

## Claremont Colleges Scholarship @ Claremont

---

Pomona Senior Theses

Pomona Student Scholarship

---

2016

# Defending the Social Good Theory of Punishment

Sydney R. Scott  
*Pomona College*

---

### Recommended Citation

Scott, Sydney R., "Defending the Social Good Theory of Punishment" (2016). *Pomona Senior Theses*. Paper 164.  
[http://scholarship.claremont.edu/pomona\\_theses/164](http://scholarship.claremont.edu/pomona_theses/164)

This Open Access Senior Thesis is brought to you for free and open access by the Pomona Student Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in Pomona Senior Theses by an authorized administrator of Scholarship @ Claremont. For more information, please contact [scholarship@cuc.claremont.edu](mailto:scholarship@cuc.claremont.edu).



SENIOR EXERCISE IN PHILOSOPHY

**Defending the Social Good Theory of Punishment**

*Author:*

Sydney Scott

*Advisor:*

Dr. Michael Green

Submitted to Pomona College in Partial Fulfillment of the  
Degree of Bachelor of Arts

April 25<sup>th</sup>, 2016

## 1. INTRODUCTION

How do we justify the institution of punishment? Most philosophers have attempted to justify punishment either as utilitarian (punishment deters crime and incapacitates criminals from inflicting further damage on society), as retributivist (criminals deserve to have done to them the same wrong that they have done to others), or as some combination of the two. However, alternate theories do have their appeal. Herbert Morris (1981) and Jean Hampton (1984) are the two most well-known proponents of a theory I have previously called the Moral Good Theory, or the MGT (Scott, n.d.). This theory, generally speaking, attempts to justify punishment on the grounds that it is a benefit to the criminal being punished. Hampton's argument for the justification may be stated as follows:

1. We can never harm people, and can only justify institutions that do no harm.
2. Punishment is not a harm to offenders if it can be shown to confer a benefit on them.
3. Punishment does benefit the offender.
4. Thus, punishment can be justified.

We accept (1) for the sake of the theory. Hampton in particular is motivated by the idea that voluntarily harming others is always wrong; if we harm criminals for no other reason than the fact that they have also caused harm, we end up no better than criminals ourselves. (2) is true by virtue of the definitions of "harm" and "benefit." It is possible that something can both harm and benefit a person in different senses; for our purposes, I argue that punishment is not a harm if it can be shown to have a *net* benefit. The conclusion (4) follows from the premises. Thus, the only premise in need of defense is (3).

Hampton argues that punishment is beneficial because it has an educational function. She justifies punishment "insofar as it can teach both wrongdoers and the public at large the moral

reasons for choosing not to perform an offense” (Hampton, 1984). This education allegedly benefits the offender, something “designed to achieve a goal that includes their own moral well-being” (ibid). To illustrate this point, consider Morris’s variation of the MGT, which he calls the Paternalistic Theory: the state should punish criminals in the same vein that parents punish children, with the goal of making the latter better people (Morris, 1981).

Moral good theorists also believe that punishment has a communicative function, and that in communicating with an offender that he has done something wrong, punishment gives him a moral education. Antony Duff (1986) states that criminals have separated themselves from the moral good in committing a crime, and in punishing them, we express our concern for their moral well-being and allow them to be reconciled both to the good and to the community. As such, punishment is beneficial.

Punishment through this lens of communication and education thus serves to protect a wrongdoer’s autonomy, according to most MGT advocates. Through hard treatment, she is given a chance to reconnect with the morally virtuous and fully autonomous person she could have been before she committed a crime. Morris states that punishment will cause a rule-breaker to feel guilt over a wrongdoing, relinquish the guilt by undergoing this deprivation of freedom, and view himself once more as a morally autonomous person worthy of respect. Only the infliction of pain, or punishment in some other form as a deprivation of freedom, would accomplish these goals. Duff adds that respecting autonomy involves never treating a criminal as though he is beyond redemption.

Furthermore, punishment according to Hampton and Morris cannot be used coercively to elicit the “correct” response from a criminal: we cannot torture a criminal until he seems repentant, or keep him in prison until he shows compliance with society’s laws, for instance.

Punishment in the educative vein is distinct from the ways in which we treat animals, which amounts to mere conditioning in order to deter problematic behaviors. There is a difference between punishment (hard treatment that, at least according to the MGT, is educational) and conditioning (creating a positive or negative association with an action to deter or encourage that action). Punishment serves to enlighten criminals that their action is wrong, and thus leads them to avoid this behavior because they understand the reasons it is wrong. Conditioning forces a criminal to conform to laws because he fears pain, with no grasp on the moral reasons for the existence of those laws.

