victims. The state cannot punish an offender, nor coerce other actions beyond the scope of extracting payment. However, scholars defending amended versions of pure restitution argue for a state that cannot punish, but may coerce actions beyond simply transferring money to the victim, such as community service. In this sense 'pure' refers to the status as being purely non-punitive rather than specifying a limitation to monetary payments.

To further clarify, the concept of pure restitution has come to encompass a family of resembling versions rather than a single account, and some further differences arise from the questions of the roles of the police and civil law in a regime of pure restitution. Some restitutionists take the extreme and untenable view of advocating the "abolition of criminal law" altogether.[138] Criminal law would

137 Barnett also published responses defending his position (see 1980 and 1998).
138 Ellin, 1994, at 1.

essentially collapse into tort law, and the police and courts (to various degrees depending on the theory) would be privatized. For instance, Ellin writes that much of the burden of "investigation, apprehension, and prosecution" would be paid for by the victim out of pocket, or by insurance companies, and this cost would then be added to the offender's restitution.[139] It is not clear to what extent Ellin and others propose to privatize these institutions, but it makes no difference for this discussion. Because this view is unpopular and liable to obvious objections, it would be unfair to attack pure restitution based on this version. For the purposes of this dissertation I will assume that the most defensible version of pure restitution will leave the institutions of the courts and police unchanged, and maintain that the theory can still account for the distinction between criminal and tort law.

2.1.1 Introducing Pure Restitution and Common Objections

Here I present Barnett's original version, followed by what I take to be the essential elements of any theory of pure restitution. I will note the strongest general criticisms of pure restitution. I then will follow this section with a more detailed exploration of the responses to objections and a discussion of whether even the most defensible version can overcome them. In his original version, Barnett wrote that under a system of pure restitution,

> "No longer would the deterrence, reformation, disablement, or rehabilitation of the criminal be the guiding principle of the judicial system. The attainment of these goals would be incidental to, and as a result of, reparations paid to [the] victim. No longer would the criminal deliberately be made to suffer for his mistake. Making good that mistake is all that would be required."[140]

Policing and trials would remain unchanged, but when a criminal defendant is found guilty, he would be sentenced to a certain amount of restitution. If he could pay it immediately then he would be absolved of further obligations. Much of Barnett's article is spent elaborating on possible "refinements" to this "basic system." He notes that insurance companies could sell crime insurance, which would pay

139 *Id*. at 2.
140 Barnett, 1977, under 'Outline of a New Paradigm.'

out if you were a victim, leaving the companies responsible for collecting the restitution from the offender. If the offender could not pay, then "he would be confined to an employment project."[141] And there could be "*direct arbitration*" between offender and victim, resulting in a lower restitution if the offender pleads guilty.[142]

While the details of Barnett's version differ from subsequent theories, the essence or Barnett's argument remains in any theory of pure restitution. Boonin summarizes his version as follows: "if an offender is responsible for having wrongfully harmed a victim, then (a) the state should compel the offender to restore the victim to the level of well-being that the victim rightfully enjoyed prior to the offense and (b) the state should not punish the offender."[143] By "restore" Boonin means that "the theory maintains that the offender must restore the victim to a condition that is as close in value to her original condition as possible."[144] Boonin's definition seems to square well with other formulations, despite differences in how scholars elaborate the details of their versions. All theories of pure restitution rest on two fundamental tenets: first, the state should compel offenders to repair the harm done to their victims; and second, the state should not punish offenders.

Here I will introduce seven common objections to any theory of pure restitution. The first is the objection that restitution is too individualistic: in most cases of crime the total resulting harm cannot be restored by giving restitution only to the primary victim(s). The crime's effects spill over to society, its harms are not neatly contained within the domain of its immediate victims. Thus a defendant of pure restitution would have to be able to explain how an offender could restore the harm done to society. The second objection is the concern that there are crimes without a specific victim that can be compensated, raising the question of how offenders could possibly pay restitution in such cases. Crimes that fall under this category include cruelty to ownerless animals, harm to public institutions, attempted

141 *Id*.
142 *Id*., Barnett's emphasis.
143 Boonin, 2008, at 223.
144 *Id*. at 224.

but unsuccessful crimes, and reckless driving that doesn't result in injury to other drivers.[145] Under a regime of pure restitution, it is not clear that the offenders in these cases of crime would owe any specific victim restitution despite the serious transgression. Thirdly, some crimes like rape and murder seem too serious to be properly resolved by restitution. The fourth objection is that restitution would provide insufficient deterrence and reprobation. An offender might be tempted to view restitution as no more than a potential fee for criminal activity; and of course such a message from the state would be unacceptably permissive of crime. Relatedly, the fifth objection is that rich offenders would have such an unfair advantage that they would be able to buy crimes. If a rich serial rapist chose to pay the hefty costs of restitution to his victims, then we would think that something is rotten with the system. The sixth concern is that poor offenders would be at a severe disadvantage. It isn't clear how they would be forced to pay the cost of their restitution without resorting to unfair measures that systematically repress poor offenders compared to ones who can afford the costs of restitution. The last objection holds that forcing offenders to pay restitution is in fact punishment.

These objections provide a structure for pure restitutionists of what a defensible version would have to overcome, so I will progress through the various scholars who address these issues (and I will use the arbitrary order of the objections that I listed above). I will also consider objections to these responses and adapted versions of pure restitution. I should note that the seven main objections that I outlined above are not exhaustive. For the sake of brevity I have grouped together some objections that could be considered distinct.

2.1.2 The Harm to Society Objection

This objection holds that because the harms of crimes are not limited to the immediate victims, pure restitution cannot account for harms done to these indirect victims. Ellin notes that this objection can come in two versions, one claiming more than the other: “On the strong version, society as such is

145 Miller, 1978, at 359.

victimized by crimes…. On the weak version, all actual victims are individual people," by which he means that individual people, including the direct victims and indirect victims comprise the group that requires restitution.[146] I take the weak version to be troublesome enough for pure restitution to merit ignoring the strong version in this discussion. Ellin further notes that there are two kinds of indirect victims, which he calls "direct-indirect victims" (e.g. witnesses of the crime, family of the direct victim) and "indirect-indirect victims" (e.g. neighbors who hear about the crime and feel less secure).[147] Boonin groups both of these kinds together under the label of 'secondary victims.' Hajdin argues (and Ellin adopts his position) that if every apprehended offender is made to pay full restitution to his victim(s), plus some extra money to fund restitution for victims of unsolved crimes, then indirect-indirect victims would have no real reason to worry about the effects of crime. As Hajdin puts it, "Why would it matter to you whether the crime rates in the society are high or low, if the system of justice were set up in such a way that your becoming a victim of a crime could not make you worse-off?"[148]

Dagger gives two forceful reasons to reject this response. The first is that there are some direct victims who will never be fully restored (rape and murder victims). This explains why indirect-indirect victims can be justifiably worried about an increase of such crimes in their neighborhood, despite even the most robust system of victim restitution.[149] There does not seem to be an adequate answer to this response under Hajdin's theory. The second reason is that even in other cases, such as car theft, a potential victim would still choose to avoid the stress and hassle of replacing a car over what is considered to be fair restitution. Hajdin might respond that in such cases the monetary compensation need only be raised to the point in which it truly is a fair trade-off, which would negate the worry that a potential victim would have a reason to prefer one outcome to another. This response is problematic for two reasons. First, Hajdin admits that for his system to be feasible the amount of restitution owed in

146 Ellin, 1994, at 7-8.
147 *Id*. at 8.
148 Hajdin, 1987, at 82.
149 Dagger, 1991, at 6.

each case cannot depend on individual assessments of every victim, but rather a general calculation based on the type of crime committed. Due to the application of his own theory, it is not plausible to raise restitution in every case to adequately account for each victim's ideal point in which it truly is a fair trade-off between suffering the harms of victimhood and reaping the rewards of restitution. Secondly, as Dagger points out, so long as people are averse to being victims of crime, "people will feel the indirect effects of crime, through attitudinal and avoidance costs, in ways for which they cannot be compensated."[150] In other words, even despite its implausibility, if it were the case that every victim were given the exactly correct restitution such that they were made no worse off than before, we cannot help our intuitive desire to avoid being crime victims. Our worries stemming from such desires do not include rational calculations of future restitution. For these reasons we should reject Hajdin's solution to the Harm to Society Objection.

I'll now move from Hajdin to Boonin, who gives two responses to the Harm to Society Objection. The first is similar to Hajdin's in that he argues that the worry can be allayed by coercing the offender to pay his secondary victims. Boonin acknowledges up front that compelling each offender to give "just enough money to every secondary victim" is "of course, completely impractical."[151] But unlike Hajdin, he argues that this is not a reason in itself to reject the response. After all, the fact that it is "impractical to punish a criminal by subjecting him to precisely the amount of suffering that he is thought to merit" does not in itself invalidate punishment. Boonin elaborates, "if the law cannot do precisely what it should or may do, this does not mean that it should or may do nothing at all. It simply means that it should do the best it can to approximate what it should or may do."[152] He gives the example of large class action lawsuits against big tobacco companies, which cannot properly compensate the harm wrongfully done to each of the millions of smokers, but "the state can attempt to

150 *Id*.
151 Boonin, 2008, at 227.
152 *Id*.

approximate this result in a reasonable manner."[153] In applying such approximation to criminal cases, he suggests that a robber, above restitution paid to the direct victim, could be compelled "to pay a lump sum to the city, which would use the money to hire an extra patrol officer, or two, or three, depending on how much police power was necessary to restore the community to its previous level of well-being."[154]

This is a strong response because it avoids the folly that hinders Hajdin and Ellin's responses, namely to attempt to argue that restitution can fully account for harm to every secondary victim. Boonin foresees three objections to this response, and addresses them well.[155] Indeed, if his theory of pure restitution could successfully rely solely on monetary restitution, then I would have little to say against his response to the Harm to Society Objection. But even Boonin concedes that there are cases when purely monetary restitution will not adequately restore the harm done to a community, even by the lowered standard of a "reasonable manner" established by approximating the total harm. He gives the example of an exceptionally talented burglar, who after many thefts "has finally been apprehended and been made to compensate his victims for the harms he has wrongfully caused them." Suppose further that the burglar has paid an extraordinarily high level of monetary restitution, enough to pay for all the stolen goods as well as many new security measures for the neighborhood. Boonin postulates that "even with their new alarm systems, organized community watches, and extra police officers patrolling the neighborhood, his victims have been made objectively less secure by his actions than they were before merely because he is so skilled at evasion and is still free to roam the streets at night."[156] This conclusion seems correct, and it gives one of many examples of why Boonin's theory of pure restitution (or any other defender's version) must resort to nonmonetary restitution. In this case, they may not fully feel secure until the burglar were locked up. I will address the problematic idea of

153 *Id*. at 228.
154 *Id.*
155 The three objections are: (1) restitution made to society differs "too much from the principle of restitution;" (2) that this actually counts as punishment; and (3) that public perception will play too big a role in determining amounts of restitution (at 228-231). His responses to these objections seem adequate so I won't detail them here.
156 *Id*. at 232.

making an offender 'pay' restitution by incarcerating him later. For now, a more unequivocal example of why purely monetary restitution cannot suffice to adequately repair harm to secondary victims is the case of a rich serial rapist. Presumably, despite if he paid millions of dollars in restitution to his primary victims and to improve the community's safety measures, the community would not feel adequately repaired if he were still free to rape and pay later. Admittedly, this conjures up another objection, the Rich Offender Objection, but this is not a problem for my current illustration. I merely aim to show that nonmonetary restitution will be necessary for the most defensible version of pure restitution, and specifically to address the Harm to Society Objection. Because the community would prefer to avoid being victims in the first place, and because the rapist has such deep pockets, the prospect of a rapist who is limited only to monetary restitution as a consequence of his crimes would be deeply unsettling for the community. They would require nonmonetary means of ensuring their safety from the rapist, presumably some kind of restriction on his movement. Boonin suggests that nonmonetary interventions such as the rapists' agreement to be tracked, and his agreement that he cannot come within a certain distance of the community could perhaps be sufficient restitution. Other illustrations of this point will come up later, and they will supplement the strength of my claim here that Boonin and other restitutionists will need to rely on nonmonetary restitution. This brings me to Boonin's second response to the Harm to Society Objection.

This response elaborates on his first response, but makes the additional claim that state-coerced nonmonetary means of paying restitution are justified by his theory of pure restitution. To elaborate his burglar example into his full description of this response, he writes that the burglar could

> ...be compelled to wear a device by which his location could be monitored by the police at all times. He could be subjected to intensive supervision, such as that accompanying probation in some cases, which often includes a curfew. He could simply be locked up. In other cases, an offender might be made to take an anger management course, to undergo therapy, to give up drinking, to stay away from certain people or people under a certain age, and so on. If one or more of these impositions are necessary for an offender's victims to be fully restored to their former level of safety and security, then he owes it to them to undergo these impositions and they could be fully justified by the

theory of pure restitution.[157]

Perhaps the most immediate objection that comes to mind is to insist that one or more of those coercive actions counts as punishment. This is the only objection to this response that Boonin addresses, and I believe that he succeeds in overcoming it. Because in other sections of this dissertation I argue that the actions he lists in the above quotation can be non-punitive, I will not summarize Boonin's response here. But I will raise one further objection that those coercive actions are not restitution.

It is not at all clear that the nonmonetary state-coerced actions Boonin lists can be justified by his theory of pure restitution, as he claims that they are without further argument. Even if some of them could be justified using pure restitution, which is doubtful, they could be justified much more easily and straightforwardly by appealing to crime prevention. To illustrate this point, I will continue using a burglar example in which I claim that the monetary restitution he is forced to pay is justified by the theory of pure restitution, but I argue that his so-called 'nonmonetary restitution' can only be justified by appealing to crime prevention. Imagine that a burglar is finally caught, and he is sentenced to pay restitution to all his victims, including funds to help hire a new patrol officer in order to repay his debt to his secondary victims. His sentence also mandates that he wear a GPS device that tracks his movements at all times and automatically alerts police when he is not where he is supposed to be. Setting aside the question of whether these sentences count as punishment, it is uncontroversial that the *monetary* restitution paid to his victims and secondary victims could be justified by appealing to the theory of pure restitution.[158] However, it would be very strange to insist that the reason that the burglar is forced to wear the device is because by doing so he is giving restitution to his victims by restoring them to a level of previous safety. The device seems unambiguously intended to help the police catch the burglar if he tries to burgle again, and also to deter him from such attempts. Both of these goals are consequentialist reasons to prevent recidivism, and have nothing to do with repairing the harm done to

157 *Id*. at 232-3.
158 *Id*. at 224.

his prior victims. To bring this point into sharper focus, imagine that the burglar had the habit of traveling while he burgled, never visiting the same neighborhood twice. Presumably, after catching him, the state would still be justified in tracking his movements, but this would not be in order to make his previous victims feel safe, but rather to prevent harm to potential future victims.

Now, one could contrive a case in which it is more plausible that the burglar's sentence to wear the device is justified by a debt to repair the harm done to his victims. For instance if he preyed on five or six particular households, repeatedly burglarizing each, then those households would be much more relieved to learn of his capture than they would if they were just one-time victims. But even in this case, it seems like the best reason to justify the device is to *prevent* him from returning to those houses, not because he is redressing the harm done to them. To further illustrate this point, imagine that someone is a victim of a drunken stranger's assault, and that the offender is an alcoholic who is violent when he drinks. The victim wants nothing to do with the offender after he is caught and convicted. The judge decides that the best sentence is to order monetary restitution on top of mandatory treatment for alcoholism. If the judge's reason for giving the order for alcoholism treatment depends, wholly or in part, on restoring the victim back to his original state, then the victim have to be aware of the sentence, and he would also have to be updated on the offender's progress to ensure that the restoration is completed, which could be reasonably assumed to occur when the victim learns that the offender has remained sober and thus poses no further risk to him. But if the victim insists that he not hear a single thing about the offender after his minimum participation in the conviction (and I'm assuming that victim's requests like this one should be respected as much as possible), then it does not seem reasonable that the judge's justification for alcoholism treatment could rely on restoring harm to the victim. The money for restoration, in bright contrast, seems unproblematically justified by restoring the victim.

However, a restitutionist could respond by claiming that preventative measures such as GPS monitoring are simply a means to give restitution back to the community. The restitution takes the form

of increased safety, and the community is composed of anyone who benefits from this increase, regardless of they were aware of the risk or the offender's crimes. This argument would hold that since the offender inflicts ongoing harm to the community by putting them at risk with potential future offenses, then restitution in the form of preventative state interventions should also be ongoing. Essentially, this is biting the bullet in two ways: by admitting 1) that both direct and indirect victims may not be aware of the restitution (i.e. preventative acts to curb the offender); 2) that the restoration of harm does not have to be limited to restoring victims affected by the original offense. The benefit of this argument is that restitutionists can deal with problematic cases in which the victims are unaware of the restitution, or ones in which people who were unaware of the original offense are given restitution to account for their increase in risk, despite not being aware of the original risk or the actions taken to curb it. In order to make this argument, the restitutionist needs to show that people can be objectively harmed by crime, even if they are not aware of it. Boonin does an admirable job arguing this point, and I will not take a stance on that particular question.[159] However, the difference between this account of restitution and the original conception of monetary restitution is striking. One cannot help wondering how much the two actually have in common. An offender paying his victim money for the harm he caused him is the simple, original conception. The amended argument would hold that the following case of preventative detention is also an example of offender restitution: An offender restoring his unwitting victim, who is unaware of both the original offense and the risk the offender's presence in his community would impose, by being 'locked up,' in Boonin's words.[160]

Ultimately, this argument fails for two reasons. First, it adds unnecessary complications to what would otherwise be straightforward justification of preventative state interventions. Secondly, it strays too far from the concept of restitution. To address the first point, Boonin's amended version of pure restitution uses crime prevention practices as a means to restore victims by reducing their risk

159 *Id*. at 253-254.
160 *Id*. at 232-3.

exposure. The problem is that these practices don't need the extra story of restoration in order to be justified: they can be justified simply out of concern to prevent crime. Crime prevention practices are best understood as a means to prevent crime, and any attempt to turn them into a means for something else should be met with skepticism. Not only does the concern of restoring victims seem like an ad hoc addition to the justification of these practices, but Boonin's theory of pure restitution gains important traction from including these practices. Without them, his theory would be untenable because of the fact that many offenders need to be monitored, restricted, treated, or detained in order to maintain a tolerably safe community. So it is difficult to see Boonin's justification of these practices under the rubric of 'restitution' as anything but piggybacking off the inherent plausibility of their crime prevention justification. This can be seen in the number of steps it would take to justify any of these particular practices. For example, take an offender who is convicted of a child molestation that happened years ago in a different city than his current home of Boston. The state would likely be justified in taking a number of interventions with this offender. One uncontroversially plausible intervention following the conviction would be that upon his return to his community in Boston, he would not be allowed to work or live with children. The obvious justification for this intervention is to prevent future harm to children in his community. But Boonin's justification is much more complicated. The offender gives restitution to the community, not with money, but because of the state's restriction on his movement. Although the members of his community remain unaware of his original crime, his restriction restores them to a previous level of objective well being by reducing their risk exposure to the offender. These steps are unnecessary. Crime prevention is an end in itself. Boonin's account uses it as a means of justifying practices under the banner of restitution, despite the uneasy fit.

To address the second point, many of the crime prevention practices that Boonin claims are justified by appealing to restitution stray so far from the concept of restitution that the justification begins to look implausible. The practice of offenders paying restitution, as it was originally conceived

and how it is currently practiced, involves offenders paying their victims. It is an act that the offender takes. Of course the act may be coerced, or the offender's wages may even be garnished in which case he does not himself complete the action of paying the victim. But it is his money, and it flows from him to his victims. As the theory of pure restitution has evolved to address objections, this simple transactional concept has morphed beyond recognition. If an offender is locked up against his will because he is a danger to the community, what is it exactly that he is paying? What actions or resources of his are going to restore his victims? The answers to these questions are not obvious. The conceptual force behind restitution resides in an offender's ability to restore his community, and this force loses justificatory power the further it lapses away from an offender's payment for his offense. Preventative detention is a way to prevent crime; it is not a way for an offender to pay his community back.

To summarize and conclude the discussion above, the second of Boonin's responses to the Harm to Society Objection claims that state-coerced nonmonetary actions such as wearing a GPS device or mandatory treatment are justified by the theory of pure restitution. But for the above reasons, we must reject this claim. These are not instances of restitution. Furthermore, because the success of his first response, that restitution to secondary victims can adequately approximate the harm done to the community, relies on resorting in some cases to nonmonetary restitution, then we must temper his conclusion. What Boonin successfully demonstrates in his response to the Harm to Society Objection is that *in some cases but not all*, monetary compensation can adequately repair harm to secondary victims. This conclusion depends further on accepting his argument that such restitution is not punitive, but I think there is good reason to accept this, and I will give supporting arguments for it at a later point. However, this response is not sufficient to overcome the objection. Furthermore, since Boonin relies on his claim that state-coerced nonmonetary actions are justified by his theory of pure restitution for his responses to other objections, I will refer to my reasoning in this section for rejecting this claim in discussions below.

2.1.3 The Victimless Crime Objection

The Victimless Crime Objection points out that there are many crimes that do not have an identifiable primary victim to whom the offender should pay restitution. This undermines the theory of pure restitution because it calls into question the state's ability to properly respond to such crimes using the justification of repairing harm to victims. In this discussion I address a variety of cases that can be considered to comprise separate objections. For instance, Boonin addresses the cases of victimless wrongdoing (e.g. illegal drugs), failed attempts, and nonharmful endangerment (e.g. drunk driving) as separate objections. But the unifying problem for pure restitution in each of these cases is the concern that there is no discernible victim, and thus no basis for the state to intervene.

Miller notes that controversial topics such as "prostitution, pornography, gambling, and drug use" are unhelpful when addressing the Victimless Crime Objection compared to four less controversial crimes: (1) cruelty to animals, (2) harm to public institutions, (3) attempts, and (4) reckless and/or drunk driving.[161] These four crimes are unambiguously harmful to society, and they are victimless insofar as there is no clear primary victim to whom restitution could be paid. In Miller's words: "On what grounds could persons be prosecuted for such acts if no tangible harm is caused to others? What restitution would be appropriate?"

Sayre-McCord[162] addresses the Victimless Crime Objection by pointing to two tangible harms that the so-called 'victimless' crimes create: first is the "marginal costs" to the criminal justice system to apprehend and prosecute the crimes. "But the harm is often much more than that, since many "victimless" crimes are criminalized precisely because there is a real, albeit diffuse, harm caused by the act."[163] Thus in addition to the costs of prosecuting the crimes, there are the diffuse harms of the crimes themselves. The harms may not fall on a single victim that may be paid, but they are spread out

161 Miller, 1978, at 360-361.

162 Sayre-McCord does not consider his theory one of pure restitution, but his theory of reparations closely resembles Boonin's theory of pure restitution. Both theories rely on the two tenets I identified as the core of pure restitution (non-punitive and aiming to enforce the reparation of harm), and both resort to ascribing power to the state to compel nonmonetary means of restoring harm.

163 Sayre-McCord, 2001, at 519.

among secondary victims. Sayre-McCord offers a further reason to reject this objection that does not rely on harms: that breaking a just and reasonable law is grounds enough for the state to force the offender to make reparations. He does not elaborate this claim, and it seems dubious to argue that the mere breaking of a just law sufficiently grounds reasons for taking a particular response. Advocates of punishment and pure restitution would both like to play such a trump card, but both would need to give further justifications.

Boonin elaborates on Sayre-McCord's secondary victims response to the Victimless Crime Objection. This supplements the move that was made to respond to the Harm to Society Objection: both responses claim that secondary victims may be adequately, if imperfectly, paid restitution for the harms wrongfully done to them. But the response to the Victimless Crime Objection goes further to identify such secondary victims as potential beneficiaries of restitution in the four contested cases that Miller lists. Many of the problem cases that critics highlight, including Miller's four, can in fact be resolved by state-coerced actions that repair harm to primary or secondary victims. For example, someone guilty of animal cruelty could pay for his victim's veterinary bills. It is uncontroversial to say that animals can be victims of crime when they are bodily harmed; and by paying for an animal's care an offender could be repairing this harm. Another uncontroversial example is of an offender who defiles a public institution with graffiti who is coerced to pay for the clean up. Yet another example is of an offender who attempts to harm his victim but fails, and pays his victim for therapy to overcome the trauma of undergoing the potential harm. However, it is not clear how paying for reparations of secondary victims can resolve the force of the Victimless Crime Objection in more problematic cases, such as an offender who kills an ownerless animal or an offender who attempts to harm his victim but without the victim noticing. These objections are further exacerbated when considering the case of a drunk driver.

The case of the offender who kills his ownerless animal victim poses two problems. First, there are no secondary victims. Suppose an offender killed a stray cat and was immediately arrested. Nobody

else saw the attack, and nobody had any ties to the cat, and other animals could not be said to suffer from worrying about an increased rate of animal attacks in their neighborhood. Any secondary victims that a defender of restitution could identify would be stretching the notion of victimhood. But we would still expect the state to take serious actions to condemn the offender's actions and prevent him from further ones. Secondly, reasonable non-punitive responses that a defender of pure restitution might endorse would include nonmonetary state-coerced actions. One example is to compel the offender to go to animal welfare classes or receive treatment with the goal of desisting from future attacks. But these responses seem best justified by prevention of future physical suffering of animals rather than reparation paid to secondary victims of the initial attack.

Secondly, the case of an offender attempting to harm someone who never learns of the attempt poses trouble for pure restitution. Boonin provides a detailed treatment of this specific worry, and concludes that the potential victim was objectively harmed, even if he did not realize it, which provides a basis for coercing the offender to pay reparation for that harm.[164] As an aside, Boonin does not bring up the practical worry that in order to pay reparation *to* the unaware victim, the victim would need to learn about the attempt. But assuming that an offender does objectively harm his unaware victim, and can be forced to pay reparation to him, the worry arises that the most appropriate kind of reparation would be nonmonetary. For example if a stalker whose victim never learns of his stalking is caught, the best course of action would be along the lines of ensuring that he does not come near his victim again, rather than simply paying for the objective harm he inflicted on his victim. As I showed earlier, pure restitution cannot justify such state-coerced nonmonetary actions of this kind. Consider further if the stalker's victim learned of the stalking, but secretly enjoyed it. His family noticed the stalker, reported him, and he was detained. In the investigation, the victim cooperated with his family and the police, never admitting to the true delight he felt while being stalked. The state would surely be justified in making the stalker avoid his victim, but by complying with his orders, the stalker would be actively

164 Boonin, 2008, at 253-254. See also Ellin at 6.

harming the victim, who acutely misses the thrill of being stalked. There is no way that such harmful avoidance may be considered restitution. While this is far-fetched, it further emphasizes the strange contortions necessary to try to explain how straightforwardly preventative actions might be justified as restitution.

Finally, and relatedly, reckless and drunk drivers pose a much more pressing and realistic illustration of the same objection above. Imagine a drunk driver who is arrested soon after he begins driving, before he encountered any other drivers besides the police. Or consider a drunk driver on a particularly empty stretch of highway who is arrested before encountering other drivers. This is another case when there are no secondary victims. The state's recourse would be to compel this offending driver to similar, if not the same treatment as a drunk driver arrested amid busy traffic. Possible actions would include revocation of his driving license and mandatory alcohol classes. Neither of these sentences would draw justification from repairing the objective harm done to the other drivers on the busy road who were, perhaps unknowingly, put at risk. This can be seen especially in the lack of difference between the state's responses to the driver arrested on the empty road and the driver arrested on the busy road. Furthermore, the objection about nonmonetary state-coerced actions comes into play with the same force. I conclude that pure restitution can justify some responses to 'victimless crimes,' but not all, and therefore it fails to overcome the objection.

2.1.4 The Irreparable Harms Objection

Some crimes, most notably rape[165] and murder, result in harms that can never be fully repaired. This is not a problem unique to pure restitution. Punishment cannot repair harm done to a murder victim any more or less effectively than restitution can. However, this objection claims that pure restitution cannot adequately respond to these cases. Boonin does a very thorough job of arguing that

165 Many rape victims resent being referred to in terms of being permanently damaged. This indeed is a paternalistic characteristic of the dialogue. But here I will simply acknowledge this unfortunate fact, and proceed because at least some rape victims do suffer permanent and irreparable damage.

simply because an offender takes on a debt that he will never be able to repay does not mean that the state cannot coerce him to make as much reparation as possible. His arguments in response to this objection are particularly fascinating, and I will not attempt to explain them here. Ultimately I reject his arguments because of one problem related specifically to cases of murder as well as problematic reliance on nonmonetary restitution that is actually not restitution.

There is one serious objection that arises uniquely in the case of murders. This objection states that because the victim is dead, there is no sensible way to explain an offender's actions of reparation to the dead victim, who "cannot be even partially restored to her previous level of well-being."[166] Boonin argues that three undeniable claims can, when taken together, overcome this objection. The first is that a debt owed to someone who then dies is still owed to that person, and the money can be transferred to the estate of the recently deceased. "This is how the law currently operates, and it is uncontroversial."[167] The second is that when someone "cannot pay another precisely what she owes him, she can still be obligated to fulfill her debt by substituting something else of comparable value."[168] For instance if I destroy your irreplaceable painting, then I can still be obligated to pay the approximate value despite not ever being able to replace the painting. The third claim is that it is possible to "put a dollar value on a person's life."[169] Such approximations of valuing life in dollars already are "embedded in the law" and can be seen "[e]very time there is a wrongful death suit."[170] Together these three claims work to overcome the objection that an offender cannot be obligated to make monetary reparations for a debt incurred by murdering someone: the money can be transferred to the victim's estate, the payment can be considered a substitute for the irreplaceable life, and the payment can be according to an approximate value for the life lost. But this response leaves Boonin to face at least one objection that he cannot overcome: the case of a murder victim who "has no descendants and no

166 Boonin, 2008, at 240.
167 *Id*. at 241.
168 *Id.*
169 *Id*. at 242.
170 *Id.*

will."[171] In this case it would not be possible to transfer payments to a victim's estate, thus undermining the first part of Boonin's response. Boonin acknowledges this objection and responds that it can be dealt with one of two ways: either the debt can be transferred to the state or it can "be claimed by interested parties."[172] Boonin does not elaborate on how an offender paying the state or interested parties can be construed as repaying the direct debt owed to the victim. The plausibility of the transferability claim that a debt can still be paid to someone deceased relies on the possibility of transferring assets to an estate directly related to the deceased. If there is no way to do so, then it seems implausible to continue to claim that such transfer can take place.

Boonin might respond that this is a question of degree, and that if the state or interested parties are the "beneficiar[ies] of the victim's legacy then any debts to the victim become debts to the state."[173] But this is a degree too far. The transferability claim is credible because the direct relatives or the estate of a deceased person are reasonable proxy recipients for him. The reason we believe that they are reasonable proxies is because of the directness of the relationship to the deceased: he was the father of his surviving children, he was the creator of his estate, the writer of his will. There is no such directness in his relationship to the state. Boonin's inclusion of the phrase "interested parties" hints at the confusion that this strategy would entail. Suppose a victim with no descendants and no will was known to hate paying taxes but who loved his local professional sports team. Arguably there is a much more direct relationship between the victim and the sports team than between him and the state. Would the sports team be eligible to claim the restitution? I know this is a ridiculous scenario, but I bring it up to highlight the confusion that would ensue without the clarity that direct relationships provide. Without descendants or a will, we cannot guess the intentions of the deceased, and the strength of the transferability claim breaks down. We no longer have reason to believe that we can pay the victim by proxy.

171 *Id.*
172 *Id.* at 243.
173 Benatar, 2013, in private comments.

Beyond this unique objection, there is a more general objection based on recourse to nonmonetary restitution. In discussing irreparable harms, Boonin never explicitly claims that nonmonetary state-coerced actions can be justified based on restitution owed to direct crime victims. (He only defends this based on restitution owed to secondary victims, which I discussed above). But his defense against the Irreparable Harms Objection implicitly relies on this claim. In other sections of his defense of pure restitution he references subjecting offenders to probation, GPS monitoring, curfew, incarceration, therapy and treatment. This full range of actions would be necessary to sufficiently handle rapists and murderers and meet the minimum requirements of a crime response regime. However, as I argued above, we should view these actions as preventative rather than examples of restitution.

2.1.5 The Inadequate Deterrence Objection

I move now to Mane Hajdin's response to the criticism that pure restitution would not adequately deter offenders. Most crimes are not solved,[174] and in a regime of Barnett's pure restitution, offenders who are caught would only owe restitution related to their crime. This objection posits that under such a regime if an offender considered buying a TV or stealing one, the rational choice would be to steal one. This is because he likely would not be caught, and if he were caught he would only have to pay a price similar to what he would have paid to buy it. Barnett, Ellin and Boonin all point out that the offender should also pay more than the price of the TV for the costs of catching and prosecuting him. Hajdin argues further that pure restitution can overcome the objection by treating offenders as gamblers. The cost to criminals should be higher than mere 'full restitution,' and instead should take account of the criminals' gamble that they won't be caught. In other words, if criminals are caught they should pay significantly more than what they owe to repay their gamble that they would not be caught. This way the criminal would be properly deterred when considering crime. Of course, an

174 Kershnar, 2012, at 74.

individualized calculation for the gamble involved in each type of crime seems like an implausible task, but Hajdin proposes a way to overcome this: "For a certain period we establish the ratio between the total amount of damages that are due to criminal acts of a certain type" that are not prosecuted and the total amount of restitution due if all cases were prosecuted, then "we increase the amount that each of the convicted offenders has to pay by that percentage."[175] What would happen with the surplus income? Money over and above the victims' pure restitution would go to a fund that would compensate victims of unsolved crimes. This way, "[c]riminals as a group would pay for all the costs of criminal activities."[176] This is a clever response, and I suspect that such a regime could arguably have as much of a deterrent effect as punishment. One objection would be the plausibility of actually extracting so much restitution from offenders, given the high percentage of poor offenders. This anticipates the Poor Offender Objection, which I find more forceful and interesting, so I will move forward to discussing rich and poor offenders.

2.1.6 The Rich Offender Objection

As I noted earlier, the possibility of an extremely wealthy rapist who continues to rape victims and pay the millions of dollars of restitution poses serious trouble for pure restitution. The only plausible response to this scenario would resort to nonmonetary state-coerced actions, which is problematic for the reasons I gave above. But even if the crime were less serious than rape and thus might not require nonmonetary responses, it would still seem problematic for wealthy people to 'buy' crime. Abel and Marsh[177] and Sayre-McCord[178] both answer that the restitution can be measured in effort, or work hours, rather than money. This would level the playing field. One might object to this response because if the goal of pure restitution is to repair harm, and a rich offender can adequately do

175 Hajdin, 1987, at 81.
176 *Id*.
177 Abel and Marsh, 1984, at 185.
178 Sayre-McCord, 2001, at 508.

that with money, then forcing him to work additional hours would be redundant. It would also be open to criticism that the intention is actually to make him suffer, which would count as punishment.

Defendants could respond that the goal is to make the currency of payment fair, rather than intending to make the offender suffer. To make an offender work, either at his normal job or one provided by the state, seems like it could be justified by the goal of repaying a debt to the victim. Working to repay debt is much more straightforward than wearing a GPS device to repay a debt. But on closer inspection, we must reject this response. Imagine a wealthy unemployed playboy who enjoys spending his time stealing expensive objects and destroying them. If he is caught he could be forced to pay restitution measured by hours of work in a government-provided job. This would overcome the trouble of his ability to 'buy' the crime. But suppose the victim simply wanted the money for restitution as quickly as possible. If the offender were forced to pay the victim but then still work additional hours, the justification of pure repayment would seem dubious for the additional hours of labor. And if the victim were told to wait until the offender completed the labor, this would undermine the explicit goal of repairing harm to the victim.

2.1.7 The Poor Offender Objection

The Poor Offender Objection draws its force from urgent practical problems rather than theoretical inconsistencies. Most offenders are poor. Unsurprisingly, there is a wealth of data showing that poor offenders cannot pay victim restitution even when coercive action is taken to compel their payment.[179] This fact leads to the argument that if poor offenders cannot reasonably be expected to afford the restitution they owe, then the theory of pure restitution should be rejected. Secondly, if poor offenders know that they cannot be forced to pay restitution because they have no assets, then there would be little to deter a poor offender from committing crime. Defendants might counter both concerns by substituting monetary payment with work hours. Poor offenders could be compelled to

179 Dickman, 2009, at 1695.

work, either in the private sector or on state projects. This would eliminate the concern over deterrence, as well as provide a currency that even poor offenders could use to repay: their effort.

Theoretically, this response works well. It avoids the problem of redundancy that plagues the same response to the Rich Offender Objection because the poor offender has no option to simply pay the restitution. The nonmonetary state-coerced action can also be justified straightforwardly as effort directly linked to repaying the offender's debt. But the drawbacks are in the details of implementation. Consider two offenders guilty of the same serious robbery, both reluctant to pay the high restitution. One has enough funds in his bank account and the other has no assets at all. Pure restitution would justify treating the two accomplices extremely differently. One of them would begrudgingly pay the restitution, while the other would be consigned to a forced labor camp. Even presuming that the poor offender's labor earns him minimum wage, the number of hours required to repay the restitution in full would be daunting. If the offender were also charged for the marginal costs added to the criminal justice system, such as the costs of his arrest, prosecution, board and housing during the labor camp, as well as restitution to the community's secondary victims, then the robbery could very easily cost this offender years of forced labor. The stark difference between the two accomplices, both guilty of the same crime, one simply having to write a check and the other consigned to forced labor for decades is reason enough to reject this response. Even if one does not subscribe to the retributivist ideal of equivalent treatments of equivalent wrongdoings, it is easy to see how the resulting situation would be untenable. An egregious divide between forced labor camps for the poor offenders and writing a check for the rich ones is a politically and historically troubling notion. Given the racial disparity currently evident in the US criminal justice system, where a disproportionate percentage of poor black males are incarcerated, the switch to a system where fewer wealthy whites were locked up, and the poor black men were forced to labor for years would be unpalatable, to say the least. But suppose that the currency were measured instead by hours of work rather than the wages of those hours. In this case the state could approximate the rough equivalent of the value of the restitution and the hardship of a certain

number of hours of labor. This response is also flawed because it crosses the line of what pure restitution could reasonably justify. If the state goes to the trouble to provide the infrastructure for offenders to work, but knows that the work will not result in full restitution, then the state would be paying for the offender to work. Why not take the money and compensate the victim instead? Surely the money spent on the security, food and housing of offenders in labor camps could instead be directed to the victims of the crime in the spirit of repairing the harm done to the victim.

However, some might not share my intuitions about proportionality, in which case they may not see the "poor offender objection" as a problem for pure restitution. But this will not change the overall conclusion that pure restitution fails due to the other forceful objections.

2.1.8 The Restitution is Actually Punishment Objection

It will be useful now to take stock of what I have and have not argued in this section. In many but not all cases of crime, the theory of pure restitution must resort to state responses that it cannot justify. One of the two tenets of pure restitution, is that "the state should compel the offender to restore the victim to the level of well-being that the victim rightfully enjoyed prior to the offense."[180] The other tenet is that the state should not punish. Many necessary responses to crime simply cannot be justified by appealing to restoring victims, or repairing the harm done. However, the goal of restitution can justify some responses to crime. What I have argued is that the theory of pure restitution is false, but this does not mean that the goal of restitution itself cannot justify non-punitive compulsory victim restitution. I now turn to whether this is possible.

Even assuming that the theory of pure restitution is false, cases of compulsory victim restitution may still be construed as punitive. Take the case of the state compelling a vandal who breaks a window to pay for repairing the window. This is clearly justified purely in terms of restitution. But one might still object that while it repairs harm, it is also punitive. Forcing the offender to pay clearly harms the

180 Boonin, 2008, at 224.

offender. However, based on the definition of punishment that I defended in the first chapter, this does not qualify as punitive. The intent is to repair the window, not to cause the offender suffering. There is some evidence that the general public supports this conclusion. In a study that tracked participants' responses to scenarios of victim restitution and punishment, Tsoudis found that people are more likely to view compulsory victim restitution as non-punitive, which "further supports the position that victim restitution is more compensation than punishment."[181]

However, consider if the offender refuses to pay restitution. It seems that any mechanisms of enforcing his compliance would be punitive. Cholbi writes that if an offender refuses to comply, then the state could respond by "forcing her compliance in a literal sense (using electronic funds transfer to garnish her wages, for instance) or by resorting to other coercive measures, such as incarceration. Such responses amount to punishing her non-compliance."[182] And Kershnar writes, "The likely means of ensuring payment is via threat of… punishment"[183] These are compelling arguments. If the state orders an offender to pay restitution, this alone does not qualify as punitive because the harm involved is not intended. However, if the he refuses to pay, and the state threatens to put him in prison, then the threat of intended harm is introduced. And if the state carries out this threat, then it seems clear that it has punished the offender for non-compliance. There does not seem to be a good response to argue against this assessment. A relevant US Supreme Court case shines light on this issue: in 1983 the Court ordered that a sentencing court cannot imprison someone who has made good efforts to pay his court-ordered restitution debt but cannot make the payment due to "the poverty or inability of the probationer to pay the fine and to make restitution." The Court's opinion shows that it considered such imprisonment to be punishment. It held that a probationer cannot be imprisoned "for failure to pay a fine and make restitution, absent evidence and findings that he was somehow responsible for the failure or that alternative forms of punishment were inadequate to meet the State's interest in punishment and

181 Tsoudis, 2000, at 494.
182 See Cholbi, 2010, at 86.
183 Kershnar, 2012, at 73.

deterrence."[184] In other words, the Court held that the probationer's poverty is not his responsibility, therefore it cannot be held against him as reason for further punishment if he makes good effort to try to pay restitution but cannot meet the full payment.

Cholbi and Kershnar are correct to point out that many cases of enforcing compulsory victim restitution are punitive. They are incorrect, however, in arguing that all cases of enforcing restitution are punitive. One notable exception is Cholbi and Kershnar's example of wage garnishment. If an offender does not pay restitution and the state threatens to garnish his wages, it is not threatening to intentionally harm him. Rather, it is threatening to forcefully do what he was ordered to do: pay the victim.

I conclude that most cases of actually enforcing compulsory victim restitution are likely punitive, with the notable exception of wage garnishment. The court order to pay restitution is not punitive, but if it comes along with a threat of punitive enforcement, then its enforcement will entail recourse to punishment.

2.1.9 Looking Forward: An Argument for Voluntary Restitution

To summarize, the theory of pure restitution cannot overcome all objections (or perhaps even any of the ones I covered), so it is false. Nonmonetary state-coerced actions often cannot be justified by the goal of repairing harm by paying restitution. And enforcing compulsory victim restitution will require punishment. Pure restitution, to put it mildly, could not be a standalone alternative to punishment. So how would it fit into a non-punitive post-arrest regime? In this section I examine the hazards of trying to enforce restitution. In looking at how restitution can be best applied, it is helpful to note examples of how it can go wrong. With this practical framework in mind, I argue that restitution is best treated as a voluntary option that should be given to offenders. It is in everyone's interest for the state to take this avenue seriously. It should facilitate and protect as many opportunities as possible for

184 See the Syllabus of *Bearden v. Georgia*, 461 U.S. 660, 662-63, 668-69 (1983).

offenders to pay restitution. One of the main but chronically overlooked benefits of such a system is rehabilitation for the offender.

To begin, the ruinous effects of mandatory restitution imposed on offenders regardless of their ability to pay can be seen today in the USA. In 1996 the passage of the Mandatory Victims Restitution Act (MVRA) "made restitution mandatory in almost all cases in which the victim suffered an identifiable monetary loss, which removed judicial discretion from the imposition of restitution orders. The MVRA also removed judges' ability to fashion restitution orders based on an offender's ability to pay."[185] Unsurprisingly, the MVRA has resulted in the explosion of criminal debt, growing eightfold in the first decade after its passage, and the vast majority of this debt is unpaid victim restitution.[186] In a moment of remarkable candor, the Department of Justice admitted that the MVRA "has resulted in a large surge in criminal debt, but it has not resulted in any appreciable increase in compensation to the victims of crime, in most cases, because of the defendants' inability to pay."[187] Because over 85% of arrested offenders are impoverished, and because their assets gained by criminal activity are seized by the state, and because they have low levels of training and skills, and because their job prospects diminish after their release, and for a wide variety of other factors contributing to their poverty,[188] "judges have likened the collection of restitution to 'get[ting] blood out of a stone.'"[189] There is evidence that suggests offenders do not try as hard to fulfill their restitution orders when they feel that the orders are unfair and unrealistic compared to when the orders are tailored to their situations.[190]

In addition to victim restitution, the state has tried to recoup the enormous costs of mass incarceration and other criminal justice expenses by imposing fees and fines on top of restitution debt. Legal debt is long term and subject to interest, and it is "typically substantial." In fact, "By 2008,

185 Dickman, 2009, at 1688.
186 *Id*. at 1691-2.
187 *Id*. at 1694, quoting: "Letter from Clarence A. Lee, Assoc. Dir., Admin. Office of the U.S. Courts, to Gary T. Engel, Dir. of Fin. Mgmt. and Assurance, U.S. Gen. Accounting Office (June 6, 2001), in GAO, CRIMINAL DEBT 2001, at 105."
188 Dickman, 2009, at 1695.
189 Id., quoting: "R. Barry Ruback, The Imposition of Economic Sanctions in Philadelphia: Costs, Fines, and Restitution, FED. PROBATION, June 2004, at 21, 25."
190 *Id*. at 1697.

defendants sentenced in 2004 had been charged an average of $11,471 by the courts over their lifetime."[191] This additional debt has more and more frequently translated into imprisonment. While this may seem implausible given that debtors' prisons were banned in 19th century Europe, and also given the 1983 Supreme Court case I cited above, it is an increasingly common practice. The reason for this is specifically the practice of arresting debtors "not for nonpayment, but rather for civil contempt of court—that is, failure to comply with a court order to pay their financial obligations."[192] This allows the courts to imprison debtors without having to charge them for nonpayment. "In short, this civil route to prison or jail is not only characterized by relatively few procedural roadblocks, but has enabled judges and others to revive an especially controversial practice—the incarceration of indigent debtors."[193]

So what are the effects of these policies and practices? They burden an already penniless population with massive debt. This debt follows offenders for years and harms them in various ways: lowering their credit rating, decreasing their likelihood of securing housing and income, bankrupting family members. And increasingly it results with imprisonment for failure to pay.[194] Put succinctly, the "epidemic adoption of fines and fees on top of restitution requirements and child support obligations has contributed to a massive problem for the poor without justification."[195] Another effect is that victims are less satisfied with the compensation they receive.[196] I will discuss this further in the next section where I argue for state compensation. Finally, by undermining offenders' confidence in their ability to actually pay their restitution, the MVRA has compromised the potential rehabilitative effects of restitution. To summarize: low rates of actual victim compensation, massive unjustified debt on the poor, a pernicious return to debtors' prisons, and a burgeoning bureaucracy to handle the Sisyphean task of collecting the un-collectable. This leads me to conclude that there are forceful practical reasons

191 Beckett and Murukawa, 2012, at 228.
192 *Id*. at 229.
193 *Id.*
194 ACLU, 2010.
195 Katzenstein and Nagrech, 2011, at 565.
196 Dickman, 2009, at 1701.

to reject implementing a system of enforced victim restitution.

The original function of offender's paying restitution to their victims was to rehabilitate offenders. "[I]n the late nineteenth and early twentieth centuries, it was imposed primarily to promote the responsibility and rehabilitation of offenders, not to compensate victims."[197] In 1979 Galaway wrote that public opinion was beginning to change, and over the past three decades we have witnessed an almost total reversal. But Galaway warned back then what we have seen now: "Although restitution may benefit the comparatively small number of victims of captured and converted offenders, it will not be an effective program for meeting the victims' needs."[198] Even the US Supreme Court acknowledged that restitution while on the surface seems oriented toward the victim, it is in fact focused on offenders, either as a sanction or as rehabilitation. The Court held that "[a]lthough restitution does resemble a judgment 'for the benefit of' the victim, the context in which it is imposed undermines that conclusion." And further the Court claimed that criminal court-ordered restitution should not be construed as a means of compensating the victim "[b]ecause criminal proceedings focus on the State's interests in rehabilitation and punishment, rather than the victim's desire for compensation."[199] That was 1986. Unfortunately, by the time that the MVRA was passed, the shift in approach to restitution had become complete.

There are proven positive rehabilitative effects of restitution. For instance, one study showed that offenders who paid a higher percentage of their restitution orders "were less likely to commit new offenses." Another study showed that juveniles who participated in restitution were less likely to reoffend than their peers who did not participate.[200] Furthermore, voluntary acts of contrition and restitution are good signals of desistance, pointing to "the absence of repeated behavior among those who had established a pattern of such behavior."[201] Ward and Salmon note that voluntary acts of

197 *Id.* at 1701.
198 Galaway, 1979, at 57.
199 Dickman, 2009, at 1703, quoting: "Kelly v. Robinson, 479 U.S. 36, 52-53 (1986)."
200 Dickman, 2009, at 1703.
201 Maruna, 2012, at 79.

restitution are an invaluable part of communal reconciliation. They write that, "[t]here are two facets to reconciliation that are clinically relevant: offenders' obligation to apologize and make reparations and the community's obligation to help the offender reintegrate back into the community once hard treatment is served."[202] While I have shown many of the potentially harmful effects of mandatory restitution, I don't want to underemphasize the potential positive benefits of opening channels of restitution for offenders. There is no reason that an indigent offender in prison should not have the opportunity to pay restitution. He could pay with his hours of labor, possibly at community service, or at voluntary employment projects in order to make money that he would then choose to send to the victim. The rehabilitative and reconciliatory benefits make voluntary restitution something that should be available for every offender.

2.2 Victim Compensation

In this brief section I argue that state compensation for victims of violent crime is an essential part of any crime response regime. I give three justifications for the claim that violent crime victim compensation is the state's obligation. First, such compensation is an essential public good, and the state is obliged to provide such goods. Secondly, because the state has a duty to protect its citizens, the state should act on the presumption that it owes compensation to violent crime victims for failing to protect them. Third, when a law is broken, the state is obliged to communicate certain values and commitments. Following cases of violent crime, compensating the victims is the best, most direct way for the state to acknowledge the seriousness of the harm, express collective sympathy for the victim, express remorse in not having prevented the harm and take responsibility in working to redress the harm.

Before addressing these justifications, I want to make some background points and clarify terms. As a reminder, when I use the term 'compensation' I am referring to payment made from the

202 Ward and Salmon, 2009, at 245.

state to a victim. Compensation is most likely monetary, but there is no reason that the state should be limited to monetary compensation. Providing access to psychological treatment seems like an unproblematic way to compensate a rape victim. Also, when I refer to 'victims' hereafter I refer to "victims who have sustained personal injury directly caused by an intentional crime against the victim's life, health or personal integrity."[203] This definition comes from the Commission's Communication on Standards and Actions concerning crime victims in the European Union, and it has the advantage of including deserving victims of sexual, racist or xenophobic crimes that may not qualify as 'violent' but still merit inclusion. For the sake of simplicity I refer to 'intentional crime against the victim's life, health or personal integrity' simply as 'violent crime.'

I limit the state's obligations to victims of intentional violent crime for two reasons. The first reason is that insurance can feasibly cover most victims of accidental violence and property crime, while "there is no general type of insurance" for violent crime.[204] Secondly, civil law provides avenues for seeking damages in all these cases of crime, but because violent crime victims usually do not pursue civil lawsuits, and because the state has certain obligations to them that I will enumerate below, they should be eligible for compensation in addition to the option to pursue a civil claim. The reason victims normally do not pursue civil claims is because they understand that, most often, the offenders will not be able to afford the damages owed.[205] For this reason civil lawsuits are usually not pursued unless the defendant is an institution with deep pockets, and even then, it is enough of an imbroglio to deter most potential claimants. Thus, without insurance, and without reliable access to civil damages, violent crime victims are left out in the cold. And as I've explained, restitution from offenders is an unreliable source of compensation. Galaway writes that, "Although restitution may benefit the comparatively small number of victims of captured and converted offenders, it will not be an effective program for

203 Buck, 2005, at 164, quoting: "COM (2002) 562, p. 84."
204 Buck, 2005, at 159.
205 See e.g. Sarnoff, 1996.

meeting the victims' needs."[206] One might argue that victims would be more satisfied to receive restitution from the offender. However, research from Australia shows that victims are primarily concerned with being compensated, and the source of the compensation does not matter to their level of satisfaction. Indeed, because the state can much more reliably compensate victims, it results in much more victim satisfaction. Victims who are compensated by the state are even much more willing to forgive the offender.[207] Unfortunately, most victims of violent crime currently do not apply for the scant public compensation funds that are available.[208] Due to the justifications I am about to explain, the state already has an obligation to compensate them, and the lack of viable alternatives supplements this argument.

I move now to the justification of violent crime victim compensation as an essential public good. George Klosko provides a novel solution to the puzzle of why we consider it obvious that governments should provide benefits to their citizens to the exclusion of citizens in other states that may be far more needy. In the course of answering this, he focuses on "people's need for a range of public goods, bearing mainly on security in different forms from different threats, if they are to lead acceptable lives."[209] He elaborates with the example of the state's protecting citizens from crime: "Because of the special relationship she bears to state X, a given individual, Smith, is required to obey X's laws. But her obedience comes at a price. It is in exchange for benefits, especially security, the state provides, which are provided only to people who are required to obey."[210] In other words, there are public goods that are uniquely available to citizens of a particular state because the existence of a particular good depends on the mandatory participation of its citizens in return for the state's providing that good. This explains why the good cannot be extended past the borders of the nation state. The law is one such example. Citizens are obliged to follow its state's laws, and the state enforces this mandate.