Other authors incorporate moral good aims into either utilitarian or retributivist theories of punishment. Sverdlik (2013) claims that utilitarianism can better account for goals of the MGT. Utilitarianism can, by definition, have a number of sub-goals if these work towards maximizing utility; as such, the theory can include a moral aim if need be. John Shook (2004) synthesizes parts of John Dewey's philosophy in forming an MGT variation under the broader context of rejecting retributivism, which strikes me as just another species of utilitarian. Dewey's MGT twist to utilitarianism enters in that he is concerned with producing moral citizens, not just people who follow the law for any reason. Dewey is unsatisfied that retributivist theories ignore "the proper moral context that gives our practices of asserting responsibility their meaning and effectively ceases to treat offenders as moral agents" (Shook, 2004). Dewey thinks that retributivism fails because it is inconsistent to judge a person as irresponsible and morally deserving of punishment, but then neglect to consider the effects of punishment on the person's *future* capacity for moral conduct. In other words, Dewey finds punishment without foresight morally reprehensible.

Robert Nozick (1981) attempts to defend a retributivist theory of punishment, though

ultimately his own theory strongly resembles the MGT. For him, retributive punishment is made up of two components: a teleological part, which is more or less the equivalent of an MGT, and a nonteleological part, in which punishment is valuable in and of itself. In the teleological component, the aim of punishment is to reconnect a wrongdoer with the correct values she has disconnected herself from. A moral good theorist would question whether this is truly retributivism, or whether this is a weakness of the retributivist standpoint that can be explained by an alternate theory such as the MGT. Yet, perhaps recognizing the shortcomings of the MGT, Nozick is not satisfied with simply a teleological goal of punishment. Punishing in order to reform a criminal would be overly optimistic, since it is highly unlikely that the majority of criminals are actually likely to reform. Thus, Nozick adds in a nonteleological aim of punishment, in which the act of punishment is inherently valuable *regardless* of whether a criminal reforms or not.

Despite the numerous authors who modify and adapt the MGT, I believe that the theory itself remains vulnerable to serious and perhaps unanswerable objections, which I will detail in section 3. Broadly, I take issue with the idea of moral education. How can we be sure that punishment, whether in theory or in practice, actually educates? In what ways is being freely attached to the moral good a beneficial situation? How can we be sure that someone who has undergone punishment is truly morally reformed, and not just conditioned to claim that he is?

I also object to the idea that punishment can be consistent with a protection of autonomy. I find such an idea counterintuitive because incarceration, fines, and other punishments are undeniably a loss of autonomy. Perhaps Hampton and Morris are only concerned about *moral* autonomy, but if so, they do not offer a justification as to why moral autonomy deserves protection and other types of autonomy do not. There is also notable conflict within variants of

the theory itself when it comes to autonomy. It seems that authors cannot decide whether autonomy is of critical importance – as in the theories espoused by Hampton, Duff, and others – or something to be overruled for a person’s own good, as in the case of moral paternalists. The MGT would be better served by discarding the idea of autonomy altogether, and arguing that punishment is something which benefits a person regardless of his own wishes and free will.

To dissolve the tensions faced by the MGT, I propose an alternate theory. Seizing on a line of argument proposed by Duff, that punishment expresses a continuing concern for the wrongdoer as a member of the community, I view the benefits that punishment can offer as social, rather than morally educational. As Duff claims, in committing a crime, an offender has separated herself from moral goodness, as well as the protective space of a law-abiding society. Punishment, according to Duff, allows her to reconnect with both the moral good and with her community. Duff focuses in large part on this moral good and a criminal’s personal reform; I am only concerned with the notion of social reintegration. In this paper, I will defend a new version of the MGT so that the theory better holds against scrutiny. I will call the improved theory the Social Good Theory (SGT).

## **2. THE SOCIAL GOOD THEORY**

Like Russ Shafer-Landau (1991), whose objections are explored in depth in section 3, I am skeptical of Hampton’s claim that punishment truly offers a moral education. However, I also take seriously the idea that we cannot harm people, seeking to avoid the vicious pitfalls of retributivism: if we harm people because they harmed society, then we have sunken to the level of criminals. My aim in this paper is not to defend or justify punishment as it exists now, but to defend the notion that punishment can in theory benefit the offender, revising the MGT and

accounting for its weaknesses.