206 Galaway, 1979, at 57.
207 Ristovski and Wertheim, 2005, at 67.
208 Alvidrez et al., 2008, at 886.
209 Klosko, 2009, at 245.
210 *Id.*

As a result, the state is obliged to provide the public good of protection to its citizens. As Klosko writes, "Insofar as the necessary cooperation is coordinated by law, citizens are required to obey the law. It is only through general adherence to the law that an environment of law and order is able to exist... Circumstances are similar with other essential public goods."[211] Klosko gives three criteria for what can be considered one of these essential public goods: "The goods in question must be (1) worth their costs, (2) indispensable to satisfactory lives, and (3) fairly distributed."[212]

The public good of violent crime victim compensation meets these three criteria. Every state, whether it is a welfare state or not, spends considerable time and energy protecting its citizens against crime. The costs of prevention are clearly worthy, and the cost of redressing harm after the crime is equivalently worthy. Violent crime victims are changed dramatically, and in many cases they need help to restore themselves to satisfactory lives. Furthermore, by focusing on the most serious transgressions of the laws that the state obliges us to follow, this compensation avoids potentially disqualifying concerns about fairness of distribution. As I noted earlier, violent crime victims cannot fall back on insurance or other means of compensation. And the seriousness of the breach of laws invokes one of the state's fundamental obligations: to protect its citizens. As Kolsko notes, a state's obligations are based "mainly on security."[213] The state may protect people by preventing crime, or it may protect victims from the ongoing negative effects of living through violent crime. It seems uncontroversial that protecting citizens is a primary obligation of any state, whether it is a welfare state or a minimal state.

The second justification for the obligation of compensation assumes that the state has an obligation to protect its citizens from violent crime. Given this duty, every instance of violent crime could be seen as a failure of the state to fulfill its duty. Of course, the state cannot ever be reasonably expected to prevent all violent crime. Thus it will probably be unclear from case to case whether the state is in fact responsible for a lapse in its responsibility. Given this uncertainty and the foundational

211 *Id*. at 252.
212 *Id*. at 258.
213 *Id*. at 245.

claim that the state's duty to protect is a very significant one, the state should then act on the presumption that it owes violent crime victims compensation for failing to protect them.[214]

The third justification stems from the argument that the state is obliged to express certain values and commitments following an instance of crime. This thinking receives forceful treatment from Feinberg and Duff, and although they were attempting to use state expression of values to justify punishment, the claim still maintains its force. And justifying compensation is far less problematic than justifying punishment because of the trouble involved with attempting to justify the harmful treatment of offenders. As I wrote earlier in this dissertation, there are many functions of the state expressing its condemnation of an offender's actions: it allows the state to disavow the offenders' actions; it serves to "speak in the name of the people" the state's symbolic nonacquiescence; it vindicates the law by "emphatically reaffirming" it when it is broken and it validates the rights of victims that were transgressed.[215] These functions are all fulfilled by victim compensation. Compensation condemns the moral wrongness of the breach of lawfulness by expressing each of the above messages. But in addition to acknowledging the seriousness of the wrong done and condemning it, compensation is uniquely well suited for expressing collective sympathy for the victim, remorse in not having prevented the harm and willingness to take responsibility in working to redress the harm. Buck gives a forceful argument that compensation should be given to victims out of "social solidarity."[216] A final advantage of state-funded compensation is that it can be paid to the victim without any action required from the victim (for instance pursuing the harrowingly complicated process of civil compensation). For this reason payment can also be expedited to the victim.[217] See Felduthusen et al. for excellent research and recommendations on how to smoothen this process for the victim. Working with compensation schemes in Ontario, they conclude that many victims would prefer to be compensated by the state, and

214 Buck, 2005, at 150-151 also argues this point. She points out that the German state compensation fund was enacted with the explicit goal "to compensate crime victims because it has failed in its duty to prevent and control crime." (n.33).
215 Feinberg, 2011, at 117.
216 Buck, 2005.
217 See Cobley, 1998, at 232.

then have the additional option to pursue civil damages if they wish.[218]

The justifications I have given may each contribute to each other, or they may be regarded as individually sufficient. For these reasons, I conclude that state compensation is the best way to repair harm done to victims of violent crime. It is not the only way, but it is essential. To conclude the whole chapter, another avenue that should be available is voluntary restitution. Mandatory restitution should be used sparingly with wealthy offenders and acknowledging that it may lead to punishment if it is enforced.

218 Felduthusen et al., 2000, at 113.

Chapter Three: Rehabilitation and Incapacitation

An alternative post-arrest regime to the current punitive model could make use of the hard and potentially harmful but non-punitive interventions that are commonly associated with punishment. Examples of such interventions include drug rehabilitation, sex-offender therapy, preventive detention and electronic monitoring among many others. These interventions are essential for any non-punitive regime to meet the first requirement for crime response which I discussed earlier, namely to maintain public safety and order. They fall under two broad categories: rehabilitation and incapacitation.

There is a wide range of offender-focused state interventions that I will broadly refer to as 'rehabilitation.' This range includes relatively non-intrusive interventions such as community service orders, and it extends to include highly intrusive interventions such as sex offender therapy.[219] Another range of interventions, conceptually distinct from rehabilitation, involves incapacitating offenders to various degrees, from requiring them to wear an electronic monitor to detaining them. Rehabilitation aims to change an offender's future behavior in the community by bringing him through a course of treatment. On the other hand, incapacitation aims to prevent crime by restricting the freedom of the offender.[220] One question that arose in Chapter Two becomes the central focus of Chapter Three: whether enforcing a non-punitive court mandate will involve recourse to punishment. Regarding mandatory restitution, I concluded that enforcing restitution orders, which by themselves are non-punitive, will inevitably include recourse to punitive measures (with the exception of wage garnishment). In this chapter I will come to similar conclusions for rehabilitation and incapacitation. Rehabilitation in some cases will need to be mandatory; in those cases, enforcing the mandate will include recourse to punishment. Similarly, incapacitation, while non-punitive itself, oftentimes takes

219 See Rotman, 1990, at 3-4 for an extensive list of rehabilitative interventions.

220 While these are conceptually distinct, there are some interventions that are both rehabilitative and incapacitative, as I will discuss later.

the form of a threat of punishment. While the threat is distinct from the punishment, any state engaged in common restrictive practices that qualify as incapacitation will have to rely on punishment to carry out the threat. In this chapter I will argue that a post-arrest regime that relies on rehabilitation and incapacitation (and any viable regime *must* rely on these) also must rely on punishment out of necessity. This conclusion is enough to further conclude that there is no viable non-punitive crime-response regime. However, this conclusion does not entail that a regime relying on rehabilitation and incapacitation will take recourse to punishment in each individual case. In fact, I argue that the opposite is true: a regime using punishment as a viable threat should be able to avoid punishment in most cases involving rehabilitation and incapacitation. Furthermore, the effectiveness of these interventions, coupled with the state's obligation to minimize suffering according to the concept of parsimony, means that the state is obliged to attempt to work to avoid recourse to punishment except in cases which necessitate its use. And as I will discuss in the final chapter of this dissertation, the recourse to punishment can and should be minimized in such a way to drastically change current public policy.

I begin my discussion of offender rehabilitation by discussing its definition, the various paradigms of offender rehabilitation throughout history, as well as the variety of treatments that fall under its scope. Then I will evaluate the justifications for rehabilitative interventions, address objections and explore the relationship between rehabilitation and punishment.

3.1 A Definition of Offender Rehabilitation

Colloquially, when people refer to offender rehabilitation, they refer to a collection of practices aiming to 'make the offender right.' In this section I aim for a precise definition. Defining offender rehabilitation in the context of this dissertation has a similar potential pitfall to defining punishment. The definition of rehabilitation could be unfair if it smuggles in justification(s) for rehabilitative practices. The question of whether rehabilitation can be justified is separate from defining it, and the

justification question should be approached similarly to the way scholars have attempted to justify punishment. Separate from smuggling in justification, another concern is that a definition might unfairly preempt the question of whether a rehabilitative intervention is also punitive, that is, it may define itself away from punishment without proper explanation of why the two practices are separate. But on closer inspection, this is not a problem. In this section I argue that an intervention cannot be both rehabilitative and punitive. This is not an outrageous position, and is in fact very similar to how I discussed restitution. While I claim that restitution is definitionally distinct from punishment, this poses no problem for discussing whether mandatory restitution involves recourse to punishment. In this chapter I apply the same skepticism to whether mandatory rehabilitation can be non-punitive.

Another problem is that the term 'offender rehabilitation' includes a messy cluster of concepts and connotations. In 1959 Allen underlined this point hyperbolically: "No idea is more pervaded with ambiguity than the notion of reform or rehabilitation."[221] While the concept of offender rehabilitation does encompass a wide variety of practices, this does not disqualify it from a precise definition. Legal punishment also encompasses a wide variety of practices, and the term 'punishment' also connotes many ambiguous meanings in everyday speech, but clearly this does not keep us from arriving at a precise definition of legal punishment. There have been many definitions offered for offender rehabilitation (hereafter I will use 'offender rehabilitation' and 'rehabilitation' interchangeably). The definitions offered often focus on the ultimate goal of treatment, such as behavior change and desistance. While definitions and characterizations abound,[222] I am not aware of an attempt to defend a precise account of the necessary and sufficient conditions of rehabilitation.

What does rehabilitation have in common with punishment? Both are authorized, or implemented by the state. Both are retributive, or given to an offender in response to his committing a crime. Of course there are offenders who volunteer for rehabilitation in private centers, but because

221 Allen, 2011, at 98.
222 See Allen, 2011, at 98. Hart, 1968, at 26, Duff, 2001, at 5. Pollock, 2011, at 316.

their treatment was not part of a sentence, it would be considered just 'rehabilitation' rather than 'offender rehabilitation.' Rounding out the commonalities, both punishment and rehabilitation are interventions intended by the state. One major difference is that the intention behind punishment is to harm the offender, which is not true of rehabilitation, in which the intention is to treat the offender.

Our definition of rehabilitation includes: authorized, retributive, intended treatment. I want to add a further condition: to qualify as rehabilitation, the imposition should include a minimum level of consent by the offender. Thus my proposed definition of criminal rehabilitation is: **authorized, retributive, consensual, intended treatment.** Both the 'consensual' and 'treatment' aspects of the definition need elaboration. For discussion of the three other requirements of the definition, I refer the reader back to my discussion of the definition of punishment.

3.1.1. The Consensual Requirement of Rehabilitation

If an offender dissents from treatment that the state forces on him against his will, or if the state treats him without his knowledge, then these interventions are not rehabilitation. Rehabilitation requires at least a minimal level of the offender's participation and involvement, and this minimal level is what I am using to mean 'consent.' Rotman writes that rehabilitation "demands fully informed and willing intelligent participation."[223] Before arguing that consent is a necessary component to rehabilitation, I should further specify what I mean by 'consent' in relationship with coercion, because the minimal level I am referring to is different from how the term is normally used. Many discussions of consent in the applied ethics literature refer to very robust notions of consent. In medical ethics, the consent of a patient is treated with utmost care to make sure that the patient's choice has not been coerced in any way. For instance, a potential kidney donor goes through a battery of tests and interviews, including extensive questioning regarding his relationship with the recipient to determine whether there is any

223 Rotman, 1990, at 9.

evidence that the choice is being coerced in any manner.[224] This firm notion of consent is not what I am using to define criminal rehabilitation. When I use 'consent' to refer to rehabilitation, I mean that the offender has the physical ability to choose otherwise. The state might strongly coerce the offender to participate, but so long as it does not physically force his participation, the program may qualify as rehabilitation. This is the minimal, weakest version of consent. It also aligns with our thoughts about rehabilitation programs outside of the legal context. Consider a drug addict who is considering enrolling in the private drug rehabilitation clinic. This addict has lost his job, much of his savings, and his sense of self worth. He is desperate to overcome his addiction. Furthermore, his wife tells him that if he doesn't successfully complete the program, then she will leave him and take their children with her. The addict's decision is being intensely coerced, yet we would still consider his enrollment in the program to be voluntary at this minimal level. He consents insofar as he remains physically able to choose not to enroll, despite being coerced from every direction in his life. Consenting to rehabilitation in the private facility is the same notion I am using to define consent for criminal rehabilitation. Of course, when the state coerces an offender, it often does so with the threat of punishment. I argue later in the chapter that the threat itself is an act of incapacitation, separate from rehabilitation. And only if it carries out the threat, then it is punishment.

There are two primary reasons that I include the consent requirement in the definition. First and foremost, without it, rehabilitation would include practices that clearly are not rehabilitative. Secondly, it provides very helpful conceptual clarity in distinguishing rehabilitation from incapacitation. I will explain that conceptual payoff in section 3.2 when I discuss incapacitation.

In order to show that the consent requirement is necessary, I will give examples of state interventions that meet all the definitional requirements of offender rehabilitation except for the

224 In many hospitals, if someone who has begun the donor process, but decides that he does not want to continue, the medical staff will provide him with a false medical excuse to make it easier for him to back out of the process. In other words, consent is treated so preciously that the medical staff will lie on behalf of the potential donor in order to save face for him and help preserve robust consent.

consent requirement, but which clearly are not rehabilitative. Zimmerman uses the example of aversion therapy from *A Clockwork Orange* to illustrate a strange category that he deems "non-educative" rehabilitation, but it seems clear that this should not qualify as rehabilitation at all.[225] In both the movie and book versions, the story features scenes of state agents forcing the main character, who is guilty of brutal crimes, to watch images associated with crime while he is induced with nausea. This is part of a regimen of making him physically ill every time he even thinks of a criminal act. This harsh aversion therapy qualifies as non-consensual even by the minimal sense of consent because the offender has no option to refuse the therapy. But besides being non-consensual, this practice meets all the other definitional requirements of offender rehabilitation: it is authorized, retributive, intended treatment. If the practice is not an instance of offender rehabilitation, then it is likely because of the difference in consent. My intuition is that this practice of clamping open a person's eyelids and inducing nausea is not rehabilitation, but even if the reader does not share my intuition, there are good reasons that this practice should not be considered rehabilitation.

The concept of rehabilitation has changed over time, but currently the debate regarding consent has shifted from whether rehabilitation should be consensual at all to what extent it should be consensual. Early penal reformers, who may be considered some of the first advocates of rehabilitation, advocated for changes that completely ignored the notion of consent and lead to unfortunate outcomes, including "extended periods of solitary confinement."[226] These early attempts at reform would not now be regarded as rehabilitative. However, the current debate centers on the question of what degree the state may coerce an offender to participate in rehabilitation. Underlying this debate is the implied premise that the offender must be able to choose otherwise, or must have minimal consent. Even the advocates arguing for more state coercion agree with this premise. They argue that "interventions offered to [offenders] are coercive *not* in the sense that individual offenders have no choice but to

225 Zimmerman, 2011, at 41. He doesn't specify whether he is referring to the book by Anthony Burgess or the film by Stanley Kubrick, but the scene occurs in both. This is a popular example in the literature. See e.g.
226 Robinson and Crow, 2009, at 18.

cooperate, but rather in the sense that there are likely to be negative consequences for non-participation."[227] Gone are the days when solitary confinement regardless of an offender's will can be considered rehabilitation, and in fact, US law claims that rehabilitation cannot be a justification for imprisonment.[228] Gone too is, hopefully, the notion that the aversion therapy in *Clockwork Orange* could be construed as rehabilitation.

Of course, the intervention from *Clockwork Orange* is an extreme example of non-consensual offender therapy. While most types of aversion therapy are physically restrictive, it is possible that they could qualify as rehabilitative if the offender consents to the treatment, and has a way to stop it. But the course of treatment in *A Clockwork Orange* is incapacitative because it physically restricts action without regard to the offender's will. The practice is the perfect example of a non-rehabilitative intervention belonging to a separate, non-consensual category; that category is offender incapacitation, as I will explain in section 3.2.

Both the *Clockwork Orange* intervention and solitary confinement-as-treatment are examples of state interventions that are authorized retributive treatment, but cannot be construed as rehabilitation because of the lack of minimal consent. To further argue for the necessity of the consent requirement, consider the example of a private drug rehabilitation program that involves a month of living in a facility under strict conditions. Patients are compelled to cut off contact with the outside world, and if a cell phone is found it is confiscated against the patient's will. The patient agrees to enter the program, and he has the option to leave, but a variety of actions during his treatment are coerced against his will. This is uncontroversially rehabilitative. However, if a disgruntled relative paid the program to kidnap the patient and force him to stay against his will, this would not be rehabilitative, but troublingly incapacitative, even if it is still a course of treatment. Restricting an offender without regard to his will is categorically different from his consensual participation in a course of treatment. Coercion in the

227 *Id*. at 12, their emphasis.
228 Harvard Law Review, "Recent Cases," 2011.

environment of social work and social service delivery is not unique to criminal justice. As Peters notes, "[c]oercive environments, after all, are endemic to many forms of social work intervention... Many social workers in other fields have very few clients submitting to services entirely voluntarily."[229] And Kleiman notes that, "Probation, fines, community service and drug diversion all have something in common: they involve… voluntary—albeit unwilling—compliance… The offender always has the option of ignoring the rules… That means that there needs to be an enforcement mechanism."[230]

One might object to the consent requirement of rehabilitation by claiming that it presupposes that the intervention is non-punitive. This objection would hold that by building consent into the definition of rehabilitation, it precludes the possibility that the state's action is punitive. But this objection is mistaken for two reasons. First, the question of whether a state action is punitive does not hinge on whether the offender consents to his sentence. If the state intends to harm an offender for his crime, then the harmful action is punitive, regardless of whether the offender consents to it or not. Therefore, the definition of rehabilitation does not preemptively shut the door on discussion of whether a particular course of rehabilitative treatment is actually the kind of intended 'hard treatment' that counts as punishment. Secondly, if the state gives an offender a sentence of mandatory rehabilitation, then there are two components to the mandate: the course of treatment, and the threat of enforcement. While a course of rehabilitation may not be punitive, the enforcement of mandated treatment includes the coercion of treatment over the offender's dissent, or the punishment of the offender for refused treatment. Note that the definition remains neutral on the question of whether offender rehabilitation is mandatory and if so how it could be enforced. Of course, these are centrally important questions, but they are separate from questions of the definition of offender rehabilitation.

I should note that professionals in the field of offender rehabilitation generally agree with the

229 Peters, 2006, at 356.
230 Kleiman, 2009, at 96-97.

consent requirement I have defended. Writing for the *Journal of Sexual Aggression*, Prescott and Levenson (who both work in sex offender therapy) write that,

> Treatment providers do not counsel clients who refuse services, as to do so would be unethical. Clients can choose whether or not to agree to the limits of confidentiality, and although there may be consequences for choosing not to participate in the programme, these sanctions are determined and enforced by the court, not the treatment provider.[231]

Another example is provided by the Correctional Service of Norway Staff Academy, which keeps correctional statistics of Denmark, Finland, Iceland, Norway and Sweden. In its definitions of terms, it explicitly states the need for the offender to consent before participating in treatment plans, community service or a community sentence.[232]

3.1.2 The Treatment Requirement of Rehabilitation

Using the word 'treatment' in the definition of rehabilitation has the downside of connoting the popular mid-20th century 'rehabilitative ideal' that strove for individualized plans for offenders to such an extent that it prompted a backlash from a diverse range of scholars and practitioners.[233] I will not rehash this well-trod territory, but will simply say that I mean to use 'treatment' in a way that remains neutral on such questions as to whether treatment can justify indeterminate sentencing, the extent to which the welfare state is obliged to fix societal ills that contribute to crime, and the extent to which treatment can 'work.' These issues are important, and I address them later while discussing justifications of rehabilitation, but for now I want to explicitly deny that my use of 'treatment' in the definition privileges particular perspectives on these issues. Furthermore, by including the consent requirement alongside 'treatment' in the definition, I hope to stave off worries that my use of 'treatment' implies that offenders are entirely passive recipients of medical attention that would 'cure' them given enough time and expertise. Clearly this 'disease analogy' is faulty for many reasons, and foremost

231 Prescott and Levenson, 2010, at 280.
232 Kristoffersen, Ragnar (ed.), 2010, at 7.
233 See e.g. Robinson and Crow, 2009, at 26-32.

among them is the fact that offenders need to take part in their own rehabilitation. It is a mistake to see rehabilitation as something *done to* offenders rather than something the state can *manage with* them.[234] Indeed, this worry was part of the motivation to include the consent requirement.

However, using 'treatment' in the definition has the benefit of capturing a wide variety of state interventions for offenders, and is thus a useful word for including the diverse practices that fall in the ambit of rehabilitation. 'Treatment' can point to practices that Duff characterized as "seeking to improve people's skills, capacities and opportunities."[235] Lynch elaborates, giving a helpful portrayal of rehabilitation as "any discourse or practices that speak to transforming or normalizing the criminal into a socially defined non-deviant citizen, including psychological programs, drug treatment programs, educational and work training programs, work and housing placement assistance, and half-way houses."[236] I use 'treatment' to capture these practices, and more specifically, to mean: a course of therapy, not limited to the field of medicine, designed to induce or help bring about behavioral and/or attitudinal changes in the recipient. Even with this explanation, the term 'therapy' remains vague, but this does not pose a problem. When scholars use 'treatment' and 'therapy' in the context of criminal rehabilitation, they mean to encompass the wide variety of practices used to spur changes in the offender. This is the difference between treatment and punishment: whereas treatment's primary goal is to change behavior and attitudes, punishment's primary goal is to harm the offender. Of course, the harm of punishment may in the long-term result in the rehabilitation of the offender. But this is a secondary goal, or justification of punishment, and should not be confused with treatment. Rehabilitation may foreseeably harm the offender, but the intention is to change his behavior while any resultant harm is a regrettable and unintended consequence.

Again, this raises the question of how we can determine the intention behind the treatment. This objection holds that because we cannot know the real intention behind a treatment, a course of

234 See McNeill, 2009 for an especially insightful account of this worry.
235 Duff, 2001, at 5.
236 Lynch, 2000, at 45.

treatment may look exactly like punishment but is labeled as 'rehabilitation' simply because someone reports that its primary intention is to change the behavior and attitudes of the offender. This objection holds that mental states are fickle and that we have no reliable way to trust whether reports of intentions are accurate. Again, as with the case of punishment, there is a way to determine the intention behind a course of rehabilitation, and it relies on public dialogue rather than internal mental states. Recall that in the case of punishment, the external location of the intention rests in the nexus of the dialogue between the legislature and the judiciary. The legislature expresses the intention, and the judiciary judges whether the practice actually corresponds with the expressed intention. In the case of rehabilitation, the intention resides both in the dialogue between legislature and judiciary, as well as in the nexus between codes of professional practice and professional review.

To illustrate this, I will take the example the recent discussion among sex offender therapists over whether the aim of community safety impedes the traditional principles of treatment like beneficence and autonomy.[237] This debate shows that professional codes and discourse can solve tricky problems about the aims of rehabilitation. It also illustrates that external dialogue (rather than hidden mental states) provides the intention behind treatment, as well as yielding criteria to determine the enforcement of borderline cases. This dialogue takes the form of professional codes of ethics and the deliberation over violations of the codes. Among the sex offender therapy community, Glaser has raised concerns over the conflicting aims of treatment. He has argued that because of the emphasis on changing the offender's behavior for the sake of community safety rather than the sake of benefiting the offender, current sex offender therapy "demands that confidentiality be broken, treatment be coercive and the offender undertake the therapy chosen by the clinician, whatever the cost to the offender."[238] Since this is a tension that arises specifically for experts in sex offender therapy, I will not presume to prescribe how to think about this issue. I merely want to point out that the dialogue demonstrates that

237 See Glaser, 2009; Glaser, 2010; Prescott and Levenson, 2010.
238 Glaser, 2010, at 264.

these experts can answer these questions. In response to Glaser, Prescott and Levenson point out three organizations that publish ethical codes and review cases of misconduct for the exact locus of tension that Glaser is concerned about, namely the tension between a therapist's duty to community safety and his duty to protect his client. These organizations are the Association for the Treatment of Sexual Abusers (ATSA), the American Psychological Association (APA) and the National Association of Social Workers (NASW).

> There is no question that some clinicians attend to the competing needs and other challenges of this work better than others, which is why organizations such as ATSA, APA and NASW enforce their ethical codes. In fact, meeting these seemingly competing demands is what most sexual offender treatment providers do every day.[239]

They note further that the professional codes explicitly acknowledge the tension, and they mandate certain professional practices to account for them. To take one example, the offender is fully informed of this tension before treatment begins, and what it may mean for him. He is able to take this into account before consenting to the treatment. The APA and NASW mandate that programs must "make clear at the outset of treatment that part of their intent is risk reduction, and that this is accomplished, in part, through limited confidentiality and collaboration with other professionals." After this is made clear, offenders may choose to then "sign an informed consent statement clarifying the boundaries of confidentiality and disclosure."[240] I have provided this brief glimpse into the dialogue between sex offender therapists and their professional codes to show that the intentions of treatment are not only explicit and codified, but the professional dialogue is able to handle the difficult problems arising from aiming for treatment and community safety.

Furthermore, this dialogue is also played out between legislature and judiciary. An example of this is a recent US court case that calls into question what properly counts as rehabilitation. In *United States v. Molignaro* the First Circuit Court decided that "rehabilitation could not be used to justify post-

239 Prescott and Levenson, 2010, at 277.
240 *Id*. at 279.

revocation reimprisonment."[241] This decision illustrates a capable and active exchange of laws and interpretations deciding whether the intention behind a particular practice (in this case using prison as a means of rehabilitation) can be justified. With this as evidence, I conclude that we should reject the objection that the true intention of treatment is opaque and that we cannot know whether any course of treatment has overstepped this intention.

3.2 The Changing Paradigms of Rehabilitation

For centuries people have focused on how to rehabilitate and reform offenders. Medieval clergy emphasized that the primary purpose of punishment was for the good of the offender.[242] The rise of the prison, especially in the USA, is intimately tied to notions of reforming offenders. While the central focus of Bentham's Panopticon, according to Foucault, was to make the prisoner constantly visible to authority and thus guarantee "the automatic functioning of power,"[243] the American prison experiment, starting in the early 19th century, was overtly rehabilitative. They were called 'penitentiaries' for this reason: to induce penance and reform.[244] The aim was to make offenders "transformed, gaining the moral fiber to reenter society and resist its temptations."[245]

Toward the end of the 19th century, in both the USA and Europe, failed rehabilitative measures and the rise of social Darwinist ideology gave rise to the idea of an incorrigible criminal underclass, born with predispositions to commit crime and therefore beyond the pale of reform. There is good evidence that Bram Stoker based the description of his character Dracula on pseudo-scientific physiognomical descriptions of common physical features of this criminal underclass.[246] In spite of this rise of thinking, in 1870 American prison professionals attending the National Congress on Penitentiary

241 Harvard Law Review, "Recent Cases," 2011, at 1115.
242 Allen, 2011, at 98.
243 Foucault, 1995, at 201.
244 Cullen and Smith, 2011??, at 160.
245 *Id*.
246 Bleyer, 2012.

and Reformatory Discipline doubled down on the goal of reforming prisoners by introducing new strategies which emphasized "expanded education and industrial training."[247]

The ideal of reforming offenders gained traction in the early 20th century as medical practitioners made advances in diagnosing and treating criminals. Individualized treatments of offenders aimed to reform the offender's condition not just via work and education, but also through "counseling, group therapy, behavior modification, token economies and therapeutic communities."[248] This paradigm culminated in the 1950s and 60s; in 1954 the "American Prison Association changed its name to the American Correctional Association" and prisons were renamed 'correctional facilities' or 'correctional institutions.'[249] In the field of philosophy this period saw the rise of theorists using the morally educative benefits to the offender as justification of punishment. To take an example that was applied to policy, the Criminal Code of Finland has the aim of "general prevention" but via the means of education: "the disapproval expressed in punishment is assumed to influence the values and moral views of individuals."[250] However, this confidence in rehabilitation was soon shattered by rising crime rates in the 1960s and 70s throughout the industrialized world.

During this time, conservative politicians called for tougher responses to crime, and saw rehabilitation as "welfare for criminals;" they pushed for "law and order," laying the groundwork for harsher sentences based on retribution.[251] In the USA this expanded punitive regime has lead to decades of accelerating growth of incarceration, resulting in the largest prison system in world history, far surpassing the number of prisoners in the Soviet gulag at its height.[252] In 1974 Robert Martinson wrote an article that has had uniquely influential sway over criminal justice policy, especially for the goal of rehabilitating offenders. He concluded that "rehabilitative efforts...so far have had no

247 Cullen and Smith, 2011??, at 161.
248 *Id*. at 162.
249 *Id*.
250 Lappi-Seppälä, 2011, at 240.
251 Cullen and Smith, 2011??, at 164.
252 Gopnik, 2012.

appreciable effect on recidivism."[253] His article was widely interpreted to mean that 'nothing works' to reform offenders, and its influence cannot be understated. Even now, when we understand that rehabilitation works significantly better than prison to reduce recidivism (as I will explain further in section 3.2), the 'nothing works' doctrine has significant influence. This is not the place to give a survey of the complicated developments of the past forty years, but I hope that I've shown that throughout history rehabilitation has been an integral goal of responding to crime, and that goal has been pursued by vastly fluctuating methods.

3.3 Justifications for Rehabilitation

In this section I will address the central questions of whether rehabilitation is justified and whether it can be part of a viable, minimally punitive regime. When the state intervenes to give an offender a course of rehabilitation, it usually does so within a punishment regime. Thus far I have defined rehabilitation as conceptually distinct from punishment, and have not discussed how it is usually implemented within a framework of punishment. Normally, offenders enter rehabilitation with the threat of punishment over them, and other times they are in prison when they receive the rehabilitation, thus being punished and rehabilitated at the same time. In current practice, punishment and rehabilitation are so entangled that it may strike readers as disingenuous or willfully ignorant to speak of justifications for criminal rehabilitation separate from the coercive threat of punishment. Sex offender therapy, drug addiction clinics, job and skills training: these are some of the most common forms of criminal rehabilitation and currently they usually occur within a framework of parole. Completion of these courses is the ticket to escaping the carceral net, while failure or refusal results in punitive parole revocation. Here I ask for the reader's patience. Just as pure restitution is a hypothetical alternative that can still be discussed and scrutinized, criminal rehabilitation separate from punishment is likewise worthy of attention. Before specifically looking at the relationship between punishment and

253 Martinson, 1974, at 25.

rehabilitation in section 3.3.3, I turn now to justifications of standalone offender rehabilitation.

There is one main justification for rehabilitation: the positive consequences of treatment. I will discuss the justifications of (1) the benefit to the offender, and (2) the benefit to society, and I consider objections to each. The benefit to society is the successful justification, so I will spend much more time discussing it. First I will briefly consider the benefit to the offender.

3.3.1The Benefit to the Offender Justification

The changes induced in the course of rehabilitation can drastically improve the lives of offenders. Some offenders who completed drug rehabilitation and turned their lives around remain profoundly grateful for the rest of their lives for the treatment they received. Other offenders may not be grateful, but because of the changes in their lives due to rehabilitation, they are able to live with a much higher quality of life. Perhaps the potential benefit to offenders is sufficient reason to give them rehabilitation. On reflection, this justification fails for two reasons.

First, it would justify rehabilitating the innocent. If the potential benefits of life changing rehabilitation were reason enough for the state to provide treatment, then the state would be obliged to provide treatment to those who have not broken the law. Presumably an argument could be made that a welfare state does have this obligation, but the obligation would not be part of a post-arrest regime, thus is not part of my discussion. Even if this argument were made, many people would not agree to the claim (or at least not want to back it up with tax dollars) that the state should provide rehabilitation simply because a person has a problem that would benefit from rehabilitation. This claim follows from the Benefit to the Offender Justification, which provides reason enough to reject it.

Secondly, the Benefit to the Offender Justification fails to justify some treatments that uncontroversially qualify as criminal rehabilitation. Consider a happy drug addict who is completely contented in his lifestyle, which includes burglarizing homes to fund his addiction. The state's obvious

rehabilitation option is to give this offender drug addiction therapy. But the benefit to the offender objection cannot justify this option because the offender is content without rehabilitation; in fact he would strongly prefer not to have it. He would reject the state's offer, and would resent the state for pressuring him to enroll. One reply to this objection would be to claim that the state can determine what would be best for the offender, and decide to offer or pressure him into the rehabilitation because the state determines that it would benefit him. But this reply fails because it requires that the state take a problematically patronizing position. Why would the state be allowed to decide what is best for an individual? The happy drug addict surely maintains the ability to decide what he wants for himself. A theorist may respond by claiming that the state is not deciding individual cases, but taking a public health perspective, and issuing a policy of drug rehabilitation for all addicted offenders. This would be similar to preventative health measures such as smoking bans. But by pressuring citizens into rehabilitation, the state is choosing a much more invasive course of action than by merely nudging citizens away from certain health decisions such as smoking. We would feel that if the state had a policy to pressure all citizens it deemed eligible into rehabilitation regardless of whether they had committed crimes, then the state would be acting with unwarranted patronization toward the non-offenders. But the justification cannot distinguish between offenders and non-offenders. To reiterate the first objection: if the justification merits rehabilitating offenders solely on the basis of the benefit to them, then it also merits rehabilitating non-offenders who stand to benefit similarly.

3.3.2 The Benefit to Society Justification

Alternatively, one could point to the prevention of future crimes, and reduced victimhood, as the justification of rehabilitation. This justification rests on the empirical claim that rehabilitation reduces crime. Studies have shown that rehabilitation is indeed successful at reducing rates of recidivism, which means preventing offenders from reoffending. A more accurate way to phrase this,

one that better accounts for offender autonomy, is to say that rehabilitation helps offenders to desist from crime, but I will assume that these two phrasings of 'preventing crime' and 'helping desistance' mean the same for this discussion. In 2004 Peter Raynor reported that intervention and treatment has been shown to produce large reductions in reconviction rates, from 9-21 per cent depending on the study.[254] In a 2007 meta-analysis, Lipsey et al. report that "[t]he average effect on recidivism found for the best forms of intervention in the research syntheses ... was in the range of a 10-25 percentage point differential between treatment and control groups."[255]

There are many different studies exploring the effectiveness of rehabilitation; they have various shortcomings, but the consensus is that treatment is better than prison alone. Michael Tonry writes that, "A vast new literature on correctional treatment asserts that many kinds of programs—sex offender treatment, anger management, cognitive skills training, vocational training, and drug abuse treatment—can reduce reoffending."[256] There is a dearth of knowledge regarding exactly which kinds of treatment work best for specific offenders,[257] but the past decade has seen a fervent and growing group of researchers working to tackle this question.[258] Often the question is reframed to be what intervention best helps offenders desist from reoffending, because the reality is that no single course of treatment can possibly 'work' in the sense of reliably reducing recidivism for every offender.[259] Recidivism is a notoriously tricky problem, but treatment does help offenders desist from re-offending. While there are dozens of studies to cite that show this conclusion, I will quote Glaser who is a critic of aspects of current practice of sex offender therapy. While he has qualms with the treatment delivery, he openly admits that research shows "sex offender treatment programmes have enjoyed considerable success, resulting in significant reductions in offender recidivism, the suffering experienced by victims and their

254 Raynor, 2004, at 199-200.
255 Lipsey et al., 2007, at 220.
256 Tonry, 2011a, at 9.
257 *Id*.
258 See Hollin and Palmer, 2009 for a detailed description of dozens of studies. Ward and Maruna, 2007, give an excellent book length evaluation of this research.
259 Ward and Maruna, 2007, at 12.

families and the costs of victim support and offender processing in the criminal justice system."[260] While this is true, the story is also still fairly gloomy: certain well run rehabilitation programs may reduce recidivism some 10-20%, but some offenses typically see rates of re-offense as high or higher than 70%. Pedophilia and burglary are two examples of crimes whose perpetrators have disconcertingly high re-offense rates. Still, rehabilitation is the most effective course of action we know of to reduce recidivism. Furthermore, given the research we have, we know that there are trends that identify better courses of rehabilitation. Generally, strength and intensity of implementation are more important than the specific differences between types of rehabilitation. Most rehabilitative interventions are more effective than prison, except for fear or punitive-based approaches. The most effective programs aim to change the "criminal behavior directly" or its "specific proximal causes" but without "fear-based or punitive approaches to do this (e.g., boot camps, intensive supervision)."[261] The more effective programs include "structured regimens, such as training, as their primary component" which are more successful than "less structured interventions based on building relationships, such as counseling, mentoring, and restorative programs."[262] Finally, "[m]ultimodal programs are generally more effective than programs that employ a single treatment strategy" and, unsurprisingly, programs that are competently administered with intensive interaction succeed best.[263] By 'intensive,' Lipsey et al., mean "about 25 weeks' program duration with 5-10 contact hours of treatment per week delivered in multiple sessions."[264]

A dominant rehabilitation structure has emerged from Canadian researchers, known as the Risk-Need-Responsivity (RNR) Model of rehabilitation.[265] To quote a summary of the three principles of this model:

260 Glaser, 2010, at 262.
261 Lipsey et al., 2007, at 219.
262 *Id*.
263 *Id*. Interestingly, the authors note that "in particular, adjunct counseling, academic and employment training, and supervision appear to be more effective in multimodal combinations than as freestanding treatments."
264 *Id*. at 220.
265 Ward and Maruna, 2007, at x.

> [F]irst the risk principle whereby services are directed towards medium to high risk offenders; second, the needs principle so that interventions target offenders' criminogenic needs; third, the responsivity principle by which interventions are matched to offenders' learning styles.[266]

However, in practice, these principles are often not yet implemented. Current common correctional practices leave much room for improvement to meet the criteria of successful rehabilitation techniques. A review of 374 correctional programs concluded that 61% "failed to reach even a basic level of adherence to the RNR principles."[267] This goes to show that policy and on-the-ground practice lag behind research.

To summarize the thrust of these findings, research confirms the empirical claim that rehabilitation can reduce recidivism. The research behind this claim supports the Benefit to Society Justification by demonstrating its feasibility. It also demonstrates that the gains are substantial, with a much higher success rate, if properly implemented, than current penal practices. Another reason for optimism is that we continue to progress in our ability to prompt positive change: in 2005 Raynor and Robinson reported that, "We know more about how to help offenders… Twenty years ago, hardly anyone would have predicted that such achievements were imminent."[268] Of course, even with high quality rehabilitation programs in place, rates of recidivism would still be higher than we would like. Also, there are some offenders who will never benefit from rehabilitation. While practitioners are able to identify many of them, it would be an intolerable public hazard to rely solely on rehabilitation for these types of offenders. For instance, if an offender with psychopathy is a high-level risk, then it may be the case that the state is justified in isolating him from society.

One objection to the Benefit to Society Justification is to claim that it would justify rehabilitating the innocent. This objection was successful against the Benefit to the Offender Justification because if the potential benefit to the offender were the reason to offer rehabilitation, then

266 Hollin and Palmer, 2009, at 147.
267 Day et al., 2011, at 75 quoting Morgan et al., 2007.
268 Raynor and Robinson, 2005, at 172.

it would follow that the state should offer rehabilitation that it normally reserves for its offenders to anybody who might likewise benefit. However, this objection fails in the case of the Benefit to Society Justification. The reason for this is that society stands to benefit from reduced crime when rehabilitating offenders, but not when rehabilitating non-offenders. There are no cases, to my knowledge, in which this justification would sanction rehabilitating innocent people. One possible candidate for such rehabilitation is an alcoholic who is veering toward breaking the law but has not yet. His behavior pattern suggests that he has a high likelihood of breaking the law soon, and thus from the interest of preventing his potential crime, the state would be justified in giving him rehabilitation. However, on closer inspection the example fails. How would the state know that the alcoholic is about to commit a crime? If he were alone in his apartment, about to be kicked out for not paying the rent, then the state could not know that he was headed for a life of vagrant street crime unless the state took problematically intrusive measures to monitor his behavior. If he were already out on the street, then the best way that the state could know that he is headed for crime is if he already committed a crime associated with homeless alcoholics: urinating in public, public drunkenness, disrupting the peace etc. Consider if someone makes a criminal threat: while this is a good indicator that the person will commit a serious offense, the threat itself is an offense. The point is that the only reliable and unproblematic way of determining that someone is an imminent risk of committing crime is if they have already been caught committing one.

A similar objection claims that the Benefit to Society Justification would allow for indeterminate and disproportionate sentences. Take the example of an offender whose conviction of assault results in a course of anger management classes. If his anger problems persisted past the usual length of treatment, then the state may be justified, using the Benefit to Society, in continuing to coerce him into more treatment. This cycle could presumably last for years, resulting in a situation that most would deem disproportionately harmful to the offender in relation to his crime. The reality is that this scenario would not occur. In a rehabilitative post-arrest regime, the state would not be able to

individually tailor every sentence to the exact length necessary so that the offender's risk to the community would be neutralized. As is the case in the current punitive regime, there would be sentencing guidelines based on expert analysis. A guideline might claim that a certain kind of assailant with a diagnosis of a particular anger problem should receive a treatment regimen within a certain range of intensity and duration that would seek to maximize the benefit to society given the resources available and the demonstrated risk. It would not make sense to keep a low level offender coming back for more treatment because of a simple assault. Of course, more individualized treatment and attention would be necessary for the offenders who present a more serious risk. It may be the case that a murderer never quite escapes the state's net of treatment. His treatment providers and risk assessment experts would make recommendations that would need to be reviewed by an external body, perhaps the courts. Abuse of risk assessment has the potential to unfairly intrude into offenders' lives. But there could be systems set in place to check this kind of abuse. There is nothing inherent to the Benefit to Society Justification that would automatically lead to sentences deemed to be disproportionate.

Another objection to the Benefit to Society Justification is the claim that it could not merit giving rehabilitation to certain offenders who need it. Consider the example of a very old offender who is deeply addicted to drugs. The state has determined that the cost of rehabilitating him is not worth the potential benefit because he is unlikely to commit further crimes, in part because of his placidity and also because he is likely to die soon. It would seem callous and discriminatory to deny him the programs which are available to similar offenders who only differ because they are younger, healthier and thus at higher risk of reoffending. But a successful reply to this objection is to simply bite the bullet. If the reason that the state gives rehabilitation to offenders and not to non-offenders is in order to prevent future crime, then we must accept the fact that there will be cases in which offenders who could benefit from rehabilitation do not qualify for treatment because they do not pose a danger of reoffending. To give another example that brings this harsh fact into focus, consider a case of an offender whose alcoholism contributes to petty theft, which the state is justified in trying to curb with

alcohol rehabilitation. But consider if he also has other issues that would warrant treatment but have not contributed to his criminal behavior. For instance, if he is also has anger management issues that have not resulted in additional crime, then the state would be justified in not treating them, because the state is not obliged to treat everyone's anger management issues. Of course, some advocates of the welfare state may argue that the state does have this obligation, but I will not be exploring that argument. Even in states like the USA without universal health coverage there exist treatment options outside of the state's purview for the old man who wants rehabilitation and the alcoholic who has anger issues. While these options in the USA are sadly inadequate, the discussion of healthcare obligations in a welfare state does not belong here. My assertion is that the Benefit to Society Justification will result in cases of offenders who do not receive treatment that they could benefit from, but that this does not provide reason enough to reject the justification.

However, before moving forward, there is one other notable caveat to this biting the bullet response. Suppose that in the course of treating an offender, the state finds that he has a condition that is unrelated to the criminal behavior but is dangerous enough to merit treatment. For instance instead of anger issues, suppose the alcoholic from the example above demonstrated a pedophiliac orientation. Or suppose that he has bipolar disorder that results in homicidal ideation. It seems clear to me that the state should be able to provide treatment to these offenders. But the treatment is not justified based on his past criminal behavior, but rather potential future crime. The state should offer all of its citizens these kinds of treatment, regardless if they have committed a crime yet. This is not a refutation of the Benefit to Society Justification, but a wider claim about what the state owes to its citizens. Here I am making a claim about the welfare state, but this is because unlike other treatments, there is clear and compelling evidence that suggests that every state should provide these treatments to all of its citizens. This evidence makes universal provision of these treatments less debatable than, for instance, anger management. I do not pretend to know which kinds of disorders warrant this coverage. But it seems clear that there are some extreme examples that stand out. I would leave drawing the cutoff line to

experts, but let me just briefly focus on pedophilia to make the point that this category of disorders warranting universal treatment does indeed exist. If the treatment is a response to a crime, then it is rehabilitative, but if it is not, then it is preemptive, and does not qualify as offender rehabilitation. The case I mentioned above, where an offender is convicted of an alcohol related crime, but is found in the process to be pedophiliac, and is then treated for his pedophilia, would not qualify as offender rehabilitation. Below I make the case that he should be treated regardless.

Pedophilia is a sexual orientation, which, like other sexual orientations, cannot be changed or 'cured.' Recent developments in brain scan technology have allowed researchers to accurately give a clinical diagnosis of pedophilia by looking at a patient's brain during sexual stimuli. As Ponseti et al. conclude, "Functional brain response patterns to sexual stimuli contain sufficient information to identify pedophiles with high accuracy. The automatic classification of these patterns is a promising objective tool to clinically diagnose pedophilia."[269] Treating pedophiles before they offend has become an increasingly high priority. The treatment involves support groups and therapy to help cope with not being able to fulfill their sexual urges. This will have the long term benefit of preventing pedophiles from acting on their impulses, or as Bleyer puts it: "one way to protect kids may be to reach pedophiles pre-emptively, to give them the therapeutic tools to control themselves and still lead fulfilling lives."[270] Bleyer goes on to report on a successful instance of this kind of program.

> "Fred Berlin, the founder of the Sexual Disorders Clinic at Johns Hopkins University... has had success treating pedophiles with therapies similar to those for drug addicts, with an emphasis on taking responsibility for one's actions, identifying triggers, and resisting cravings, as well as developing empathy for potential victims and addressing cognitive distortions that may support unhealthy behavior. His patients have also had successful outcomes with testosterone-lowering medication, otherwise known as chemical castration."

In order to protect children from pedophiles, we should be able to encourage pedophiles to acknowledge their orientation in order to seek help and support. For the sake of potential child victims

[269] Ponseti et al., 2011.
270 Bleyer, 2012.

and the pedophiles struggling against their nature by themselves, these treatments should be available to all pedophiles, not just ones who have committed offenses. Thus, I hope to have shown that there is a category of treatment, one that does not necessarily fit within the definition of offender rehabilitation, that should be available to all citizens, not just offenders, because the risk of offending is so costly. Notice that the justification for such preventative treatment is not the benefit to the offender, but the benefit to others, which mirrors the justification structure of post-conviction treatment.

To refocus on the Benefit to Society Justification, a further objection states that it is perverse to offer offenders beneficial treatment that is not available to non-offenders solely because the offender has committed a crime. In the literature this is known as the problem of 'less eligibility' because, according to Bentham, offenders should not be deemed "more eligible" than their innocent citizen counterparts.[271] This objection holds that such a distribution would create warped incentives for people who desperately want treatment but have not been convicted of a crime (due to not having committed one or not being caught yet) to commit crime in order to be convicted and given treatment. Again, I believe this comes down to questions of what a state is obliged to provide to its citizens. I agree, as I detailed above, that all pedophiles should be eligible to receive state treatment, whether or not they have offended. The reasoning I use follows a 'benefit to others' justification, and indeed if there is a clear benefit to society in addition to the treated individual, it is likelier chance that one could show that the state is obliged to provide the treatment. There is also good evidence that the most effective way for a state to manage its drug addicts is to treat the problem as a public health issue rather than focusing on punitive measures for the addicts. If a citizen is so desperate for treatment that he is willing to commit crime in order to find help, then it is likely that the society is failing in many ways that do not have to do with the criminal justice system. A state should be able to provide for the general well-being of its citizens to an extent that crime does not provide a temptation in order to receive treatment.

271 Bentham, 1791, quoted in Robinson and Crow, 2009, at 13.

3.3.3 Rehabilitation and Punishment

In most instances of rehabilitation, the state has some hand in coercing the participation of the offender. While the offender can choose to participate or not, the state attempts to compel the offender to choose to complete the rehabilitation. But how exactly does the state pressure offenders into rehabilitation? The answer is usually via punishment. The two main questions of this section are (1) whether rehabilitation can be coerced without threat of punishment, and (2) whether rehabilitation which has been coerced using the threat of punishment is in fact punitive.

Lewis, somewhat sensationally, sums up the compulsory nature of rehabilitation: "If a tendency to steal can be cured by psychotherapy, the thief will no doubt be forced to undergo the treatment. Otherwise society cannot continue."[272] Are there any alternatives to punishment for coercing offenders to participate in rehabilitation? One possibility is to threaten offenders with preventative detention in order to motivate their participation. But on closer examination, this fails. The best-case scenario would be that an offender is in preventive detention because he is a danger to the community and needs to go through rehabilitation before any possibility of being released back into the community. The state could then stipulate to the offender that he will likely remain in preventive detention until he decreases his risk to the community through participation in treatment. This course, I will later in this chapter argue, is non-punitive. However, this would certainly not be an option for many offenders, as preventive detention is only justified for very dangerous offenders. It seems that, as was the case with restitution, the way to enforce rehabilitation is with the threat of punishment. And carrying out the punishment if the offender refused to comply would be necessary in order to ensure that other offenders took the threat seriously.

Secondly, if an offender participates in a course of rehabilitation because he has been coerced by the threat of punishment, does that make the rehabilitation itself punitive? I argue that it does not. The rehabilitation itself is still a practice aimed at changing the behavior and attitudes of the offender,

272 Lewis, 2011, at 92.

not at harming him. The fact that the offender participated because of a threat of punishment does not change this aim. It would be strange to claim that by employing the threat the purpose of the threatened punishment would somehow be permanently transposed onto the rehabilitation such that the aim of changing the offender's behavior would be replaced with the aim of harming him. In this context, there are three discrete actions available to the state: to punish, to threaten, and to rehabilitate. Each of these has its own discrete aim. The threat of course is aimed at coercing the offender to participate by making his refusal illegal; I will argue later that such a threat is a legal restriction of behavior, and thus qualifies as offender incapacitation. It does not make sense to assume that these purposes can transfer from one to the other.

One might reply that the threat of punishment remains throughout the offender's rehabilitation, and this changes the nature of the rehabilitation. Certainly most, but not all, cases of coerced rehabilitation carry the continued threat of punishment if the offender refuses or fails in his participation. So how would the continued threat of punishment affect the purpose of the rehabilitation? It seems that the two remain discrete, despite occurring simultaneously. If at any time the offender were punished for refusing to participate, then that would be the moment that the punishment would begin. It does not make sense to claim that it happened earlier. Coercion in the environment of social work and social service delivery is not unique to criminal justice. Recall Peters' claim that I quoted above: "Many social workers in other fields have very few clients submitting to services entirely voluntarily."[273]

Of course, punishment and rehabilitation may happen simultaneously. It is very common to find drug treatment programs and mental health treatment within prisons. There is good reason to see the co-occurrence of rehabilitation and punishment as distinct actions happening together, albeit dysfunctionally. Research shows that punishment is often inimical to rehabilitation when they are

273 Peters, 2006, at 356.

administered jointly.[274] This suggests that the goals of the two do not change, even when their courses coincide; in fact the goals of harming someone and treating them remain in stark contrast.

A Danish study that investigated voluntary drug treatment programs held within prisons concluded that the negative effects of punishment were the single major obstacle for rehabilitation. Instead of being able to focus on changing the offenders' future behavior out in the community, "the counsellors' main goal in providing prison-based drug treatment appeared to be the alleviation of the negative consequences of imprisonment."[275] In the words of Robinson and Crow, rehabilitation under this view is "conceptually divorced from punishment, such that it is not understood as an objective or quality of a positive process of punishment, but rather as an *antidote* to punishment: or, more precisely, the potential harmful effects of punishment."[276]

Punishment and rehabilitation are discrete even when they co-occur or the threat of punishment overlays the rehabilitation; one further piece of evidence for this comes from the US law. The Code of the Laws of the US states bluntly that "imprisonment is not an appropriate means of promoting correction and rehabilitation."[277] Furthermore, in 2011, the "First Circuit held that even upon revocation of supervised release, courts may not impose imprisonment with the aim of facilitating rehabilitation."[278] This means that when an offender is on parole, the state may not put him back in prison as a means of rehabilitating him. This suggests that the First Circuit Court agrees with my reasoning that punishment and rehabilitation are discrete. The threat of punishment, which is ever present for someone on parole, does not mean that it changes the non-punitive aim of any rehabilitation program he is completing.

As I have argued, compelling offenders to participate in rehabilitation is indeed a necessary intervention that the state's crime response regime will need to have available. But the current punitive

274 Kolind et al., 2010, at 46.
275 *Id*.
276 Robinson and Crow, 2009, at 7-8.
277 US Code, 18 U.S.C. § 3582(a) (2006).
278 Harvard Law Review, "Recent Cases," 2011, at 1112.

focus often obfuscates the goals of treatment. As I argue elsewhere, the state has an obligation to minimize the punitive aspect of the intervention as much as possible. If punishment is justified as a last resort of coercion, then it should be used as a last resort rather than the primary focus. Sex-offender therapists have to constantly grapple with this balance. But as Prescott and Levenson note, research has shown that offenders' "efforts at change often begin only with outside pressure while the client develops his or her own internal motivation." But they immediately warn that, "[t]reatment methods that are purely coercive rarely produce meaningful change. This is precisely why methods that respect the inherent autonomy of each client (e.g. motivational interviewing) have gained currency in recent years."[279] In conclusion, punishment is necessary as a final recourse for a post-arrest regime that relies on rehabilitation. But coercing an offender to participate in rehabilitation does not in itself count as punishment. Furthermore, punishment should be limited as a last resort rather than the predominant priority.

3.4 A Definition of Offender Incapacitation

In this section I will defend a definition for non-punitive offender incapacitation as well as give a brief overview of the kinds of practices that qualify as incapacitative under my definition.

As I discussed in the first chapter, when the state quarantines somebody to prevent the outbreak of disease, it does not intend to harm him. When the state restricts an offender without intending to harm him, the restriction is not punitive, a claim that I will further support in this section. Also in this section, I defend three conceptually distinct definitions for punishment, offender rehabilitation and *non-punitive* offender incapacitation. Under these definitions, no single act may qualify as more than one of these interventions. No single act may be both punitive and rehabilitative or incapacitative; no single act may be both rehabilitative and punitive or incapacitative etc. I have previously defended the definitions of punishment and rehabilitation on the merits of their respective requirements, and I aim

279 Prescott and Levenson, 2010, at 278.

here to do the same for non-punitive incapacitation. Additionally, the division of these three concepts yields benefit in organizing and clarifying the intentions and justifications of each discrete state action, whereas current dialogue is often plagued with vague or imprecise use of these three terms.

To reiterate and clarify, I am claiming that there is both punitive and non-punitive incapacitation. There can, of course, be punitive incapacitation, which is what most imprisonment is. But I will argue that non-punitive offender incapacitation, as a standalone concept, is worthy of scrutiny separate from punishment. And while it is definitionally non-punitive, its use as an alternative to punishment merits the same kinds of questions that arose for restitution and rehabilitation. As I showed in the previous discussion, state interventions are often composed of more than a single discrete action. A court sentence may contain separate parts, each of which on its own belongs to different categories. Each of the separate categories, including non-punitive incapacitation, deserve scrutiny just as the composite sentences do.