To that effect, I propose a new and strengthened version of the MGT, the Social Good Theory (SGT). Here, I accept the same basic framework as the MGT, as stated in section 1:

1. We can never harm people, and can only justify institutions that do no harm.
2. Punishment is not a harm to offenders if it can be shown to confer a benefit on them.
3. Punishment does benefit offenders.
4. Thus, punishment can be justified.

Again, (3) is the premise that is in need of defense. As stated, I do not think that MGT supporters adequately defend punishment being a benefit to offenders, since the idea of a moral education is vague at best.

I do believe that punishment can still benefit offenders, but that these benefits are social as opposed to educational. In essence, an offender removes himself from society when he commits a crime, and punishment is what allows him to be integrated back in. Being reintegrated into society is inherently beneficial, since existing within a society is preferable to existing outside of it, and so the process of punishment confers a benefit. I will thus need to defend two ideas: that punishment allows a criminal to rejoin society, and that being a part of society is beneficial.

The first idea, that punishment allows a criminal to rejoin society, has two parts. The first part is that in committing a crime, a person estranges herself from society. Let us view society as a social contract, in which a group of people live together under an agreed-upon set of rules. Committing a crime is the equivalent of breaking one of these rules. Given that the rules were mutually decided upon and consented to, we can see how our criminal has ostracized herself. She is morally estranged in that she has broken away from society, defined as a rule-abiding set of



people. She also departs socially, given that citizens are unlikely to continue their associations with her. Even without a formalized institution of punishment, her lawful peers and neighbors would frown on her, knowing her to be untrustworthy and self-serving. They may actively deny her membership in social institutions or inflict revenge upon her. We cannot rely on the passage of time or the inner goodness of citizens' hearts to normalize relationships between the criminal and her society: there must be another method. The second part is that punishment is this method, and in fact the only method, in which the criminal may rejoin the society from which she has been estranged. Morris (1981) claims, and I accept here in formulating the SGT, that when a criminal commits a crime, he incurs a type of debt. When someone obeys the rules of the social contract, he is in a neutral state. The moment he breaks a law, he has gained an unfair and illegal benefit, acting on an impulse that everyone else is able to suppress. There are some crimes that may not benefit a criminal at all; accidental crimes, for instance, such as vehicular manslaughter. In these, a criminal is indebted not due to anything he has unfairly gained, but because he has unfairly taken away from society, even if it was unintentional. When we punish him by inflicting a negative treatment on him that non-criminals do not receive, we cancel out this debt.

Criminals gain moral reintegration into society by being reset to net zero in terms of things gained or taken from society (the result of the crime) and things lost (the punishment). They gain social reintegration into society when others accept punishment as a satisfactory payment of the debt and agree to take them back. Members of society are unsatisfied unless they believe that the wrongdoer *understands*, in the empathic sense of the word, the seriousness of what she has done. Such understanding can only come about when she undergoes some deprivation of freedom herself, which communicates to her the wrongness of what she has done.

We can see this idea playing out in our own lives. Interpersonally, we often do not forgive friends who have brought us harm until we are satisfied that they have suffered, emotionally or socially.

Of course, there are other types of debts or unfair benefits that we do not negate through punishment: being born into a high socioeconomic class, winning the lottery, and so forth. The only debt we seek to cancel out through actual punishment is the type incurred when someone breaks a law; our laws are codified statements that delineate what a society finds permissible and what it doesn't. Punishment can take many forms – paying a fine, incarceration, or mandated community service, for instance: what matters is that it is against the criminal's will and a deprivation of some freedom, in order to pay for what he has gained through the crime. The SGT also requires that the degree of punishment fit the degree of the crime, in order for the debt to be paid in full, not less and not more. Punishment is also restricted to people who have actually committed crimes, since they are the only ones needing to pay off a debt.

Note that the sentence is always determined *before* a criminal undergoes it: this sentence, barring extenuating circumstances, cannot be modified as the criminal is carrying it out. Thus, we determine an adequate payment of the debt, through the court and law enforcement systems, ahead of time. This, too, controls for at least some bias in the idea of social reacceptance. If we allow a criminal to reenter society only on the condition of social approval after he has done his time, the system might be skewed: contractors may be more likely to accept people of some races and not others, for instance. If we consider debt repayment the sentence that is laid out ahead of time, we limit our unjust biases to prejudice in the law enforcement and court systems, something that is not relevant to the SGT and that I will not address here.

Assuming that punishment allows a criminal to be socially reintegrated, we must show

that being in society is beneficial before we can argue that punishment confers any benefit. Let us again view society as a contract, where all live together under a set of rules. This situation is advantageous for all those partaking in it: in exchange for compliance with the rules, a person receives the guarantee of safety, as well as other positives such as social inclusion, education, community infrastructure, and whatever else the government would provide. Being in this situation is always preferable to being outside of it. This idea is backed by a long philosophical tradition, from Plato to Hobbes. Aristotle accepts that “a human being is by nature a political animal. Anyone without a city because of his nature rather than his fortune is either worthless or superior to a human being” (Reeve and Miller, 2015). Imagine the case of someone who is not accepted by his society: he cannot find employment to sustain himself, he may not be able to obtain proper food, shelter, or health care, and he likely does not have many opportunities for socialization. Thus, if someone commits a crime and marks himself as a social defect, we can accept that it is better for him to escape this dismal situation. He can do so by being punished: thus, punishment is beneficial.

It could be the case that a criminal rejects his social reintegration, just as a criminal could reject being morally educated. This should not faze the SGT. After the criminal has undergone his sentence, his debt is paid, whether or not he wants it to be. His subsequent return to society *is* his reintegration, and even if he returns to a life of crime a moment later, he has still had a moment of rejoining society. This society, as we have said, is a beneficial situation, regardless of whether or not the criminal accepts it as such. The benefit of punishment, therefore, is independent of how the criminal feels or reacts to it.

In further defending the SGT, I seek to overcome other problems faced by the MGT, and will explain objections in greater detail below.

### 3. OBJECTIONS TO THE ORIGINAL MGT

#### A. *Incarceration is a necessary punishment but also causes inevitable harm*

One of Shafer-Landau's criticisms is summed up as follows: "Educationists have, and possibly must, take incarceration as the best available punitive sanction. Yet it is unclear how morally educative such a punishment will be" (Shafer-Landau, 1991).

Let us assume that we do accept punishment as the correct choice of action in order to morally educate a wrongdoer. The next step would be to define what form punishment takes. The general consensus from MGT proponents and critics alike is that punishment involves hard treatment, a deprivation of the wrongdoer's freedom. From here, Shafer-Landau argues that MGT proponents are obligated to support incarceration as the form of punishment most conducive to moral education.

He arrives at this conclusion using the process of elimination. Alternative punishments to incarceration are, simply put, infeasible. Exile is difficult and does not accomplish the goal of education better than incarceration would. We could impose fines on a criminal, or have him participate in forced labor, but this again does not guarantee education. Public humiliation might compel a wrongdoer to change his behavior, but this could be too arbitrary a punishment to match to a crime. In addition, a psychopathic criminal would be unmoved by public humiliation. Forms of coerced therapy could also reform a criminal's behavior, but most MGT proponents would say that this is a type of conditioning; the criminal cannot truly absorb and understand what he has done wrong. I also doubt that total annihilation of personal agency in forms of punishment like torture or coercive therapy could be educational.

Shafer-Landau admits that the type of punishment an MGT proponent might favor would include punishments that are tailor-made to fit the crime, such as "an arsonist [working] in a burn

victim ward” (ibid). He claims this is an infeasible form of punishment given that judges don’t have the time to make such detailed or creative sentences, and that some crimes do not have such obvious punishments – what about, for example, a tax cheat? We could argue that if we are unable to come up with an appropriate punishment, we are not justified in punishing at all. But Shafer-Landau dismisses this notion, indicating that we still believe crimes merit punishment (or, for the MGT proponents, that criminals deserve a chance to reform). Thus, we are left with only incarceration as a practical way of dealing with wrongdoers; the only customization to the crime we can hope for is varying the length of prison-time with the severity of the offense.

Having driven the MGT into this corner, Shafer-Landau attacks. He claims that incarceration cannot meaningfully provide any sort of education to a wrongdoer. Morris claims that this sort of hard treatment is beneficial because it 1) makes the offender aware of his wrongdoing, 2) conveys how seriously society takes its values, 3) allows the offender to repay his debt to society, and 4) provides an opportunity for forgiveness and redemption. Shafer-Landau rejects 1) and 2), retorting that criminals know full well the extent of their crimes and the degree to which society cares about them. Perhaps punishment of this sort provides some other, more empathic knowledge of the severity of the crime, but if the MGT supporter really wants to pursue this route, he would have to be in favor of *lex talionis* as well. He rejects 3) on the basis that repayment of a debt is not education, and that it is difficult to justify punishment as debt repayment since different people may feel the burden of a social contract to different degrees, and the benefit from committing a crime will affect each person differently. Thus, we cannot consistently speak of a “debt” for each criminal, since the theoretical size of this debt would vary. Lastly, Shafer-Landau rejects 4) in claiming that some crimes are too awful to be forgiven through punishment. Even if forgiveness from the community were forthcoming, it would do so

because the criminal had had a change of heart, and not because of the incarceration itself.