Before parsing what makes non-punitive offender incapacitation unique, which I will refer to hereafter as 'incapacitation,' a good start to defining it is to ask what it has in common with punishment and rehabilitation. All three share the qualities of being authorized, retributive and intended. These three, shared conditions confirm that all three are post-arrest state interventions. So what makes incapacitation unique? As I alluded to while discussing the definition of rehabilitation, incapacitation differs because it is enforced without regard to the offender's will. This does not mean that it is necessarily enforced against the offender's will, although it may be. It may also be enforced without the offender's foreknowledge, or even with the offender's consent—the point is that none of these references to the offender's knowledge or preference matter because the action would be imposed regardless. Thus for an action to qualify as incapacitative, the offender's will must not play any role, which is how I am using the term 'non-consensual.' Notice that this is also true for punishment, but it is unnecessary to name in its definition because intentionally harming someone implies non-consent.

Secondly, non-punitive incapacitation restricts the offender without intending to harm him. The

restriction requirement does not necessarily mean physical restriction. The state may also legally restrict an offender from doing something as a result of his breaking the law. The obvious purpose of restriction is to prevent future offenses, but in order to avoid smuggling a justification into the definition, the term 'restriction' is limited to meaning a physical or legal restriction of the offender's behavior. Thus we arrive at the definition: incapacitation is **authorized, retributive, non-consensual, non-punitive, intended restriction.** Below I attempt to elaborate this definition and defend it against objections.

First, one might object by claiming that the non-punitive caveat is ad-hoc. I previously defended my reasoning behind keeping rehabilitation definitionally distinct from punishment, so I will be brief here. Incapacitation fills the conceptual void to capture the many non-punitive interventions that are non-consensual and restrictive. I will elaborate on the range of practices that fall into this category below, but here are some incapacitative examples: preventive detention, non-consensual chemical castration, and electronic monitoring. Simply because the categories of punishment and incapacitation are conceptually distinct does not privilege any particular justification for either. Non-punitive incapacitation by itself needs to be justified. Its role in sentencing, especially in relation to punishment also needs scrutiny. An advocate of non-punitive incapacitation as an alternative to punishment needs to explain how the practice can remain non-punitive, which is a separate discussion from the definition. By making incapacitation definitionally non-punitive, I am keeping the two conceptually distinct for the purpose of discussing a particular range of interventions. Surely many criminal restrictions of liberty are punitive; but there is a group of interventions that are both non-punitive and incapacitative, and defining incapacitation this way allows for the useful discussion of inquiring whether this group is justified.

A second objection to this definition is to claim that 'non-consensual restriction' is redundant, and that the only reason to include 'non-consensual' is to keep it conceptually distinct from rehabilitation. But this objection fails because there are consensual criminal restrictions. The distinction

is useful because it can account for substantive differences in interventions that otherwise might be subsumed under an inaccurate label. Take for example estrogen hormone treatment for male sex-offenders that is known as 'chemical castration.' This treatment drastically reduces male sex-drive; it is a reversible administration of anti-androgen drugs, which reduce male hormone uptake. It is designed to reduce male libido and sexual activity. The American Civil Liberties Union has argued that such treatment is cruel and unusual punishment because if it is mandated then it violates an offender's right to reproduce and to refuse treatment.[280] While I will not be wading into a debate over whether this treatment is constitutional, the debate demonstrates why 'non-consensual restriction' is not redundant, as I explain below. It also gives credence to the separation of incapacitation from punishment.

Chemical castration, if it is consented to, is rehabilitation. However, there are instances when chemical castration is court-ordered. In such cases the state makes it illegal for offenders to refuse treatment. The legal restriction is itself a form of incapacitation, but this is separate from the course of treatment. Unless the state administers the treatment regardless of the will of the offender, which would be a case of physical incapacitation, the course of treatment is rehabilitation because the offender may choose to refuse. But it is accompanied by the legal incapacitation of the state's threat to punish the offender for refusing.

Chemical Castration

Administered regardless of offender's will	= Physical Incapacitation	
Administered with threat of punishment	= Rehabilitation	+ Legal Incapacitation
By request of offender	= Rehabilitation	

In contrast, offenders may request the drug. The ACLU cites the case of a Texas inmate, Larry Don McQuay, who for years "begged the state to [chemically] castrate him."[281] There have been other high-profile instances of offenders requesting to receive the treatment. The BBC recently reported that

280 Spalding, 1998.
281 *Id*.

Christopher Hughes, a serial rapist and pedophile, "has asked to be chemically castrated to stop him reoffending."[282] This demonstrates that restrictive treatment can be consensual, and hence it does not qualify for offender incapacitation as I have defined it because of the non-consensual requirement. Another high profile example is the Whatton prison in Nottinghamshire where over 100 sex offenders requested to be chemically castrated. The Guardian reports that the program is "entirely voluntary," and that Prof. Don Grubin, who runs the program, says of the offenders referred to him, that:

> A small number have been referred for anti-androgen treatment which leaves men with pre-pubescent testosterone levels similar to those in men who have been physically castrated. This so-called chemical castration is reversible and therefore some people argue that the phrase "castration", which suggests permanence, is not accurate. [283]

The Telegraph further quotes Prof. Grubin saying that "We know the treatment works to reduce sexual arousal and fantasies."[284] This shows that chemical castration can be a successful form of restrictive, voluntary rehabilitation. Chemical castration provides a useful example of how a single intervention, depending on how it is administered, can be either rehabilitation, or purely incapacitation. I have not come across any instances of the treatment being forced on the offender regardless of his will, which would be the only way that it qualifies as incapacitation, but it is possible. All of the accounts that I have found explicitly intend the treatment as a preventive measure against future offenses, which disqualifies it as a punitive measure. The ACLU cites a California legislator who "talked of prevention, not punishment. 'We are simply trying to stop sexual offenders from committing additional crimes,' he said."[285] Even if the offender is mandated to take the treatment under the threat of punishment, but can refuse it, then it qualifies as rehabilitation. Of course, as I discussed above, mandatory rehabilitation cannot be considered a 'pure' alternative to punishment because it will often involve recourse to punishment if the offender refuses treatment. The ACLU documents a Florida law mandating chemical castration for sex offenders with the threat of being punished for refusing it: "If the defendant fails to

282 BBC, September 2012.
283 Curtis, 2012.
284 Holehouse, 2012.
285 Spalding, 1997.

submit to or refuse to appear for treatment, the defendant is guilty of an additional second degree felony."[286]

A third objection challenges the scope of the definition, specifically pertaining to what I have deemed 'legal incapacitation.' This objection holds that this practice is better described simply as a 'prohibition,' and should not be included within the range of offender incapacitation. It points out that legal restrictions may be broken, and are thus similar to prohibitions, whereas physical incapacitation (usually) cannot be transgressed because of the state's closer grasp. However, this objection fails. There is no reason to limit the concept of incapacitation simply to physical restriction; it may also be a legal restriction. Both physical and legal restrictions may be broken (think of jailbreaks), and both may be tightened according to state control. It is true that most legal restrictions are simply prohibitions created by laws. However, there is a specific kind of legal restriction that is a result of an offender's past criminal behavior. This is the kind that falls under the definition of offender incapacitation, specifically fulfilling the retributive requirement. These restrictions are also administered regardless of the offender's will, fulfilling the non-consensual requirement. As a means of further clarification, below I attempt to describe practices that I take to be straightforwardly incapacitative given the definition I have defended. This of course is not an exhaustive listing, but rather brief evidence that the ambit of this category is wide and varied. Above I explained that chemical castration, if administered without any regard to the offender's will, is incapacitative. However, this is not a good representative of the category, because non-consensual chemical castration seems rare and perhaps non-existent, and furthermore it is subject to serious bioethical objections that I do not want to address. There are much more common and less controversial examples of incapacitation, in both the physical and legal varieties.

One example is rescinding an offender's driver's license after his driving related offense. This qualifies as incapacitation because the revocation of the license makes it illegal for the offender to

286 *Id*.

drive, therefore legally restricting his behavior. The goal of the restriction is to prevent the offender from further driving related offenses, and it is a direct result of his past offenses, meeting the retributive requirement. If this restriction is the true motivation for the state's action, then it is non-punitive because the aim is preventive rather than to harm the offender. It also (typically) happens without any regard to the offender's will. Similar examples of incapacitation via legal restriction are: house arrest with electronic monitoring which makes it illegal to be certain places at certain times; restraining orders, which make it illegal for someone to come within a certain distance of another person; restrictions on sex offenders in the community which make it illegal to live or work within a certain distance of places where children gather (an infamous example of this is in Miami where recent laws forced sex offenders to camp underneath a causeway on an island in a bay because it was the only place they could live that fit the restrictions);[287] the multifarious laws governing what offenders on parole can and cannot do; and the disenfranchisement of felons in the USA, which takes away offenders' ability to vote. Before moving forward, I should emphasize that the act of making these activities illegal is incapacitative, but the follow-up action if an offender violates any of these restrictions is most likely punitive. The threat of punishment is a restriction, but not itself punishment. I also should acknowledge that it is dubious that some of the interventions I listed above can actually be justified on grounds of prevention. For instance, the restriction of voting seems like a straightforwardly punitive action without any viable justification for public safety. And while others do seem to be able to be justified based on prevention, their relationship with punitive consequences needs to be scrutinized.

A more extreme and controversial example of incapacitation is preventive detention. This is the practice of detaining somebody "not for crimes they have committed, but rather to prevent future wrongdoing."[288] This definition raises the question of whether preventive detention is retributive, i.e. in response to an offense, which is a requirement for offender incapacitation as I have defined it.

287 See e.g. Skipp and Campo-Flores, 2009. A brilliant fictionalization of this Miami encampment of sex offenders is the novel *Lost Memory of Skin*, by Russell Banks.
288 DeVore, 2011, at 727.

Typically, people who are preventively detained are past offenders who are deemed dangerous enough to be detained to prevent future offenses. There is no just or reasonable way to ascertain whether someone is sufficiently dangerous to warrant preventive detention without evidence of past offenses. As DeVore points out, the three groups of offenders who are currently most likely to be subject to preventive detention in the USA are: "suspected terrorists, sex offenders, and criminals with severe mental illness."[289] There is a robust debate over the legality and ethics of holding suspected terrorists, and I will not take a side. It is debatable whether they have committed a crime that signals the need for preventive detention. But in the cases of dangerous mentally ill offenders (e.g. psychopathic killers) and sex offenders, it is clear that the detention, while aiming to prevent future offenses, is also in response to a past offense. DeVore notes that the practice is arguably unconstitutional and unjustified. However, he also notes a primary reason for its implementation: "people in the three categories listed above are commonly thought to be undeterrable. If this is true, preventive detention may be the only way to avoid the crimes they would otherwise commit."[290] Another way to read 'undeterrable,' especially in light of the problems with aiming for deterrence (which I describe in Chapter 5), is 'immune to rehabilitation.' Allen and Laudan give a more precise definition of preventive detention: "incarceration of persons not as punishment for harms already wrought but as a device for preventing future harms that the state surmises they are likely to inflict if not incarcerated."[291] They also note the controversy over whether certain acts, such as those that signal a 'suspected terrorist' are in fact crimes. They take the side that an "act or event that in itself produces no harm should nonetheless be criminalized, if it significantly increases the risk that future harm will ensue."[292] In other words, they claim that an act that produces significant risk of future harm, even without actually harming anyone, is a crime. This is a plausible account, and it is backed up by judicial history. Indeed, conspiracy to

289 *Id*.
290 *Id*.
291 Allen and Laudan, 2011, at 782.
292 *Id*.

commit crime has been criminalized in England since at least 1305. "What is true of conspiracies is likewise true of other inchoate crimes, such as solicitation of a criminal act, facilitation of a criminal act, incitement to a criminal act, threat of a criminal act, and attempt."[293] Given that preventative detention fits all the requirements of the definition of incapacitation that I have defended, it provides a good example of physical incapacitation.

Furthermore, the act of keeping an offender in jail prior to trial, by denying bail, is an act of incapacitation. It is retributive because it is in response to a crime for which the state has good reason to believe that the offender is responsible. It is clearly a physical restriction; and it is usually reserved for offenders suspected to be dangerous. There is good reason to believe that these offenders do in fact pose significant risk to society. In the USA,

> 13% of all homicide arrests from 1990 to 2002 involved persons who were either on active bail or were fugitives from bail... Supposing that most of those arrested for homicide are guilty and that most homicides that take place would not have occurred if the actual perpetrator had been incapacitated, a policy of denying bail to all defendants accused of a violent crime or with a history of violent crimes, it follows that we could prevent the deaths of around 1,500 citizens every year.[294]

The goal of preventing the large number of murders committed by suspected offenders on bail awaiting trial for violent crimes has nothing to do with intentionally harming the suspected offender, but is purely preventive. Having given examples of current practices which qualify as incapacitation, I turn now to the discussion of whether such practices may be justified.

3.5 The Benefit to Society Justification and Objections

In this section I will describe the one plausible justification for incapacitation, as well as the various objections to it. I will also discuss the relationship between incapacitation and punishment. As I have mentioned before, the only justification for incapacitation is preventing future offenses. This Benefit to Society Justification in relation to incapacitation faces similar objections as it did in relation

293 *Id*. at 789.
294 *Id*. at 796-797.

to rehabilitation.

First, one might object that this justification would justify incapacitating innocent people. But this objection fails, because as I argued above, if we accept that nascent planning or risk creation for future crimes can be criminalized, then there is no way the state could deem someone dangerous enough to justify incapacitating without that person being convicted of a past crime.

Secondly, the Benefit to Society Justification could justify severe restrictions that are disproportionate in relation to the offender's original offense. This is certainly true. It justifies keeping a criminally insane offender detained indefinitely if he is shown to be dangerous enough. Critics argue that such treatment may be disproportionate to their desert. Aside from the criminally insane, many justice reform advocates believe that an example of disproportionate restriction is of the Miami offenders forced to live on an island under a highway. They see it as a step too far in restricting the residency conditions of offenders who have been released from prison where they 'served their time.' And beyond these real-world restrictions, the justification may provide reason for a state's egregious abuse of power. If sentences of life imprisonment were imposed for minor offenses such as motorist speeding, it would almost certainly benefit society by reducing deaths on the road, but this appalling consequence would be reason enough to reject the justification.

In response to this objection, it seems clear that the extreme cases, such as the life sentence of incapacitation for speeding, would never be justified based on reasonable risk assessments, and in any case would be inflicting worse suffering than they are preventing. For the more modest concerns, the only plausible response would be to bite the bullet and admit that the benefit to society may justify unacceptable practices. However, it may also be the case that the bad consequences of rejecting the state's ability to incapacitate could outweigh the bad consequences of disproportionate restriction. As noted above, the hundreds of murder victims who are killed by suspected violent offenders out on bail each year is a compelling reason to deny bail to offenders accused of dangerous crime. Extrapolating from this example, it seems equally unacceptable to let psychopathic offenders with violent tendencies

roam the streets out of fear for disproportionately restricting them. The best way forward would seem to be admitting that the justification makes room for inexcusable abuse, and for a system of regulation to be set in order to check for abuse. We could expect that with enough attention to risk assessment and transparency, a state could do a decent job in preventing grossly excessive incapacitation.

A theorist might reply two ways, but both of these responses fail. First, one could claim that the harms of biting the bullet would outweigh the actual benefit to society. This response fails because it mischaracterizes the context of the incapacitative interventions, and also because it miscalculates the harms. The second reply is to claim that while the benefits to society may outweigh the potential abuses, an even better route would be to attempt to punish such offenders according to proportional desert rather than to incapacitate them. This argument fails because of a pollyannaish view of our ability to punish according to proportionality.

To begin with the first reply, the idea of a post-arrest regime 'letting' abuses of power transpire unchecked is deceptive. There is no reason to believe that a system based on rigorous assessments of dangerousness with stringent oversight and commitment to minimizing foreseeable[295] inflicted harm may not address most of the concerns over excessive incapacitation. Of course even in such a well run system, there will inevitably be 'false positives,' or cases when an offender is deemed to be too dangerous to let loose but who in fact would not have reoffended. But the overall benefit of being able to incapacitate dangerous offenders clearly outweighs the harms of false positives. To give evidence for this claim, the level of false positives is relatively low, and likely to dip lower as the science and implementation of risk assessment advances. Take De Keijser's following summary of a study of risk assessment used to determine whether Dutch offenders would be preventatively incapacitated. He is explicitly critical of the effectiveness of risk assessment, but as I will explain after the quotation, he has no real grounds to be as critical as he is. He explains that the study compares those offenders who

295 I refer again to Kolber, 2012, for an argument of why we need to justify reasonably foreseeable harms inflicted on offenders, not just the punitive, intended harms.

actually reoffended and those who did not with their respective diagnostic levels of risk.

> Of the actual recidivists, 84 percent did indeed have a high risk score… These are correct predictions. Not bad. However, 29 percent of those who did *not* recidivate also received a high risk score. These are false positives, and they tend to be overlooked… If the risk score was used as the exclusive basis for deciding on prolonging [the incapacitation], then for almost one-third of those who would actually desist from crime the prolongation is unjustified… Compare, for instance, the diagnostic value of a witness recognizing the suspect from a correctly performed police lineup:… 75 percent correct versus 5 percent false positives… Can that be justified? Yes it can… As a consequentialist, one may thus defend the position that, once within the reach of criminal justice, in order to protect society against more than four out of five dangerous persons, it is necessary to accept that almost one out of three persons who are not (anymore) dangerous will also be incapacitated. To me this seems a difficult position to defend.[296]

In a very brief and unconvincing explanation of why he sees this as such an indefensible position, he cites the difficulty of knowing when such "mischief" (presumably referring to false positives) overrides the benefit. To address this worry, it seems plain to me that the level of mischief that would be tolerable to a society would be much higher than the level portrayed in the Dutch system. The benefit of preventing such a high percentage of dangerous criminals from reoffending seems clearly worth the cost of to incapacitating other dangerous criminals who may not have actually reoffended. To add a few more reasons to accept my evaluation, the predictive value of the risk assessment tool is systematically underrated because of the high percentage of offenders who commit crime but do not get caught. This means that many of those who did not reoffend according to the study may have actually reoffended but not been caught. In 2004 in the USA, only 46.3 percent of violent crimes resulted in an arrest, and in 2011 the percentage was almost identical, at 47.7 percent.[297] Keep in mind also that these offenders deemed dangerous are guilty of past crimes that help demonstrate their dangerousness, and that they are being kept in conditions that try to minimize their suffering rather than intentionally inflict it. If offenders who in the past demonstrated truly dangerous behavior, and who have a one in three chance of no longer being dangerous, need to be confined to a reasonably comfortable living situation from

296 De Keijser, 2011, at 205.
297 Federal Bureau of Investigation, 2004 and 2011, *Crime in the United States, Clearances*.

which they cannot leave in order to prevent "four out of five dangerous persons" from reoffending, I feel comfortable claiming that such a trade off is one that any society should and would make, and far from what De Keijser deemed as "intolerable inaccuracy." The only other attempt De Keijser makes to rebuke such a trade off is by quoting von Hirsch who claims that society has an obligation "to do *individual* justice." By this I assume that von Hirsch means punishing individuals proportionally according to their desert. This is the route taken by the second reply, to which I now turn.

The second reply claims that rather than the two bad options of: (1) incapacitating none of the dangerous offenders, and (2) incapacitating some of them beyond what they deserve, we should instead focus on trying to punish them according to desert. Setting aside for a moment the concerns of justifying punishment and making the (quite large) assumption that it can somehow be justified, and focusing only on those violent offenders deemed highly dangerous by risk assessment tools: even then, the option of incapacitation is preferable over proportional punishment. The reason for this is because of the massive uncertainty in how to try to assess what a particular offense actually deserves in terms of proportional punishment. Without getting into the details of the enormous and ever-burgeoning amount of very compelling work done on the problem of proportionality in punishment,[298] I feel confident claiming that it is absolutely certain that the level of scientific confidence we have in our ability to predict the level of a particular offender's dangerousness far outstrips our ability to tell the amount of deserved suffering we should inflict on him. This is an issue of practicality: given the limits of our current abilities, we should choose the diagnostic tool that allows for acceptable precision over the one that is mired in inconsistency. The further benefit of this route is that it aims to minimize harm to the incapacitated offender, thereby avoiding the very messy task of needing to justify intentionally harming him.

[298] See Tonry, 2011c, for a good overview. For some of the most compelling original contributions to this discussion see Kolber, 2009a, 2009b, and 2012.

Chapter Four: Restorative Justice and Rituals

Restorative justice scholarship has exploded over the course of the past couple decades. Especially since the 1989 publication of John Braithwaite's now-canonized *Crime, Shame and Reintegration*, scholars from an impressively varied array of disciplines have been contributing to the conversation with accelerating prolificacy. Despite this intense focus, there has been little agreement on just what restorative justice is. Because a discussion of alternatives to punishment would be remiss to neglect this new and burgeoning body of work, I will address it in this chapter. But on further inspection, as I will argue, restorative justice plays a less important role in the discussion of the morality of non-punitive alternatives than its current popularity might lead one to believe. The usefulness of restorative justice for this discussion boils down to providing a compelling example of a ritual that may be used in a minimally punitive post-arrest regime. This ritual is the victim-offender conference, and while it has been extensively studied, it is appropriate only for a portion of criminal cases, thus its role will be restricted to that limited range of cases.[299] In the first section of this chapter I discuss the restorative justice movement, its relationship with punishment, its limitations and its justifications. However, the victim-offender conference is one example of a broader alternative to punishment, that of non-punitive rituals. While the restorative justice conference is currently the trendiest example of non-punitive rituals, it is not the most important: criminal trial is arguably the most important non-punitive criminal justice ritual, and any crime response regime, whether punitive or minimally punitive, will require it. Along with trial, I briefly discuss therapeutic jurisprudence. Other important rituals include reentry ceremonies and apology rituals. In the second part of this chapter I discuss trial, reentry ceremonies and apology rituals, as well as their justifications and why very few rituals (with the obvious exception of criminal trial) should be mandated.

299 The range of cases that could be handled with restorative justice conferences is debated, but it is certainly well below 100% of criminal cases, as I will show below.

4.1 Restorative Justice

It is fair to claim that the core of the restorative justice movement is the practice of holding conferences to mediate between the victim and the offender.[300] These go by many names in the literature, including: victim-offender mediation, accountability conferences, family group conferences, diversionary conferences, restorative justice conferences, victim-offender conferences etc. I will refer to them hereafter simply as 'conferences.' There is of course a wide variety of conference structures and practices, but they "share common features including a community-based sanctioning focus, non-adversarial and informal processes, and decision-making by consensus."[301] Of course, simplifying a vast, theory-based movement to its current practical application is an unfair way to characterize the movement. But I am not so much concerned with the broad movement and theory, but rather its candidates for on-the-ground alternatives to punishment. While it is beyond the scope of this dissertation to evaluate the various restorative justice (RJ) theories, it is appropriate to scrutinize conferences as alternatives to punishment. And before delving straight into a discussion of the conferences, I should give a brief impression of the wide variety of debates taking place on a theoretical level among RJ scholars, and explain further why these debates need not affect this study of non-punitive alternatives.

4.1.1 Definitions and Theory

Braithwaite gives perhaps the broadest possible definition of RJ: "Restorative justice is about struggling against injustice in the most restorative way we can manage."[302] Such a broad interpretation of RJ is common, and it allows for further general claims about RJ:

> Braithwaite: RJ "is not simply a way of reforming the criminal justice system, it is a way of transforming the entire legal system, our family lives, our conduct in the workplace, our practice of politics. Its vision is of a holistic change in the way we do

300 Okimoto et al., 2009, at 158.
301 Bergseth and Bouffard, 2007, at 434.
302 Braithwaite, 2003, at 1.

justice in the world."

> Walgrave: RJ is "every action that is primarily oriented towards doing justice by restoring the harm that has been caused by a crime."[303]
>
> Von Hirsch et al.: RJ "has emerged as a potent alternative paradigm in criminal justice."[304]

While still general, I admire the following definition of restorative justice, because it seems to capture the cluster of concepts associated with RJ more succinctly than usual:

> Zehr: [R]estorative justice is a process to involve, to the extent possible, those who have a stake in a specific offence and to collectively identify and address harms, needs and obligations, in order to heal and put things as right as possible.[305]

What should we take from these characterizations? First and foremost, they show the generality and ambitious scope with which RJ is often described. From just these four (among hundreds of candidates) we can glean that RJ has a debatable role within (or external to) the criminal justice system. We also get a glimpse into some of the key principles of RJ: restoration for harm, inclusion of stakeholders, and collective decision making.

If your introduction to RJ is through characterizations such as these, then it would probably not be a surprise to learn that the RJ literature is rife with internal debates. There are generally two kinds of debates: those arising from critiques of the RJ movement as a whole, and those arising from specific questions of scope and clarification. In the first camp, Von Hirsch et al. provide a nuanced and harsh critique of the RJ movement, claiming that it has "(1) Multiple and unclear goals... (2) Underspecified means and modalities... (3) Few or no dispositional criteria... (4) Dangling standards for evaluation."[306] They go on to defend a model of RJ that looks very similar to the characterizations that they criticize, explaining how their version can be specified and evaluated in a way that avoids some of their critiques.

303 Walgrave, 2003, at 62.
304 Von Hirsch et al., 2003, at v.
305 Zehr, 2002, at 37; quoted from Lokanan, 2009, at 290. Lokanan also asserts that Zehr's definition is the best available.
306 Von Hirsch, Ashworth and Shearing, 2003, at 22-23.

The debates become more interesting when they take on more focused questions. Theo Gavrielides offers an incredibly useful account of the various debates within RJ scholarship. He describes six different debates, and in order from the most broad to the most specific they are: (1) RJ and its relationship with the criminal justice system; (2) RJ and its position within the criminal justice system; (3) a definition for RJ; (4) stakeholders in RJ; (5) RJ and punishment; and finally (6) RJ principles and their flexibility.[307] Keeping these six layers of debate in mind is very helpful for parsing the RJ literature, which can often seem muddied by debates in which the two sides talk past each other. It also clearly demonstrates the complexity of the subject matter.

For the purposes of this dissertation, I will only be dealing explicitly with the question of RJ and punishment. I will avoid the other debates by assuming their outcomes or denying their relevance to my inquiry. Regarding the first two debates, I assume that RJ conferences can fit within the current court-dominated paradigm of criminal justice. I assume that eligible offenders convicted by the courts could participate in and receive their formal sentencing from these conferences. I should note that by sidestepping these debates I am avoiding discussion of some very useful critiques of the criminal justice system. For instance, Braithwaite argues that criminal justice ignores very important questions that RJ may be able to answer, namely "questions about whether the injustice would be better addressed by a family, by providing economic incentives for just behaviour, by just speech, as opposed to criminalisation."[308] Braithwaite also critiques the field of philosophy of punishment by claiming that the field has "a puzzling lack of interest in when it is appropriate for state versus non-state actors to do the punishing."[309] These crucial questions are outside the scope of this dissertation. Regarding the third debate over defining RJ, I will not offer a broad definition, and will restrict my discussion to the brief examples I gave above. I will offer a more focused definition of the conferences in the following section. And regarding the questions of stakeholders and principles, I will assume that a conventional

307 Gavrielides, 2008, at 170.
308 Braithwaite, 2003, at 6.
309 *Id*.

reading of victim-offender conference structure will be able to sidestep these debates. My analysis of the conferences as alternatives to punishment will not change based on the exact composition of the conference participants. By resorting to vague terms such as 'community members' I am purposefully leaving these questions unanswered, as they are best addressed by criminologists and others who have direct experience with such conferences. And the conventional conference structure manifests fairly uncontroversial principles. Most will agree that RJ's principles include restoration, accountability[310] and community decision making among others.

4.1.2 Conferences

As I stated before, it is fair to claim that conferences are the core of RJ. They are the practical application of all the theorizing. As Okimoto et al. put it, "conferencing between involved parties is common to the majority of restorative programs."[311] A crime response regime aiming to replace or minimize punishment should be clear about the benefits and limitations of conferences, as well as their relationship with punishment.

First I will briefly sketch what these conferences look like. I attempt to describe the basic components of conferences such as who takes part, who is eligible, how they are run and what the outcomes are. Most conference programs focus on crimes in which there is a clear offender and victim, and often, as is the case of a fistfight, involved parties can be both offender and victim. Calhoun and Pelech write that "[c]onferences have been held for a wide range of behaviors but mostly for interpersonal violence (e.g. assault, assault with a weapon, armed robbery) or property-related offenses (e.g. breaking and entering, theft, fraud)."[312] The participants include the victim, offender, community members and a mediator. Of course there is debate over exactly whom these terms refer to and what

310 In contrast to the conventional, punitive dialogue in which "accountability" is equated with punishment and even prison, Calhoun and Pelech point out that, "[f]rom the restorative perspective, "accountability" generally connotes some aspect of the broad relational and reintegrative goals of the approach;" 2010, at 292.
311 Okimoto et al., 2009, at 158.
312 Calhoun and Pelech, 2010, at 290.

their role should be, but this general list is uncontroversial.[313]

Next, conference programs require that the offender admit guilt before he is eligible to participate in a conference. "[C]onferencing models emphasize the voluntary inclusion of anyone affected by the wrongdoing and the need to provide a safe forum for truth-telling, encounter and a negotiated reparation agreement."[314] For instance, the only criterion that Calgary Community Conferencing (CCC) requires

> for accepting referrals is that the person who wronged accept responsibility for the wrongdoing... Conferencing activities proceed from the referral stage only when the youth and the individual(s) harmed voluntarily agree to become part of the process.[315]

Note that both the offender and the victim need to consent to participate in the process, which is the norm for conferences. This voluntary nature of the conference model restricts the range of cases that conferences may deal with. Given that, as Daly notes, "the adjudication process rests on a fundamental right of those accused to say they did not commit an offence," the state's post-arrest regime will need to be able to accommodate accused offenders who deny their guilt, which is a simple and powerful reason that conferences may not completely usurp the 'traditional' trial system. The apparatus of investigation, prosecution and trial exists because accused offenders have this right to deny their guilt. And while some offenders who admit guilt up front may be diverted to the RJ conferences, there will remain many more who are not eligible, either because their victims refuse to participate, or their crimes make them ineligible for their local conference program. As Daly observes, "RJ is limited by the abilities and interests of offenders and victims to think and act in ways we may define as restorative."[316] This is part of the reason that I am comfortable assuming that any viable application of RJ will have to work with the current criminal justice system, and not completely replace it—because RJ conferences could not deal with the same range of cases that traditional criminal justice can handle.

313 See, e.g., Cullen and Jonson, 2012, at 130; and Daly, 2003.
314 Calhoun and Pelech, 2010, at 289.
315 *Id*. at 290.
316 Daly, 2006, at 143.

The actual conferences are typically informal processes that involve open dialogue. Daly describes the conference process from the South Australia Juvenile Justice (SAJJ) project on conferencing:

> A victim and *admitted* offender and their supporters come together to discuss the offence, its impact and an appropriate penalty (agreement or outcome). The discussion evokes feelings of remorse in the offender, which leads to a genuine apology and a desire to repair the harm... Participants then discuss an appropriate penalty (or agreement of outcome). Everyone has a say, and participation by the professionals is kept to a minimum.[317]

Usually after the end of conferences the offender signs a legally binding agreement "which itemises the group's outcome decision."[318] Of course, this is the ideal version of events, what is hoped for from the outset. Daly goes on to describe in detail where practice and theory diverge in her experience with the SAJJ. Some common complaints about conferences are that they can be poorly run, the offender can be insincere, and the victim can leave feeling like justice was not restored. Generally, despite some serious pitfalls that professionals are busy pinpointing and working out, conferences have been a success over the past twenty some years since they were first implemented as an alternative to the 'traditional' court proceedings of the current dominant criminal justice paradigm.[319] Next I will briefly summarize these successes.

In a 2008 meta-analysis of the literature, Wenzel et al. conclude that "[g]enerally, the evidence suggests that victims feel more satisfied with restorative justice programs compared to court processes."[320] Beyond the literature that they cite, many other studies have come to a similar conclusion. In a summary of research on victim-offender mediation, Umbreit et al. found:

> [T]he vast majority of VOM participants (typically over 80%) across setting, cultures, and types of offense reported believing that the process was fair to both sides and that the resulting agreement was fair. Where comparison groups were employed, those individuals exposed to mediation came away more likely to feel that they had been treated fairly than those going through traditional court proceedings.[321]

317 Daly, 2003, at 220-221, original emphasis.
318 *Id*. at 221.
319 Many scholars point out that the RJ conference structure is based on ancient practices found throughout the world.
320 Wenzel, et al., 2008, at 377.
321 Umbreit et al., 2004, at 87.

Two other large meta-analyses that are widely cited and esteemed[322] echo these findings and also focus on the effectiveness of conferences in reducing recidivism. The first, Sherman and Strang, concludes that "*RJ seems to reduce crime more effectively with more rather than less serious crimes*... RJ works better with crimes involving personal victims than for crimes without them."[323] Pause for a moment to consider how striking this finding is: we normally associate RJ conferences with offenders guilty of less serious crimes, and indeed most conferences deal with such offenders. But they conclude that RJ works to reduce crime more effectively when it deals with the most serious offenders, such as rapists and murderers. I will refer back to this finding in my discussion below. These reductions are in comparison with control groups of offenders going through the normal courtroom channels. The second large study, from Bonta and Andrews, concludes that offenders who have gone through RJ conferencing have recidivism rates on average seven percent lower than control groups. They admit that this effect is significantly less potent than the impact of well-run rehabilitation programs.[324]

To assess the morality of conferences I mostly focus on the question of whether they are punitive. The question of whether they are justified is less pressing because of two reasons. First, many of the reasons for which the ritual of trial is justified also apply to conferences. For example, both practices condemn the offense and this may be enough to justify them. Secondly, RJ conferences are composed of two discrete state actions, namely holding the conference and sanctioning the offender's communal sentence. I will assume that the act of holding the conference is no less justifiable than the act of holding a trial, and thus skip discussing independent justifications for the conferences. The question of whether the state is justified in authorizing and carrying out the outcome of a conference, i.e. the offender's communal sentence, is intimately tied to whether that sentence is punitive, which is what I turn to now.

322 See Cullen and Jonson, 2012, at 142-146 for a good summary of these studies.
323 Sherman and Strang, 2007, at 8, original emphasis.
324 Bonta and Andrews, 2010, at 456.

4.1.3 RJ and Punishment

As Gavrielides notes, one of the central arguments within RJ scholarship is whether RJ should or could be punitive. Most of the theoretical debate has centered over the more conceptual inquiry of whether RJ conferences may include punishment and remain restorative. I will focus on this more specific formulation of the question: may the outcome of the RJ conference permissibly include punishment for the offender? Many RJ scholars "argue that there is no place for punitiveness when the goal of the justice process is to repair the harm that the offense has caused."[325] For example, Braithwaite and Strang argue that within the context of the goals and principles of RJ, "[r]esponding to the hurt of the crime with the hurt of punishment is rejected."[326] The reasons for this position stem from the intent of the RJ process. The conference, as the arguments typically go, is meant to restore its participants as much as possible. Even if the outcome of the conference includes some burden for the offender, the intention behind this burden is to help repair the victim, the community, and the offender's standing.

This argument immediately runs into a familiar practical concern: how are we to determine the intention behind a particular sentence arising from a conference? Unlike the traditional court system, RJ cannot point to legislative and judicial intent. Within most RJ conference models, a conference sentence arises from the conference participants. Of course, one could easily imagine conference models in which this is not the case. These could include maximum or minimum restrictions on conference sentences, and firm guidelines for the suggested outcome based on the type of crime. A judicial review of the sentence could be in place, with the ability to amend sentences that stray too far from the guidelines. If these restrictions were in place, the conference would retain its role as an occasion for discussion, public condemnation of the offense, apology and all the other valuable symbolic and communally expressive functions of RJ. However, it would lose much of its role as an event for communal decision making. The restrictions and guidelines would be set by separate bodies,

325 Gromet and Darley, 2006, at 397.
326 Braithwaite and Strang, 2001, at 1.

namely the legislature and judiciary, and these regulations would largely prescribe the offender's future. This solution would sidestep the problem of identifying the intention because it could be located the same way as in traditional criminal justice (the way that I discussed in the first chapter). However, a RJ conference with a prescribed outcome is a fundamentally unorthodox conception, and it would be unfair for an advocate of RJ to rely on this solution. Conferences normally allow for its participants to determine the outcome, which means that the intention behind the sentence rests solely with the participants. Thus the question of the intention behind the conference outcome leads to empirical questions of what typical conference participants intend.

Myriad studies show that members of the general public prefer punitive measures for offenders, especially for serious crimes. This leads to a dilemma for Braithwaite and other RJ scholars who want to keep punishment out of RJ. Put succinctly by Gromet and Darley, the problem is that "[i]f people are not willing to send more serious crimes to 'pure' restorative justice procedures, then this severely limits the viability of restorative justice as [a non-punitive] alternative to the traditional court system."[327] By "pure," they mean purely non-punitive. We know from many different studies that when participants are forced to choose between punitive and restorative measures for serious offenders, most choose punitive.[328] However, the details get more complicated when participants are given the chance to mix punitive and restorative measures, which is what the case would be in a RJ conference that had recourse to prescribing a partly punitive course for the offender.

Gromet and Darley composed a study that introduced participants to a range of offenders and gave the participants the choice of three routes to recommend for the offenders: (1) purely punitive prison sentences; (2) purely restorative conferences; and (3) to mix punitive prison sentences and restorative measures. Given this choice, the study reveals that most participants prefer a mix for serious crimes (e.g. burglary, attempted murder, rape). Participants equally prefer the purely restorative

327 Gromet and Darley, 2006, at 399.
328 *Id*. 2006, at 400.

measures and the mix for mid-range crimes (e.g. mugging). And participants strongly prefer purely restorative measures for the least serious crimes (e.g. vandalism and bike theft).[329] Also, those who preferred the mix of prison and restoration gave significantly more lenient prison terms than those participants who preferred pure prison terms.[330] This is a noteworthy finding because it shows that an offender will likely be treated more leniently if he participates in a conference. This echoes previous research that shows that people are less harsh in sentencing when the offender demonstrates remorse. The authors conclude that:

> "As crimes increase in their severity, the system can exercise the option to send offenders to prison in addition to their fulfilling the conditions of the conference agreement. This finding is especially important because restorative justice procedures have been shown to be most effective for crimes that are high, rather than low, in seriousness."[331]

Note the repetition of the surprising finding that I cited above: RJ procedures work better to reduce recidivism with serious offenders. This mix of prison and restorative elements resulting from an RJ conference is currently rare in practice because most RJ conferences are used for mid to low seriousness of offenses despite the greater successes in reducing recidivism when using RJ with more serious offenders. The authors also admit that there are practical problems that would need to be tackled, for instance, whether the prison comes before or after an offender's restorative obligations.

Unsurprisingly, other scholars have advocated for this precise mixture of prison sentences and restorative justice. In such a "restorative regime, restorative prisons can be built."[332] These prisons would be made to help the prisoner via rehabilitation. Prisoners would have the chance to reflect on their needs and to work to reduce their risk to society. Using the conceptual distinctions I defended in the last chapter, we could surmise that the action of forcing an offender to go to such facilities would be either incapacitative or punitive. This is because the action of forcing the offender to go to the prison,

329 *Id*. at 406-407.
330 *Id.* at 406 and 422.
331 *Id*. at 410.
332 *Id*. at 302

even if it is a restorative prison, does not depend in any way on the offender's will. Then once inside such a restorative prison, he would have the chance to participate in rehabilitative programs. Would conference participants' intention be incapacitative or punitive? The intent would likely be punitive. Given the research I cited above, participants would likely act punitively, especially for serious crimes. Of course, we would have no sure way of knowing besides directly consulting the participants, but we can justifiably suppose that if their intentions are limited to either punishment or incapacitation, they are likely to be punitive. Furthermore, it seems unlikely that offenders who need to be incapacitated would qualify as eligible to participate in conferences. Presumably decisions of whether to incapacitate criminals should be left to experts who can calculate risk and not to lay people in a conference.

An RJ advocate might respond by claiming that people's punitive attitudes are influenced by the current ubiquity of the punitive regime. If the regime transitioned over time and became principally restorative, then people's attitudes would also switch over time and their intuitions would lead to non-punitive responses. While this is certainly possible, until that transition occurs, it would be unfair to consider conferences resulting in prison sentences to be alternatives to punishment. And even if the regime becomes principally restorative, it may still be the case that conference participants continue to intend to harm the offender when they choose to send him to prison.

Perhaps conference participants may intend to rehabilitate the offender by sending him to prison, presenting a third option of conference intention beyond punishment and incapacitation. The conference will always include the offender, and because the decision is arrived at communally, one might say that if the sentence includes a term of restorative prison, then the offender has consented to it. This may perhaps qualify the prison term as rehabilitative and non-punitive given the definition of rehabilitation that I defended in the previous chapter. But this is mistaken; when offenders are sent to restorative prison, it is without regard to their will or consent. Many scholars have wondered how voluntary the offenders' participation is when he may stand to benefit from choosing the conference

route as opposed to normal prosecution.[333] This choice between unequal results certainly amounts to a kind of coercion to choose the conference route. But under the minimal understanding of consent that I defend as a requirement for rehabilitation, coercion does not necessarily result in non-consent. If the offender can possibly choose otherwise, then his course of action may be considered rehabilitative. However using this minimal conception of consent, one could still not claim that an offender consents to his conference sentence of a prison term. The reason for this is that if he refused to sign the conference agreement then he would be transitioned back to the court system where he would surely be incarcerated. Claiming that an offender consents to his prison term by participating in the conference would be perverse because in reality he would not be able to avoid prison regardless; if he cannot choose otherwise, then he has not consented even using my minimal conception of consent.

Of course, one might defend the claim that a prison sentence resulting from a conference may be rehabilitative by rejecting my consent requirement of rehabilitation. This position entails two separate claims: that consent is not required for rehabilitation, and that sentencing the offender to prison qualifies as a kind of this non-consensual rehabilitation. As I argued in the previous chapter, I do not believe that the first claim can be defended. And furthermore, even if someone were to successfully defend such a conception of rehabilitation, it is unlikely that the second claim could be justified. When RJ advocates speak of restorative prisons, they speak of giving the offender time and opportunity for rehabilitative programs within the confines of prison. This way of speaking betrays that there are two discrete state actions involved: first is imprisoning and second is offering rehabilitation programs. An advocate of rehabilitative prison would need to show why imprisoning the offender is necessary to the rehabilitation.[334] Why not simply offer the programs? If the offender refuses, then perhaps he should be imprisoned, but then the act of imprisoning would seem straightforwardly punitive or incapacitative. Finally, I take the ability of the consent requirement of rehabilitation to help parse whether RJ

333 Gavrielides, 2008, at 176.

334 They would also need to confront empirical evidence showing that prison terms are often therapeutically counterproductive. See Kolind et al., 2010.

conferences are rehabilitative as further evidence of its viability.

So far I have been focusing on the question of whether conference participants who give a jail sentence act punitively. The reason for this comes in part from the empirical evidence, which presents a dilemma for RJ: conferences work best for serious offenders, but conferences are almost never used for these offenders. I have argued that conferences for offenders who would otherwise be facing serious jail time will most likely result in punitive sentences. Prison also provides a clear-cut example to discuss whether RJ is punitive. But most conferences do not result in prison; do these face the same problems as putative alternatives to punishment? The answer is that they do not. Research shows that the public is much more willing to grant purely restorative, non-punitive sentences in conferences for less serious offenders. For example, "Eight in ten (77%) adults believe the most appropriate sentence for nonviolent, nonserious offenders is supervised probation, restitution, community service, and/or rehabilitative services; if an offender fails in these alternatives, then prison or jail may be appropriate."[335] Furthermore, these sentences could be considered consensual. If a conference recommends that an offender guilty of vandalizing property spend a certain number of hours cleaning up graffiti, then he could choose otherwise. Even if the state made it clear that he would be punished for failing to complete the sentence, the sentence itself is separate from the threat, which is incapacitative and separate from the actual punishment. The offender had a substantive choice: between doing community service and not doing it, with the latter resulting in imprisonment. We can easily construe that he had a substantive ability to choose otherwise, in comparison to the choice between court-ordered prison and conference-ordered prison, which would be perverse to consider an ability to choose otherwise.

Now, this may seem jarring because it means that the difference in consent is based on whether the outcome will be the same or different. And it leaves room for a critic to offer the following objection. What if an offender were given the choice between 100 and 101 days in prison. Surely the

335 Hartney and Marchionna, 2009, at 1.

introduction of a choice between slightly different outcomes would not mean that he consented to prison. But this is not the notion of consent I am defending. As I defined consent in the previous chapter, the offender must have the choice between two substantively different options. The choice between 100 and 101 days in prison is not a substantively different choice, and likewise the choice between a conference-ordered prison sentence and a court-ordered prison sentence is not substantively different. A critic might press on the notion of 'substantive difference,' but this is not a problem for my account. In application, the distinction may be determined, implemented and monitored by regulatory bodies within the criminal justice system. There are plenty of other even trickier concepts that the criminal justice system handles, and institutions structured to guide practices and check abuse act to regulate the challenges surrounding these concepts. The notions of 'due process,' and 'reasonable doubt' are two similar examples.

I will end this discussion with an example of how clear definitions of punishment and rehabilitation can aid with confusion within RJ scholarship. Lokanan laments that oftentimes the term 'punishment' is used imprecisely in the RJ literature. Some scholars (like Daly) use it to mean any burden for the offender that results from the RJ process. Others take it to mean the more precise conception of a state action intended to harm the offender. Lokanan reacts to these differing accounts by claiming that "indiscriminate conceptualizations of punishment create conceptual muddle and linguistic imprecision in an already difficult area."[336] But instead of giving a rigorous account of what punishment is, he rather splits punishment into two categories: restorative and retributive. His distinction between two types of punishment is a strange move. It is, to my knowledge, unique in the literature. I mention it in order to demonstrate the conceptual trouble that occurs when RJ scholars are not clear in their use of the term 'punishment.' Lokanan admirably acknowledges the need for clarification, and in attempting to explain he makes some errors that are useful for understanding the relationship between RJ and punishment. He explains the differences between his two conceptions of

336 Lokanan, 2009, at 293.

punishment here:

> The fundamental difference between restorative and retributive punishments is twofold. First, punishment is not punitive in the restorative process. Rather, punishment is instrumental, as it is not imposed with the offender's negative moral status in mind, but rather the desirable consequences it is believed or hoped the disposition will have on the offender...Punishment is not intentional; it is a side-effect of the reparation process. Second, the emphasis on pain is practically non-existent. When punishment is applied, it is done with conditions that control and reduce the level of pain and in a context where restoration and healing are the main goals.[337]

This quotation provides good fodder for analysis. The most striking part of his explanation bears repeating with emphasis: "*punishment is not punitive in the restorative process.*" I admire his directness in evaluating his RJ colleagues and noticing confusion over the term 'punishment.' And I also agree with his basic approach of keeping two separate categories for acts that are intended to harm the offender and others which are aimed at getting the offender to acknowledge the harm and make amends for it. However, it seems obvious to me that the punitive category of actions qualifies as 'punishment' and the other does not. What Lokanan describes as restorative punishment fits well with a definition of rehabilitation. It is instrumental, aiming to change the offender. Pain inflicted on the offender is not intended, but rather a regrettable side-effect that is minimized as much as possible. This is another good example of how conceptual clarity in defining punishment and rehabilitation can have wide ranging benefits, which include injecting some clarity amidst the RJ literature.

4.2 Other Rituals

So far in this chapter I have focused on one non-punitive ritual, the victim-offender conference. Whether a conference is actually non-punitive depends on its participants, as well as the seriousness of the crime they are considering and their intention. It also depends on whether the conference participants receive guidelines and constraints on their sentencing choices from a legislative body.[338]

337 *Id*.

338 I only briefly noted this option before, but it seems to me the clearest and most responsible way forward for incorporating conferences into a minimally punitive regime.

The other rituals that I turn to now are more straightforwardly non-punitive. While some may take it as obvious that trials, therapeutic jurisprudence, apologies and reentry ceremonies are non-punitive, here I will explain why they are. I will also consider the implications of coercing each of these rituals. I conclude that unlike trial and specialty courts such as drug courts, apologies and reentry ceremonies should rarely, if ever, be coerced.

I do not think that this is the place for pursuing a rigorous philosophical definition of ritual. There are no substantive questions raised in this dissertation that depend on such a definition, and generally the term is used uncontroversially within criminal justice literature. Rituals are symbolic practices that may or may not have further functional uses. Shadd Maruna is one of the foremost scholar on rituals in criminal justice, and he writes that, "[a]bove all, a ritual is a medium of communication, with its own symbolic grammar and syntax."[339]

4.2.1 Trial and Therapeutic Jurisprudence

Trial is the most important ritual in criminal justice. It serves many functions: determining guilt, condemning crime, redeeming innocence, expressing various commitments such as impartiality, transparency and human rights. Many trials also serve to determine the proper sentence for a guilty offender. Trials are not intended to harm the offender. It may result in a sentence that is punitive, but the trial itself remains non-punitive. Just as the threat of punishment from not participating in a rehabilitation program does not transform the program from rehabilitative to punitive, the threat of punishment for not participating in trial does not change the underlying character of the trial. Punishment as a result of not participating in trial cannot be considered part of the trial, as they are discrete actions. This dissertation is not the place to attempt even a cursory description of trial and its purposes; of all the topics that I am discussing, trial is the most exhaustively documented and analyzed. Duff writes that, "the criminal trial is the formal culmination of the criminal process" and that it is "a

339 Maruna, 2011, at 7.

formal process through which an alleged wrongdoer is called to answer to his fellow citizens by the court that speaks in their name." He goes on to claim that the criminal trial "is precisely focused on the wrongdoing, not just on the harm that might also have been caused, and focuses on that wrongdoing as a public rather than a private matter: the polity as a whole calls the alleged perpetrator to account for a wrong that concerns all citizens."[340] I assume for the sake of brevity that of the many purposes of trial, some or all of them serve to successfully justify coercing people to take part. It seems uncontroversial for a state to claim that the importance of establishing guilt and protecting innocence is far more important than an individual's preference to participate, and that it is justified in acting accordingly. Furthermore, I assume that these purposes also help serve to justify detaining offenders who are considered a flight risk. Ensuring that offenders participate will mean keeping them detained if the state has good reason to think that they might run away. Before moving forward, I should note that of the many purposes that trial serves, perhaps its most important is aimed at establishing fact. This emphasis on external affairs in part explains why trial is so different from apology and reentry rituals. Trial aspires to establish historical, objective record. In contrast, apology and reentry rituals aim to signify an inner transformation of self. This difference helps explain why trial can reasonably be a coerced ritual, and why there are objections to coercing the other two.

'Therapeutic justice' is a term that denotes a rising movement to filter offenders to specialized courts instead of traditional courts. "The types of specialized courts that exist at present include domestic violence, community, mental health, alcohol abuse, homeless, and a host of others."[341] But by far the most prominent of these courts are drug courts, and for the purpose of illustration, I will limit my discussion to drug courts. Douglas Husak writes that in these courts, "eligible offenders are given the option of subjecting themselves to normal adjudication or undergoing a treatment regime designed and overseen by the specialty court... Failure to complete this regime results in prosecution of (when

340 Duff, 2011b, at 76.
341 Husak, 2011, at 217.

courts are postconviction) the imposition of the deferred sentence for the underlying offense."[342] In this way, drug courts may be seen as rehabilitation supervisors, with the explicit threat of punishment overhanging the process. It is clear that participation in the court-supervised treatment is not punitive. The threat of punishment makes drug courts a source of transparent legal incapacitation. Drug courts are punitive only when judges send an offender back to traditional court to be prosecuted, or impose a deferred sentence. The therapeutic justice movement conceptualizes the courts as alternatives to punishment that have strict discipline in ushering offenders through a treatment regime, and that "real *punishments* are held in reserve in case defendants fail to complete their treatment."[343] Husak notes the uneasy alliance of goals: rehabilitative until the offender fails to comply, then harshly punitive (most often, offenders who fail out of drug courts receive harsher treatment than they would have if they had not entered it). In a minimally punitive regime, it seems like the goal of motivating participation could justify punishments for failing to comply with a treatment regime, but these punishments would be nominal, or only as punitive as needed to motivate behavior. Jeremy Travis, citing research on drug courts, concludes that "the severity of the sanction is less important in securing compliance with social norms than the consistency and predictability of a system of incentives and modest sanctions administered in a respectful manner."[344] In a minimally punitive regime, an offender who serves his modest punishment should get another chance at completing treatment under the drug court. This would better account for the often-zigzagging road offenders take to desistence. Others have also noted the communicative strength of drug courts in comparison to traditional courts. Addressing this comparison, O'Hear writes that drug courts communicate "more directly and forcefully to the particularities of the offense and what the offender must do to reconcile himself with the community. Indeed, even though drug court is typically defended in utilitarian, forward-looking terms, its communicative qualities suggest the possibility of justification on retributive, backward looking

342 *Id*.
343 *Id*. at 221.
344 Travis, 2005, at 348.

grounds.

4.2.2 Apology

Apology has been categorized as a "ritual of inclusion."[345] It functions to bond people. More specifically, Petrucci writes that apology is composed of the following core elements: "(i) an expression of remorse or regret, such as 'I'm sorry'; (ii) an overt acceptance of responsibility for the harmful act; (iii) some type of offer of compensation, repair, or restitution; and (iv) a promise to avoid such behavior in the future."[346]This is a stringent set of requirements for an act to be considered an apology, and many people would understand the first requirement, an utterance of remorse or regret, as the bare requirement necessary for an act to be considered an apology. It seems that an act that fulfills Petrucci's requirements amounts to something like a 'full apology.' Furthermore, an apology can fulfill the requirements to be a full apology, but still be ineffective. For instance, an apology can be insincere. Petrucci notes that, "[t[hree concepts surface as central to effective apologies: communicating emotion (such as remorse or sadness), a face-to-face interaction, and the timing of the apology."[347] If an offender refuses to admit guilt until after he is convicted, then calls the victim to apologize without expressing emotion, then it seems unlikely that the victim will perceive the apology as genuine. He may surmise that the offender is apologizing in order to lessen his sentence. For these reasons, it is easy to see why conferences are an ideal place for apologies. They happen face-to-face, after the offender has admitted full guilt, and before the offender has been sentenced. They also happen in the context of a community, and, "[r]estorative justice research suggests that most participants do spontaneously apologize for their crimes when confronted by their family members and the victims of their actions."[348]

345 Ward and Maruna, 2007,at 15, citing Braithwaite and Mugford, 1994, at 153.
346 Petrucci, 2002, at 341.
347 *Id*. at 343.
348 Maruna, 2011, 15.

What are the benefits of apology? Victims, in seeking to understand what happened, tend to find ways to blame themselves and hold themselves accountable as having some responsibility for not preventing the crime. Apology can clear this burden and show that the crime was totally out of their control.[349] Additionally, victims generally want to hear apologies. Multiple studies confirm that victims valued apology as part of conferences. Research has also shown that victims who accept apologies are less angry.[350] This gives rise to questions about whether the victim should accept the apology, which I discuss below. But apology also benefits the offender and society. "Several studies have found that offenders are interested in apologizing."[351] A New Zealand study showed that "after 3 years youth who did not apologize were three times more likely to be reconvicted than offending youth who did apologize,"[352] but of course we don't know if this is causation or correlation. It is also important to note that these findings are preliminary, and there is a dearth of research isolating the effects of apology in criminal justice. Petrucci, who cites the above studies, also notes that claims about apology in criminal justice are "essentially untested."[353] Most findings cite apology as part of conferences, which are much better documented. As I noted above, conferences have an effect on reducing recidivism, and apology is usually an integral part of conferences. It seems fair to claim that usually apology has beneficial effects. However, this should not be reason to coerce apology. Given that effective apologies are sincere ones, the beneficial effects to the victim and offender would dissipate if the apology did not signify a genuine change of heart. Victims perceiving that an apology is insincere would resent the apology, and the offender would not have undergone the internal confrontation of his wrongdoing. Unlike the state's obvious reasons for coercing participation in trial, the very reasons that apologies are desirable explain why they are impotent when coerced.

Are there any downsides to apology? There is research that shows that victims feel compelled

349 Petrucci, 2002, at 352.
350 *Id.* at 356.
351 *Id*. at 354.
352 *Id*. at 357.
353 *Id*. at 344.

by social norms to accept apologies.[354] I speculate that this may undercut the dignity of victims who would rather not accept the apology but feel compelled to do so. This leads to the further questions of (1) whether victims have a duty to accept apologies, and (2) whether offenders have the right to make amends and apologize. Some scholars have claimed that when a genuine, full apology is offered, the recipient has a duty to 'accept' this. The terminology is a bit vague in this arena. Some writers distinguish between the terms 'accept' and 'forgive,' and others use 'accept' to include forgiveness. Lazar separates the terms: "When a full apology is made, we can say that it is reasonable to expect the victim to at least accept the apology, though forgiveness is another matter entirely."[355] And Radzik argues that acceptance amounts to forgiveness: "An apology cannot repair a wrong unless it is accepted. But a victim cannot accept an apology without thereby committing himself to forgive the wrongdoer. That is what "I accept your apology" means."[356] It seems like an impossibly burdensome task to explain how a victim has a positive duty to hear, let alone accept an offender's apology or to forgive him. Part of the reason for which conferences may only function in a limited capacity is because all participants need to take part voluntarily, which means that victims who would rather not confront their offender may abstain. When Lazar separates acceptance from forgiveness, presumably he means that to accept an apology is tantamount to listening to the apology, leaving the reaction to the apology ambiguous. In normal usage, apology acceptance is a reaction to hearing the apology; and if this reaction is not full-blown forgiveness, it is at least a commitment toward reconciliation. To expect victims to hear the apology, let alone to expect them to react a certain way, is unreasonable because it places an undue burden on the victim without justification.

One possible route to attempting to justify victims' duty to hear or accept apologies is by advancing an offender's right to make amends. This right would entail the state enabling offenders to pursue making genuine amends, which may include restitution, apology and rehabilitation. Sayre-

354 *Id*. at 356.
355 Lazar, 2008, at 367.
356 Radzik, 2003, at 332.

McCord explains this argument:

> [C]riminals—in addition to having a right to be treated as responsible agents—might also deserve the opportunity to make amends... A new injustice comes in (except perhaps in the most extreme cases) if a person who commits a crime is denied the opportunity to make amends. There are limits here—some crimes are so heinous that, in committing them, a person has forfeited this right—but most crimes surely don't fall into this category... While I am somewhat ambivalent about seeing this as a right the criminal has, I think there is a point to the idea that people should be given the opportunity to "pay their debt to society."[357]

I will join Sayre-McCord and remain ambivalent about whether offenders have this right. There are good arguments for it, and Radzik defends them convincingly. She also holds the position that this right cannot usurp the victim's right to refuse to interact with the offender. She writes that, "Although wrongdoers have a duty to make amends, they are not morally permitted to attempt amends if doing so would cause extra suffering to the people they have offended."[358] Victims should also be allowed the right to refuse to forgive as a protest against the wrong done to them. As Lazar puts it, "There is, to my knowledge, no plausible conception of justice which could justify demanding that a person feel a certain way."[359] In some cases of irreparable harm, refusal to forgive may be the most appropriate response. Some victims of the Holocaust famously refused reparations as a protest in this spirit. However, to take a trivial case, if a neighbor forever refuses a young boy for vandalizing his mail box, we would find it odd and look down on his response, but we still cannot ask him to give up his right as a victim to refuse to forgive. Furthermore, an offender's right to make amends need not require the participation of the victim. He could apologize to a surrogate if the victim refuses to see him. He could also participate in acts of reparation, rehabilitation and restitution without the victim's involvement, thus enacting his right to make amends in these ways.

357 Sayre-McCord, 2001, at 512.
358 Radzik, 2003, at 338.
359 Lazar, 2008, at 366.

4.2.3 Reentry

A trendy observation in criminal justice literature is that 'they all come back,' referring to the fact that virtually every offender reenters society eventually.[360] How to treat offenders after they complete their sentences is a perennial top priority and challenge, as it is likely the most important factor in reducing recidivism. A state can only do so much to deter, punish and rehabilitate an offender, but if that offender is left to walk back into his old neighborhood without job prospects and if his strongest connections are to friends who are still committing crime, then the state stands nearly powerless to prevent his recidivism. Managing offender reentry is a fascinating subject with a voluminous literature, and I will limit my discussion here to a short appeal to the need for more robust reentry rituals, and a brief discussion of reintegrating the worst offenders. When I use the term 'reentry ritual' I am referring to ceremonies that recognize an offender as having 'paid his due' and is deemed officially ready to be fully reintegrated. There are many other reentry procedures, some of which may be deemed rituals. For instance the parole system can be viewed as a liminal apparatus to manage offenders' transitions. In this vein, Travis proposes creating "reentry courts," in which offenders would receive the individualized treatment that drug courts can offer. "Just as drug court judges celebrate every clean urine test… a reentry court judge would celebrate a job offer… or the completion of a community service project. On the other hand, the reentry court judge, just like his drug court counterpart, would be able to order curfews, electronic monitoring," etc.[361] I will not be considering specific procedures related to helping and preparing offenders to reenter society. I will rather assume that after the transition has been made, reentry rituals may signify the completion of the journey.

Maruna points out that when someone becomes an offender, it is a credential that signals something about him as a person. "Other credentials work this way too. You pass your Ph.D. and you become a 'doctor' with all of the connotations of special insight and capability implicit with that title. Quickly, the person moves from something one has achieved (finished a big piece of research) to what

360 See Travis, 2005, for an especially informative book-length treatment of this subject.
361 Travis, 2005, at 351.

someone is ('My son, the doctor').”[362] Maruna argues that criminal credentialing is unidirectional: there currently exists no way for an offender to signal to society that he has made his reparations and should be welcomed as a fully reinstated member of society. Maruna argues that certificates of rehabilitation, acting as a kind of letter of reference for employers and others, could help to fill this gap. This gap is especially large in the USA, and outside of it there are good examples of desistance credentialing. “In France, 'judicial rehabilitation' rituals take place in the same court rooms that sentence individuals to prison and (not coincidentally) 'resemble citizenship ceremonies.'”[363]

I should note that reentry rituals are not alternatives to punishment in the sense of replacing punitive post-conviction sentences, which is the sense I have thus far used in this dissertation. However, if one considers the totality of the state's post-conviction interventions with offenders, then one must include reentry procedures and rituals. Maruna and others point out that for all our intense focus on the initial interventions after a criminal conviction, we have neglected the counter-balance to the downward credentialing of labeling people as 'offenders.' This suggests that reentry rituals are alternatives in the sense of providing state-sanctioned countervailing force that allows offenders to regain symbolic ground which was taken from them by the state and thus may only be replaced by the state. Reentry rituals restore symmetry to the arc of an offender's fall and recovery.

However, one might point out that in some crimes, no real symmetry is possible. Some crimes are irreparable, and so a full restoration of an offender's standing in society is impossible. As Sayre-McCord points out, there are crimes in which “neither hope nor sense should be attached to the idea that the victim might be made whole.”[364] He goes on to say that even though reinstatement in society is impossible, if an offender has gone through adequate effort to work toward amends (although full amends are out of reach), then there may still be space for the state to ritually forgive the offender.

Sometimes the proper response to people who have worked to make amends for a

362 Maruna, 2012, at 78.
363 Maruna, 2012, at 79; citing Herzog-Evans, 2011.
364 Sayre-McCord, 2001, at 519.

> horrible offense is to forgive them... The institutional forgiveness—conveyed, perhaps, by pardon—represents a collective choice by society concerning what should count as sufficient grounds for readmitting someone to full and equal status in society. Such a decision will no doubt have to be sensitive in some way to victims' attitudes.[365]

There are plenty of anecdotal accounts of offenders languishing in prison after they have been fully transformed as individuals. They not only present no further risk to society, but also may have given their lives to the effort of reforming other prisoners and working as much as possible toward their own amends. To insist that these offenders finish their terms of imprisonment simply out of commitment to an original sentence is a lopsided view of the goals for the system. When an offender reaches the standing that the state can reasonably hope for as the final result of its intervention, then the state should be willing to pardon the offender and admit that it has no need for further intervention, even if there is no possibility of full reparation because of the severity of the original offense.

There is little worry about whether the state should coerce reentry rituals, because presumably such rituals require significant work on the offender's part to deserve such recognition. If an offender is unwilling to put in the effort to deserve the award, then there is no reason that a state would coerce him to participate in the award ceremony.

365 Sayre-McCord, 2001, at 520.

Chapter Five: Toward a Minimally Punitive Regime

In this final chapter, I give an overview of the conclusions from the previous chapters, an argument for why a purely non-punitive crime response regime is untenable, and an argument for why the state should punish as little as possible.

5.1 Overview of Conclusions from Chapters 1-4

In the first chapter, I summarized a few of the reasons that theorists, with increasing frequency and forcefulness, have concluded that all the putative justifications of punishment have failed. It is plausible that no justification of punishment has succeeded, and also plausible to suspect that a full, successful justification may never emerge. Also in the first chapter, I gave a positive argument for the necessity of the intention requirement of punishment. This requirement demands that for a state action to be considered punishment, it must be performed with the intention to harm the offender. This requirement has received less attention from scholars than it merits. It is important because it is the key to determining what counts as punitive. Any discussion of alternatives to punishment must begin by distinguishing punishment from alternatives, and justifying that distinction. Every alternative that I have explored relies on its status as non-punitive due in part to the distinction arising from the intention requirement of punishment.

In defending this distinction, I considered the objection that intentions are vague and inscrutable and thus not a reliable guide for distinguishing state actions. I countered this by showing that intentions behind punishment can and should be transparent. The public record of legislation and its interpreters in the judiciary provide adequate basis for identifying these intentions, at least on most occasions. This record also provides adequate reason to trust that state actions taken in the name of these intentions can be scrutinized to make sure that they actually conform to the stated intentions

In the second chapter, I argued against the possibility that pure restitution could entirely replace punishment. My argument relied in part on my claim that some practices, such as incapacitative imprisonment, cannot be justified by appeal to restitution. When an action is so straightforwardly preventative, it would be mistaken to claim that it is justified by repaying the community rather than the clearer goal of quarantining the offender against future attacks. I outlined other worries about the theory of pure restitution, including the objection that it could not adequately account for a murder victim who has no assets, not descendants and no will. I argued that a regime of pure restitution is impossible because in order to actually enforce restitution mandates, the state would need to resort to punishment (except in rare cases such as wage garnishment). I conclude that despite Boonin's impressive defense of his theory of pure restitution, it cannot hold up under scrutiny. I also argued against the widespread use of mandatory restitution, largely due to the fact that most offenders will not be able to pay the restitution.

In the same chapter I go on to argue that the welfare state has an obligation to provide compensation to violent crime victims. This argument hinges on the observation that most property crimes can be protected against with private insurance, while violent crimes most often cannot. While victims also have the option to pursue compensation via civil suits, the vast majority chooses against this route, mostly because people understand that the offender will not be able to afford the compensation. If the state rewarded primary compensation for crime victims, it would not preclude victims who choose to pursue a tort. Research from Ontario suggests that this system can work well, and that many victims do not place emphasis in formal fault-finding or restitution paid from the offender. Many would rather be reliably compensated by the state than to further confront the offender.[366] I also argue that the state has a positive obligation to provide compensation. I give three justifications for this obligation: first that this compensation is an essential public good; next that the state failed its duty to protect, and owes compensation for that failure; and finally that compensation is

366 Felduthsen, 2010, at 113.

a unique and essential way for the state to express certain values and commitments after the crime, namely to express sympathy, remorse and willingness to take responsibility for redressing harm.

In the third chapter, I defended a definition of offender rehabilitation that relies on a minimal notion of consent. Rehabilitation is authorized, retributive, consensual, intended treatment. I argued that mandatory rehabilitation will not be able to remain purely non-punitive because in order to enforce the mandate the state will need recourse to punishment. I also drew the conceptual borders between rehabilitation, punishment and incapacitation. I advanced a definition for offender incapacitation as authorized, retributive, non-consensual, non-punitive intended restriction. The state may restrict an offender physically or legally. Preventative incarceration is the most common example of physical incapacitation. An example of legal incapacitation is making it illegal to refuse a rehabilitation program. If the offender may still choose to refuse, despite heavy coercion, then he has enough volition in the matter to qualify under the minimal sense of consent required for rehabilitation. The benefit of these conceptual distinctions is being able to accurately explain and scrutinize the state's discrete actions that comprise a complex intervention such as chemical castration. In the latter part of that chapter I argued that rehabilitation and incapacitation may both be justified by appealing to the benefits to society.

In the fourth chapter, I showed why restorative justice is too limited to deal with many cases, let alone replace punishment altogether. It should be regarded as a valuable but restricted alternative. Given public opinion surveys, we can surmise the intentions of average conference participants to determine whether a conference will result with a punitive sentence. RJ conferences that result in prison sentences are most likely punitive, but ones that do not result in prison are much more likely to be non-punitive. But, since most conferences are for non-serious offenders and hence do not result in prison, they should still be regarded as an alternative to punishment. I also argued that like restorative justice conferences, apology and reentry ceremonies should not be coerced.

5.2: The Case Against A Purely Non-Punitive Regime

Together, these conclusions can be brought together to make larger claims about crime response. We cannot have a post-arrest regime that relies solely on pure restitution, nor one that relies solely on rehabilitation, nor one that relies solely on incapacitation, nor one that relies solely on restorative justice and other rituals. A mix of all of these, without any recourse to punishment, will not suffice either. The candidates that do the heavy lifting for replacing punishment, that is to say, the ones that come closest to being viable pure alternatives, are mandatory restitution, mandatory rehabilitation, and incapacitation. But it turns out that mandatory restitution and mandatory rehabilitation cannot function as purely non-punitive, because the state will require punitive bite in order to enforce their mandates. And incapacitation is only justified for dangerous or incorrigible offenders. If there is no viable candidate to replace punishment, then there seems to be no way to have a purely non-punitive regime.

Perhaps a mix of all of these alternatives would somehow suffice to get by without punitive interventions. For this to be the case, there would need to be some reason to think that the shortcomings of mandatory restitution and mandatory rehabilitation can be overcome by the addition of one or more of the alternatives. But the practical problem that leads to punitive recourse is this: how may a state enforce its mandates to comply with restitution or rehabilitation orders when it deals with stubborn and recalcitrant offenders? If one of the alternatives offers a solution to this problem, then it is possible that a mix of alternatives could yield a non-punitive regime. But no alternative can provide a solution, leaving us to conclude that a non-punitive regime is out of reach.

Of all the candidates, only incapacitation seems like a viable contender to fill the role of dealing with offenders who refuse participation in restitution or rehabilitation. An advocate of this theory would claim that if an offender refuses to participate, then the state may deem him dangerous enough to keep away from society, at least until he comes around to cooperating. But this is an indefensible theory because the state would not be justified in incapacitating many of the recalcitrant offenders.

Take a case of simple assault, say during a drunken bar fight. If one of the brawlers refused to participate in any restitution or rehabilitation (perhaps paying for broken bones or property, going to anger management classes or alcohol treatment), would the state be justified in locking him away as a preventative measure? The question of who is dangerous enough to merit preventatively detaining is surely a difficult one, but one need not be an expert to know that the offender's behavior that night at the bar certainly does not pose a great enough risk to merit his preventative detention.

One reply to this response is to claim that I have picked an example of a case that does not merit an enforceable mandate in the first place. If the offender is not enough of a danger to society to preventatively detain, then perhaps his actions do not merit the state's coercing him with interventions. There certainly are cases when offenses are unserious enough that the state should refrain from coercive action. However, there are a great many cases of offenses that are serious enough to require certain actions from the offender to make amends, but not so serious to merit preventative detention. If the offender is unwilling to meet the state's requirements, then the state needs to be able to enforce these requirements. This is because nobody would see them as actual requirements if people were free to ignore them. The bar fight is an example of this kind of case because it resulted in serious harm that demands amends from the assailant. The state is justified in coercing these actions, but if he is unwilling to follow the state's mandate, then the state needs to have a way to enforce the mandate, or else nobody would have reason to believe that the mandates are indeed required.

5.3: The Case for a Minimally Punitive Regime

Given the uncertainty over the moral status of punishment, we should only punish as much as we can be confident of justifying. The urgent need to justify our practice, which I described in the first chapter, forces us to err on the safe side of justification. To give a brief reminder of how pressing this need is, consider how every action the state takes that harms its citizens should be justified. Take for

example the numerous institutional and procedural safeguards that states take in order to justify practices such as eminent domain. Citizens may sue to keep their property, and courts then must rule on whether the state's justifications for seizing the land are adequate. An example of the extremely painstaking and careful deliberation that may undergird a contentious eminent domain decision is the US Supreme Court case *Kelo v. City of New London*.[367] Eminent domain is just one of many actions the state takes that harms its citizens. Taxation is another example, and the heated debates it inspires in legislatures throughout the world attests to the seriousness with which people treat its harms. For all the disputes these actions inspire, they have a crucial advantage over punishment when it comes to justification. Their primary intention is to benefit society, so their merit can be debated in terms of cost and benefit. The harms are explicitly regarded as lamentable outcomes that should be minimized as much as possible given the agreed course of action to benefit society. Punishment not only inflicts harm as serious (and usually much more serious) as these other state actions, but the nature of inflicting intended harm exacerbates the problem of justifying it. I can think of no other state action that undertakes to harm citizens with the intended outcome of precisely the harm itself. As Gross puts it succinctly, "[c]riminal punishment is by far the worst thing that the law allows."[368]

In addition to this, as I hope I've shown, there is no clear consensus on whether the institution of punishment can be justified. As Tunick puts it, "legal punishment is essentially contested."[369] As I pointed out in the first chapter, even doyens of this corner of philosophy whose work has heretofore assumed that punishment can be justified, have recently admitted that the search for a justification has been unsuccessful and may be futile.[370] Examples abound of other practices throughout history in which a practice that involved intentionally harming other people once had wide respectability, then were contested, then were abolished or minimized as much as possible. Punishment seems like a prime

367 545 U.S. 469 (2005).
368 Gross, 2012, at 7.
369 Tunick, 1992, at 186.
370 See Garvey, 2011, at 498; and Dolinko, 2011, at 427 and 428.

candidate for such a practice, and it seems clear that we should move with haste to be on the safe side of justification, and history. An objection could make the claim that we do not have an obligation to minimize punishment until we know for certain that there is no way to justify punishment *in toto*. But this claim is indefensible because although we may never achieve such certainty, we need not achieve it to know that we should minimize punishment as soon as possible. Given the enormous scale of suffering involved, we should act with parsimony pending further justification. Imagine if the world's governments held out from officially abolishing slavery for another two to three decades in order to be more certain that it is morally wrong. Those additional millions of hours of human suffering would have been the result of moral reluctance to act quickly in order to be 'safe rather than sorry.' Because it is at least possible that most punishment is unjustified, it seems unconscionable to take the same foot-dragging approach to minimizing punishment. While unlike slavery, we can be sure that at least some punishment can be justified, but punishment resembles slavery in inflicting suffering on millions of people. And like slavery was once contested, the practice of punishment is also contested.

We should therefore punish as little as we need (which is as much as we could confidently justify) until we have relative consensus that a different level of punishment can be justified. This is the concept of parsimony. Von Hirsch even admits that if the requirements of a criminal justice system "can be performed without having to resort to the coercive and unpleasant features of punishment" then the institution of punishment "would lose its *raison d'etre*."[371] How can we be confident of the level of punishment that we may justify out of necessity? *The only method is to build a thorough understanding of the non-punitive practices that a post-arrest regime may rely on to fulfill the requirements of a criminal justice system, and to clarify when the use of these alternatives is justified, and when punishment will be required in their stead.* This dissertation is an attempt to head down this road. To remind the reader, three of the basic requirements any criminal justice system must fulfill are to maintain safety, communicate condemnation of offenses, and redress harm to victims. In this section I

371 Von Hirsch, 1999, at 79.

argue that a minimally punitive regime that relies on the alternatives I have explored in previous chapters can fulfill these requirements.

To help demonstrate this, I will imagine how two different offenders might travel through a minimally punitive regime. Of course, this is a speculative enterprise, but I hope that I can at least demonstrate the following: first that such a regime would result in serious changes in practices in order to minimize suffering, and secondly that the differences are not too big to merit panic over deterrence. After this brief speculation, I will address deterrence in a separate discussion.

First, let's take a low level offender, Mike, who is convicted of a nuisance crime that I will label 'drunken and disorderly conduct,' as well as vandalism and destruction of property. The police found Mike in a stupor and looking for his keys on the ground, but only in the circles illuminated by the streetlights. He had been cutting off the heads of parking meters and broke a shop window. In a sentencing hearing, the judge learned that Mike has recurrent alcohol problems. The Judge ordered Mike to get an assessment by an addiction treatment practitioner, and in the meantime to also work 30 hours for the city's 'Blight Reclamation Street Team,' boarding up broken windows and cleaning streets. He also is told that he has the option to write an apology to be published in a local weekly paper. He has the option of including his name in the piece, as is regulated by a piece of legislation guiding public displays of apology. The sentencing order cites a law that restricts the hours of community service that can be given for certain offenses, and directs the service to be related to the community that was victimized as much as possible. The judge gives Mike one month to meet with the practitioner and begin the first five hours of the service. If he does not meet these requirements, he will be immediately put in the county jail for five days. The county jail is a medium security facility made specifically to hold offenders with short sentences. It is clean and safe. Mike meets his requirements, returns to the judge, and the judge orders him to attend ten meetings of Alcoholics Anonymous on the recommendation of the practitioner. There is no penalty for not attending, and the city gives him automatic phone call reminders for the meeting times, as well as a promise of a travel reimbursement

if he completes all ten sessions.

Now let's consider a terrible offender, Zeb, who is convicted of raping and murdering his neighbor. Zeb was held in custody before and during his trial. The judge publishes a formal condemnation in the paper, and before the trial the victim's family was compensated by the state's violent crime victim fund, with access to counselors, school loan forgiveness, paid leave and a direct deposit of money. The sentencing guidelines order a risk assessment, which deems Zeb dangerous enough to incapacitate for two years before another assessment is completed. He is sentenced to complete an intensive regimen of rehabilitation while detained. The incapacitation is very secure (i.e. escape-proof), but inside it is also clean, safe and comfortable. If he gets far enough behind on his commitments to the treatment requirements, he is punished by being put into isolation for twelve hours per day for a week. After his first two years, he is given different, more lenient requirements for rehabilitation, and is allowed to begin working 30 hours a week for the incapacitation center's kitchen. He has access to conjugal visits, and education. After 12 years in the incapacitation center, he begins a transition to the community, living in a half way house where he has to check in each evening at 6pm. Certain restrictions on his behavior mean that he can be punished for short period of times in the jail. Eventually he makes the full transition back to the community.

While this is a small, speculative exercise, I intended to suggest that a minimally punitive regime could deal with all levels of offenders and situations without any problem meeting the requirements of a criminal justice system. Of course, how such a regime would actually function is beyond the scope of this essay, and I do not presume that the above recommendations would be appropriate—surely experts in criminology and treatment would guide how such a regime would function. But by condemning the crimes, redressing harm to the victims, and keeping the community safe, both the minimally punitive sentences I described indicate that it is at least possible that a larger system could also meet these requirements. We have no reason to resist moving in this direction through phasing out the unnecessarily punitive aspects of the current regime.

But critics will doubtless say that there is reason to resist this change. The biggest worry that most critics would have is whether such a regime would adequately deter crime. But as the research on deterrence shows, there is no good reason to believe that additional severity sentences results in any additional general deterrence.[372] Tonry writes that "the clear weight of the evidence for more than thirty years is that harsh punishments, compared with lesser punishments, have few if any deterrent effects."[373] Indeed, analysts have noted that harsh sentencing policies are frequently adopted because they serve politicians as a way to communicate to the public that "*something* is being done—rather than because policy makers believe they will have any effects on crime."[374] And public officials would have no reason to believe that making sentences harsher adds general deterrence, because studies have repeatedly shown that legislating harsher sentencing policies, even when they are publicized, have no impact "on short or long-term crime rates."[375] Of course, in the complete absence of a criminal justice system, crime rates soar, as historical evidence of police strikes indicate.[376] But what a critic would need to show is evidence for why minimizing punishment, while maintaining the structures of police and courts, would result in any crime increase. There is good reason to think that money saved by minimizing punishment could go to hiring extra police, and that this would have an overall increase in the level of general deterrence.

Additionally, what we know about deterrence indicates that specific deterrence works better than general deterrence. Specific deterrence occurs when an offender knows that if he does a particular action, he will most likely be punished for it. This is the kind of deterrence that is achieved at drug courts, with impressive success.[377] David Kennedy even demonstrates that this kind of deterrence can be achieved by police *before* an arrest: it turns out that the best way to clear away an illegal drug market is to gather all the evidence needed to convict each of the key sellers, then bring them into a

372 See e.g. von Hirsch et al., 2009, at 39, Bottoms, 2004, and Lewis, 2004, at 51, and Lippke, 2007, at 42-43.
373 Tonry, 2011b, at 155.
374 Tonry, 2011a, at 5.
375 Doob and Webster, 2009, at 51.
376 Von Hirsch et al., 2009, at 46.
377 See e.g. Robinson and Dignan, 2004, at 331-332.

room with their parents and other family, show them the evidence, and tell them that if they are seen selling once more then they will be immediately arrested and tried, and to tell all the other sellers that the neighborhood is now off limits for selling. As if by magic, the street corner is cleared of drug dealing.[378] Mark Kleiman has written a thorough, impressively researched book detailing specific policy proposals that decrease crime called *When Brute Force Fails: How to Have Less Crime and Less Punishment*. Unsurprisingly, many of his suggestions focus on creating a system that relies on specific deterrence. As he writes, "uncertainty, delay, and inaccurate perception all limit the effectiveness of deterrent threats," and therefore policy makers should increase ways of making outcomes more certain in the minds of offenders.[379] A minimally punitive regime could take advantage of the effectiveness of specific deterrence, and abandon the ineffective, harsh sentencing policies aimed at general deterrence. With all of this evidence that deterrence would not be a problem for a minimally punitive regime, the onus is on the critic to show otherwise.

The final objection is that a minimally punitive regime would look so similar to a punitive one that the transition would effectively be in name only. Especially after reading my argument that practices such as coerced rehabilitation and preventative detention are non-punitive, one might wonder if the shift in emphasis to avoiding *intended* harms would result in negligible differences in practical outcomes. If the state is still coercing offenders to join rehabilitation programs that are segregated from the rest of the population, or incapacitating offenders deemed too dangerous to remain in society by imprisoning them, then one may wonder whether the resultant regime would look very different from current punitive models. If not, then this exercise in shifting emphasis would be a load of hot air without practical payout sufficient to justify its implementation. However, there is very good reason to believe that the opposite is true: that the difference would be so stark that the main practical impediment would be convincing the public to endorse such drastic changes.

378 Kennedy, 2010.
379 Kleiman, 2009, at 95.

Here is short list of current practices that would likely be eliminated as a result of this switch: the death penalty, life sentences, felony disenfranchisement, solitary confinement, restriction from labor, mandatory restitution or fines imposed on indigents, imprisonment for drug possession, mandatory minimum sentences and three-strikes laws. Inn addition, the use of home detention curfew would significantly rise,[380] and there would be far-reaching changes to prison conditions that would align with the recommendations made in Lippke's (2007) *Rethinking Imprisonment*. Tonry details particular policy steps to take in order to do less harm, which, along with radically decreasing the use of prison and abolishing punitive laws such as mandatory minimums, he also proposes the "creation of emergency mechanisms for shortening sentences of many currently imprisoned offenders; and reconstitution of permanent institutions to oversee release of offenders whose continued confinement serves no valid public purpose."[381]

Portugal and Norway provide glimpses of this minimally punitive future. In the nomenclature currently used by the drug policy writers, Portugal has 'de-criminalized' all drugs. It has not legalized them, but simply stopped punishing people for possessing them.

> That doesn't mean drugs are legal in Portugal. When caught, people are summoned before an administrative body called the Commission for the Dissuasion of Drug Addiction. Each panel consists of three members—usually a lawyer or a judge, a doctor, and a psychologist or a social worker. The commissioners have three options: recommend treatment, levy a small fine, or do nothing. In most respects, the law seems to have worked: serious drug use is down significantly, particularly among young people.[382]

Portugal still punishes drug dealers, but it has begun treating drug possession laws as simply one component of a public health crisis intervention. It even provides clean drug paraphernalia to drug addicts in an effort to cut down on disease.

Norway, famously, has perhaps the least punitive penal system in the world. The Scandinavian

380 See Mair, 2004, at 233-240.
381 Tonry, 2011b, at 146.
382 Spector, 2011, at 36.

penal systems are generally regarded as anomalously liberal,[383] and among them Norway stands out. Its most publicized differences from other, more punitive regimes are its open prisons and its short prison sentence lengths. Bastoy prison is on an island four kilometers from the mainland. Prisoners live in plushly furnished cottages and work outdoors, unsupervised, during the day. There is almost no guard presence on the island, and prisoners are mostly free to do as they please, even go swimming, despite the fact that the mainland is close enough to swim to. There are no fences. The prison is only open to serious offenders who have shown a maturity and willingness to change in other, higher security prisons. A chainsaw murderer was even allowed to work in the forest with a chainsaw. Bastoy is one of Norway's 'open prisons' and news stories covering it from flabbergasted US or British media with a tone of bewildered outrage are prevalent. Another oddity of Norway's system to outsiders is its maximum prison sentence of 21 years. Anders Breivek, who killed 69 people in 2011, is the focus of considerable attention, and his case has highlighted Norway's maximum sentence. After 21 years, if the authorities deem that Breivek is still a danger to society, it may extend the term again. Thus it is likely that if Breivek continues his hateful and far-right militant speech and thinking, he will stay in jail for the duration of his life. Still, the sentence also allows him to possibly be paroled after 10 years. The difference between this sentence and sentences he would receive elsewhere highlights how stark the contrast would be between punitive regimes, and minimally punitive regimes. It is undoubtable that if he were convicted in one of the 29 states in the USA where the death penalty is legal, he would be put to death. Norway has arguably made most of the transition to becoming minimally punitive. The fact that it achieved this transition is further reason to think that it is possible. Notably, it has a remarkably low rate of crime and recidivism. While the average recidivism rate of Norwegian prisoners is around 20% (between one half and one third of what the rate seen in countries like the US and UK), the rate of recidivism from prisoners leaving Bastoy is a remarkable 16%.[384]

383 See Pratt, 2007.
384 Sutter, 2012.

It may be the case that the rest of the world never catches up to Scandinavian implementation of criminal justice, but we can at least hope to head that way. In order to turn that direction, theorists should be leading, and there has been too little done to scrutinize alternatives to punishment. This dissertation has been an attempt to add to that effort.

Works Cited

Abel, Charles F. and Marsh, Frank. H. (1984). *Punishment and Restitution.* Greenwood Press, Connecticut.

ACLU. (2010). "In for a Penny: The Rise of America's New Debtors' Prisons." http://www.aclu.org/prisoners-rights-racial-justice/penny-rise-americas-new-debtors-prisons

Allen, Francis. (2011). "Legal Values and the Rehabilitative Ideal." In: *Why Punish? How Much?* Ed. Michael Tonry. Oxford University Press.

Alexander, Michelle. (2012). *The New Jim Crow: Mass Incarceration in the Age of Color Blindness.* The Revised Edition. The New Press, New York, NY.

Alvidrez, I. J., Shumway, M., Boccellari, A., Green, J. D., Psyd, V. K., & Merrill, G. (2008). "Reduction of State Victim Compensation Disparities in Disadvantaged Crime Victims Through Active Outreach and Assistance: A Randomized Trial." *American Journal of Public Heaith*, *98*(5), 882–889.

Banks, Russell. (2011). *Lost Memory of Skin*. Random House.

BBC News South East Wales. (2012). "Rapist Christopher Hughes asks for chemical castration." Retrieved from: http://www.bbc.co.uk/news/uk-wales-south-east-wales-19583439

Bearden v. Georgia, 461 U.S. 660, 662-63, 668-69 (1983).

Beckett, K., & Murakawa, N. (2012). "Mapping the shadow carceral state: Toward an institutionally capacious approach to punishment." *Theoretical Criminology*, *16*(2), 221–244.

Benatar, David. (2013). Private comments.

Bergseth, Kathleen J. and Bouffard, Jeffrey A. (2007). "The long-term impact of restorative justice programming for juvenile offenders." *Journal of Criminal Justice*. 35(4): 433-451

Bleyer, J. (2012). "How Can We Stop Pedophiles? Stop treating them like monsters." *Slate*. Posted Monday, Sept. 24, 2012. http://www.slate.com/articles/health_and_science/medical_examiner/2012/09/stop_childhood_sexual_abuse_how_to_treat_pedophilia_.single.html

Bonta, J., & Andrews, D. A. (2007). Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation. *Rehabilitation.*

Boonin, David. (2008). *The Problem of Punishment*. Cambridge University Press, New York, NY.

Braithwaite, John. (1989). *Crime, Shame and Reintegration*. Cambridge University Press, Cambridge, UK.

(2003). “Principles of Restorative Justice,” in A. von Hirsch, A. Roberts and A. Bottoms (eds) *RJ and Criminal Justice*, 1-20. Oxford: Hart Publishing.

Braithwaite, John and Strang, Heather. (2001). “Introduction: Restorative Justice and Civil Society,” in Strang and Braithwaite, eds. *Restorative Justice and Civil Society*. Cambridge University Press, Cambridge, UK.

Buck, Katharina. (2005). “State Compensation to Crime Victims and the Principle of Social Solidarity.” *European Journal of Crime, Criminal Law and Criminal Justice*, *13*(2), 148–178.

Calhoun, A., & Pelech, W. (2010). “Responding to young people responsible for harm: a comparative study of restorative and conventional approaches.” *Contemporary Justice Review*, *13*(3), 287–306.

Cholbi, M. (2010). “Compulsory Victim Restitution Is Punishment: A Reply to Boonin.” *Reason*, *2*(1), 85–93.

Cobley, C. (1998). Financial compensation for victims of child abuse. *Journal of Social Welfare and Family Law*, *20*(3), 221–235.

Corlett, J. A. (2003). Making More Sense of Retributivism: Desert as Responsibility and Proportionality. *Philosophy*, *78*(02), 279–287.

Cullen, Francis T. and Jonson, Cheryl Lero. (2012). *Correctional Theory: Context and Consequences*. SAGE Publications.

Cullen, Francis T. and Smith, Paula. (2011). “Treatment and rehabilitation.” In: ed. Michael Tonry, *The Oxford Handbook of Crime and Criminal Justice*. Oxford University Press.

Curtis, Polly. (2012). “Should sex offenders be chemically 'castrated'?” *The Guardian News Blog*. Posted Tuesday 13 March 2012. http://www.guardian.co.uk/politics/reality-check-with-polly-curtis/2012/mar/13/prisons-and-probation-criminal-justice

Dagger, R. (1991). “Restitution: Pure or Punitive?” *Criminal Justice Ethics*, 1–19.

Daly, Kathleen (2003). “Mind the Gap: Restorative Justice in Theory and Practice.” In, Andrew von Hirsch, et. al., eds., *Restorative Justice and Criminal Justice: Competing or Reconcilable Paradigms?* Oxford and Portland, Orgeon: Hart Publishing. 219-236.

(2006). “The Limits of Restorative Justice.” Prepared for Dennis Sullivan and Larry Tifft (eds.) *Handbook of Restorative Justice: A Global Perspective*. New York: Routledge.

Davis, M. (2008). “Punishment Theory’s Golden Half Century: A Survey of Developments from (about) 1957 to 2007.” *The Journal of Ethics*, *13*(1), 73–100.

Davis, Angela Y. (2003). *Are Prisons Obsolete?* Seven Stories Press, New York, NY.

Day, A., Ward, T., & Shirley, L. (2011). “Reintegration Services for Long-Term Dangerous Offenders: A Case Study and Discussion.” *Journal of Offender Rehabilitation*, *50*(2), 66–80.

De Keijser, J. (2011). "Never Mind the Pain, It's a Measure! Justifuying Measures as Part of the Dutch Bifurcated System of Sanctions." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

DeVore, C. (2012). "Symposium: Preventive Detention." *The Journal Of Criminal Law & Criminology*, *1*, 0–2.

Dickman, M. (2009). "Should Crime Pay?: A Critical Assessment of the Mandatory Victims Restitution Act of 1996." *California Law Review*, *49*(1982), 1687–1719.

Dolinko, David. (2011). "Punishment." In: *The Oxford Handbook of Philosophy of Criminal Law.* Eds. John Deigh and David Dolinko. Oxford University Press.

Doob, AN, and Webster, CM. (1999). "Studies of the Impact of New Harsh Sentencing Regimes." In: *Principled Sentencing: Readings on Theory and Policy*. ed. von Hirsch, Ashworth and Roberts. Hart Publishing

Duff, R.A, (1999). "Punishment, Communication and Community." In: Punishment and Political Theory, ed. Matt Matravers. Hart Publishing.

(2001). *Punishment, Communication, and Community*. Oxford University Press.

(2008). "Legal Punishment", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/legal-punishment/>.

(2011a). "Penance, Punishment, and the Limits of Community." In: *Why Punish? How Much?* Ed. Michael Tonry. Oxford University Press.

(2011b). "Responsibility, Restoration, and Retribution." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

Ellin, J. (1994). "Restitution Defended." *Ethics*, *4*(July 1978), 1–10.

Feinberg, Joel. (2011). "The Expressive Function of Punshment." In: *Why Punish? How Much?* Ed. Michael Tonry. Oxford University Press.

Feldthusen, B., Hankivsky, O., & Greaves, L. (2000). "Therapeutic Consequences Of Civil Actions For Damages And Compensation Claims By Victims Of Sexual Abuse." *Can. J. Women & L.*

Federal Bureau of Investigation, 2004 and 2011, *Crime in the United States, Clearances.* Retrieved from: http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/clearances

Foucault, Michel. (1995). *Discipline and Punish*. Vintage, 2nd Edition.

Galaway, B. (1979). "Differences in Victim Compensation and Restitution." *Social Work*, 57–58.

Garvey, Stephen P. (2011). "Alternatives to Punishment." In: *The Oxford Handbook of Philosophy of Criminal Law.* Eds. John Deigh and David Dolinko. Oxford University Press.

Gavrielides, T. (2008). "Restorative Justice—The Perplexing Concept: Conceptual Fault-Lines And Power Battles Within The Restorative Justice Movement." *Criminology and Criminal Justice*, *8(2)*, 165–183.

Glaser, B. (2009). "Treaters Or Punishers? The Ethical Role Of Mental Health Clinicians In Sex Offender Programs." *Aggression and Violent Behavior*, *14*(4), 248–255.

(2010). "Sex Offender Programmes: New Technology Coping With Old Ethics." *Journal of Sexual Aggression*, *16*(3), 261–275.

Golash, Deirdre. (2005). *The Case Against Punishment: Retribution, Crime Prevention, And The Law*. NYU Press.

Gopnik, Adam. (2012). "The Caging of America: Why do we lock so many people up?" *The New Yorker*, January 30, 2012.

Grann, David. (2009). "Trial by Fire: Did Texas Execute and Innocent Man?" *The New Yorker*. September 7, 2009.

Gromet, D. M., & Darley, J. M. (2006). "Restoration and Retribution: How Including Retributive Components Affects the Acceptability of Restorative Justice Procedures." *Social Justice Research*, *19*(4), 395–433.

Gross, Hyman. (2012). *Crime and Punishment: A Concise Moral Critique*. Oxford University Press.

Hajdin, M. (1987). "Criminals as Gamblers: A Modified Theory of Pure Restitution." *Dialogue*, *26*(01), 77–86.

Hanna, N. (2009). "The Passions Of Punishment." *Pacific Philosophical Quarterly*, *90*, 232–250.

Hart, H. L. A. (1968). Prolegomenon to the Principles of Punishment. *Proceedings of the Aristotelian Society*, *60*, 1–26.

Hartney, C. and Marchionna, S. (2009). "Attitudes of US Voters toward Nonserious Offenders and Alternatives to Incarceration." *FOCUS: Views from the National Council on Crime and Delinquency*, (June).

Harvard Law Review. (2011). "Recent Cases: Criminal Law —Federal Sentencing — First circuit holds that rehabilitation cannot justify post-revocation imprisonment." *3582*(2).

Holehouse, Mathew. (2012). "Paedophiles chemically castrated in British jail." *The Telegraph*. 13 Mar 2012.

Hollin, C. R., & Palmer, E. J. (2009). "Cognitive Skills Programmes For Offenders." *Psychology, Crime & Law*, *15*(2 & 3), 147–165.

Honderich, Ted. (1969). *Punishment: The Supposed Justifications Revisited.* Penguin Books.

Husak, Douglas. (2011). "Retributivism, Proportionality, and the Challenge of the Drug Court Movement." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

Katzenstein, M. F., & Nagrecha, M. (2011). "A New Punishment Regime." *Criminology & Public Policy*, *10*(3), 555–568.

Kershnar, S. (2010). "Punishment and Torture," in Jesper Ryberg and J. Angelo Corlett, eds., *New Perspectives on the Ethics of Punishment* (Palgrave Macmillan, 2010), 130-148.

(2012). "Does Necessity Justify Punishment? Assessing the Main Threat to David Boonin's Restitution Theory." *26*(2), 71–79.

Kelo v. City of New London. 545 U.S. 469 (2005).

Kennedy, David. (2011). *Don't Shoot: One Man A Street Fellowship, and the End of Violence in Inner-City America*. Bloomsbury USA, New York.

Kleiman, Mark. (2009). *When Brute Force Fails: How to Have Less Crime and Less Punishment*. Princeton University Press.

Klosko, G. (2009). "Cosmopolotanism, Political Obligation, and the Welfare State." *Political Theory*, *37*(2), 243–265.

Kolber, A. J. (2009a). "The Comparative Nature of Punishment." *The Boston University Law Review*. 89: 1565.

(2009b). "The Subjective Experience of Punishment." *Columbia Law Review*. 109: 182.

(2012). "Unintentional Punishment." *Legal Theory*, *18*(01), 1–29.

Kolind, T., Frank, V. A., & Dahl, H. (2010). "Drug Treatment Or Alleviating The Negative Consequences Of Imprisonment? A Critical View Of Prison-Based Drug Treatment In Denmark." *International Journal of Drug Policy*, *21*, 43–48.

Kristoffersen, R., & Hildebrant, S. (2010). *Correctional Statistics of Denmark, Finland, Iceland, Norway and Seweden*. Retrieved from http://brage.bibsys.no/krus/handle/URN:NBN:no-bibsys_brage_19605

Lappi-Seppälä, Tapio. (2011). "Sentencing and Punishment in Finland: The Decline of the Repressive Ideal. In: *Why Punish? How Much?* Ed. Michael Tonry. Oxford University Press.

Laudan, L., & Allen, R. J. (2012). "Deadly Dilemmas III: Some Kind Words For Preventive Detention." *The Journal Of Criminal Law & Criminology*, *101*(3), 781.

Lazar, S. R. M. (2008). "Corrective Justice and the Possibility of Rectification." *Ethical Theory &*

Moral Practice, (January), 355–368.

Lippke, R. (2002). "Toward A Theory Of Prisoners' Rights." *Ratio Juris*, *15*(2), 122–145.

(2007). *Rethinking Imprisonment*. Oxford University Press, New York, NY.

Lipsey, Mark W., Landenberger, Nana A., and Chapman, Gabrielle L. (2008). "Rehabilitation: An Assessment of Theory and Research." In *The Blackwell Companion to Criminology*. Ed. Colin Sumner. John Wiley & Sons, 2008.

Lokanan, M. (2009). "An Open Model For Restorative Justice: Is There Room For Punishment?" *Contemporary Justice Review*, *12*(3), 289–307.

Lynch, Mona. (2000). "Rehabilitation as Rhetoric: The Ideal of Reformation in Contemporary Parole Discourse and Practice." *Punishment & Society*. *2*(1): 40-65.

Martinson, Robert. "What Works? Questions and Answers About Prison Reform." *The Public Interest*. *35*(Spring 1974): 22-54.

Mair, George. (2004). "Diversionary and non-supervisory approaches to dealing with offenders." In: *Alternatives to Prison: Options for an insecure society*, Bottoms, Rex and Robinson eds. Willan Publishing, Devon, UK.

Maruna, S. (2011). "Reentry As A Rite Of Passage." *Punishment & Society*, *13*(1), 3–28.

(2012). "Elements of Successful Desistance Signaling." *Criminology & Public Policy*, *11*(1), 73–86.

Matravers, Matt. (2000). *Justice and Punishment*. Oxford University Press.

McDermott, D. (2001). "The Permissibility of Punishment." *Law and Philosophy*, *29*(February), 403–432.

Mcneill, F. (2009). Probation, Rehabilitation and Reparation. In: 2nd Annual Martin Tansey Memorial Lecture, 2009-05-01, Dublin. Retrieved from: http://strathprints.strath.ac.uk/

Menninger, Karl. (1966). *The Crime of Punishment*. The Viking Press, New York.

Miller, F. G. (1978). "Restitution and Punishment: A Reply to Barnett." *Ethics*, *88*(4), 358–360.

O'Hear, M. (2011). "Drug Treatment Courts as Communicative Punishment." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

Okimoto, T. G., Wenzel, M., & Feather, N. T. (2009). Beyond Retribution: Conceptualizing Restorative Justice and Exploring its Determinants. *Social Justice Research*, *22*(1), 156–180.

Peters, C. M. (2006). "Social Work and Juvenile Probation: Historical Tensions and Contemporary Convergences." *Social Work*, *56*(4), 355–365.

Petrucci, C. J. (2002). "Apology In The Criminal Justice Setting: Evidence For Including Apology As An Additional Component In The Legal System." *Behavioral sciences & the law*, *20*(4), 337–62.

Pollock, Joycelyn M. (2011). *Ethical Dilemmas & Decisions in Criminal Justice*. Wadsworth, Cengage Learning.

Pratt, J. (2007). "Scandinavian Exceptionalism in an Era of Penal Excess: Part I: The Nature and Roots of Scandinavian Exceptionalism." *British Journal of Criminology*, *48*(2), 119–137.

Ponseti, et al. (2011). "Assessment of pedophilia using hemodynamic brain response to sexual stimuli." *Arch Gen Psychiatry.* 2012;69(2):187-194. doi:10.1001/archgenpsychiatry.2011.130. doi:10:1001/jama.2010.920

Prescott, D. S., & Levenson, J. S. (2010). "Sex Offender Treatment Is Not Punishment." *Journal of Sexual Aggression*, *16*(3), 275–286.

Radzik, L. (2003). "Do Wrongdoers Have a Right to Make Amends?" *Social Theory and Practice*, *29*(2), 325–342.

Ramsay, P. (2011). "A Political Theory of Imprisonment for Public Protection." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

Raynor, P., Robinson, G. (2005). *Rehabilitation, Crime and Justice*. Palgrave Macmillan, New York, NY.

Raynor, P. (2004). "Rehabilitative and reintegrative approaches." In: *Alternatives to Prison: Options for an insecure society*, Bottoms, Rex and Robinson eds. Willan Publishing, Devon, UK.

Ristovski, A., & Wertheim, E. H. (2005). "Investigation of compensation source, trait empathy, satisfaction with outcome and forgiveness in the criminal context." *Australian Psychologist*, *40*(1), 63–69.

Robinson, Gwen, and Crow, Ian. (2009). *Offender Rehabilitation: Theory, Research and Practice*. SAGE Publications.

Robinson, Gwen, and Dignan, James. (2004). "Sentence Management." In: *Alternatives to Prison: Options for an insecure society*, Bottoms, Rex and Robinson eds. Willan Publishing, Devon, UK.

Rotman, E. (1990). *Beyond Punishment*. Greenwood Press, New York, NY.

Santos, Fernanda. (2013). "Advances in Science of Fire Free a Convict After 42 Years." *The New York Times*. April 2, 2013.

Sayre-Mccord, G. (2001). "Criminal Justice and Legal Reparations as an Alternative to Punishment." *Social, Political and Legal Philosophy*.

Sandel, Michael J. (2012). *What Money Can't Buy: The Moral Limits of Markets*. Farrar, Straus and Giroux.

Schall v Martin, 467 U.S. 253 (1984), at 269-270.

Sherman, Lawrence W and Strang, Heather (2007). *Restorative Justice: The Evidence.* London: The Smith Institute.

Spalding, Larry. (2008). "Florida's 1997 Chemical Castration Law: A Return To The Dark Ages." *Florida State University Law Review*.

Skipp, C., Campo-Flores, A. (2009). "A Bridge Too Far." *Newsweek*. Retrieved from: http://content-cache.knowledgeplex.org/ksg/cache/assets/2801/3222/3222118.html

Stuntz, William J. (2011). *The Collapse of American Criminal Justice*. Belknap Press of Harvard University Press.

Sutter, John D. (2012). "Welcome to the World's Nicest Prison." *CNN*. Posted Thu May 24, 2012. http://www.cnn.com/2012/05/24/world/europe/norway-prison-bastoy-nicest

Tasioulas, John. (2006). "Punishment and Repentence." *Philosophy*. *81*, 271-322.

Tonry, M. (2011a). "Can Twenty-first Century Punishment Policies Be Justified in Principle." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

(2011b). *Punishing Race*. Oxford University Press.

(2011c). "Proportionality, Parsimony and Interchangeability of Punishments." *Why Punish? How Much?* Ed. Michael Tonry. Oxford University Press.

Travis, J. (2005). *But They All Come Back: Facing the Challenges of Prisoner Reentry*. The Urban Institute Press, Washington, DC.

Tsoudis, O. (2000). "The Likelihood of Victim Restitution in Mock Cases: Are the "Rules of the Game" Different from Prison and Probation?" *Social Behavior and Personality*, *28*(5), 481–498.

Tunick, M. (1992). *Punishment*. University of California Press. Berkeley and Los Angeles.

Umbreit, M. S., Coates, R. B., & Vos, B. (2004). "Restorative Justice versus Community Justice: Clarifying a Muddle or Generating Confusion?" *Contemporary Justice Review*, *7*(1), 81–89.

US Code: The Code of the Laws of the US 18 U.S.C. § 3582(a) (2006).

United States v. Salerno 481 U.S. 739 (1987).

Von Hirsch, A. (1993). *Censure and Sanctions*. Oxford University Press.

(1999). "Punishment, Penance, and the State: A Reply to Duff." In: Punishment and Political Theory, ed. Matt Matravers. Hart Publishing.

(2011). "Punishment Futures: The Desert-model Debate and the Importance of the Criminal Law

Context." In: Retributivism Has a Past? Has it a Future? ed. Michael Tonry, Oxford University Press.

Von Hirsch, A., Ashworth, A. Roberts, J. (1999). "Deterrence." In: *Principled Sentencing: Readings on Theory and Policy*. ed. von Hirsch, Ashworth and Roberts. Hart Publishing

von Hirsch, Andrew and Shearing, Clifford and Ashworth, Andrew (2003). *Specifying Aims and Limits for Restorative Justice: A 'Making Amends' Model?* In, Andrew von Hirsch, et. al., eds., *Restorative Justice and Criminal Justice: Competing or Reconcilable Paradigms?* Oxford and Portland, Oregon: Hart Publishing. 21-42.

Walgrave, Lode (2003). *Repositioning Restorative Justice*. Devon, UK: Willan Publishing

Ward, T. and Maruna, S. (2007). *Rehabilitation*. Routledge, New York, NY.

Ward, T., & Salmon, K. (2009). "Aggression and Violent Behavior The ethics of punishment : Correctional practice implications." *Aggression and Violent Behavior*, *14*(4), 239–247.

Weingarten, Gene. (2009). "Fatal Distraction: Forgetting a Child in the Backseat of a Car is a Horrifying Mistake. Is it a Crime?" *The Washington Post*. Sunday, March 8, 2009.

Wenzel, M., Okimoto, T. G., Feather, N. T., & Platow, M. J. (2008). "Retributive And Restorative Justice." *Law And Human Behavior*, *32*(5), 375–89.

Wood, D. (2010). "Punishment: The Future." *Philosophy Compass*, *5*(488).

Zaibert, L. (2010). "Punishment, Restitution, and the Marvelous Method of Directing the Intention." *Criminal Justice Ethics*, *29*(1), 41–53.

Zehr, Howard. (2002). *The Little Book of Restorative Justice*. Good Books.

Zimmerman, Michael J. (2011). *The Immorality of Punishment*. Broadview Press.