As such, Shafer-Landau rejects one tenet of the MGT. Its best option for punishment, he claims, is incarceration, but incarceration itself is not morally educational. I will take this criticism seriously, showing that punishment does provide a benefit regardless of its form, even incarceration.

The SGT can dispel a number of these criticisms simply through its framing. For instance, the new theory does not bother with moral education. If criminals receive a moral education through their punishments, all the better, but the SGT justifies punishment on the ground that it allows a criminal to rejoin society. To that end, we can adapt Shafer-Landau's criticism to ask if incarceration really can reconcile a criminal with the social good, and if so, if it is the best form of punishment for doing so.

There are certain types of punishment that the SGT rules out: exile, by definition, does not help a criminal rejoin society. One could argue that the SGT would allow for exile, since this is, in a way, a form of re-entering a (new, foreign) society. Yet such thinking conflates exile as punishment with exile as entry into a new society. Integration into *any* society must be preceded by punishment, so exile cannot be the punishment itself. It also seems unlikely that a foreign society would happily accept the outcasts of another who are incapable of following a set of rules, whether or not the foreign society shares those rules. The SGT may not necessarily rule out coercive therapy or torture, but these would infringe on our agreement not to cause net harm. I agree with Shafer-Landau that punishments customized to the crime have a kind of poetry, but may be too difficult or expensive to carry out.

We are left once more with incarceration, and are charged with the task of explaining how prison can possibly help a criminal rejoin society. To have a plausible SGT, we must avoid

thinking of or establishing prison the way it exists in the United States, which generally seems like a harsh environment where no moral education or reform could ever take place. Ideally, we use prison to deprive criminals of certain freedoms, like the freedom to move about unrestricted or make choices about one's life, but this environment should not be the equivalent of physical or psychological torture. The goal of a prison is to highlight the difference between our contracted society, which is an ideal place to live, and its alternative, a nonideal place because its residents do not have the same amenities or freedoms as those on the outside. Punishment is undoubtedly a harm to a criminal; the idea behind the SGT is that it still provides a net benefit, given the desirability of being reintegrated into society. Some harsh treatment is necessary for this reintegration so that everyone else can accept the debt has been repaid, and as I have shown above, incarceration is the only viable option for doing so. Members of society, who understand that prison represents a serious deprivation of freedom, will thus be satisfied that the debt has been repaid.

I also take seriously Shafer-Landau's criticism that it is difficult to speak consistently of debt, since some people benefit more from a crime than others would. Furthermore, the same punishments for the same crime could be much more of a setback to one person than to another, such as making a rich person and a poor person pay identical fines. It seems that when we lay out sentences for each crime we may have to contextualize each punishment based on how much a criminal has gained and how much he must lose. This is something we see in our current justice system, but it does weaken the idea that like crimes must receive like punishments. The nuanced idea of debt remains a challenge for the SGT, one which I will not be able to explore in depth given the scope of this paper.

*B. Why benefit the wrongdoer?*

Suppose that we can develop a system of punishment which does, in fact, benefit offenders, either through moral education or for social reasons. The next question is why we want to make punishment beneficial at all, given that the offender has done something wrong. MGT proponents may be concerned with the overall welfare of society, but they are hardly utilitarian. There are punishments that the MGT supporters disapprove of, like torture, which would potentially increase a society's utility but is a detriment to the offender's welfare. Their primary goal, therefore, is to make sure that the offender benefits.

Shafer-Landau points out how counterintuitive this premise is in the first place. Society only has limited resources, and it seems odd to allocate its wealth to those who have violated the law when it could be benefitting innocents (e.g., by providing better education, infrastructure, health care, and so on). He raises two points here that an MGT supporter must answer. The first is that we must show that society has an obligation to benefit criminals at all, even if it has unlimited resources. The second asks, more pragmatically, why we should benefit criminals at the expense of benefitting other, lawful citizens.

Shafer-Landau focuses his objection around the first point, given that it must be answered before the second point becomes relevant. According to him, Morris answers the question by saying that criminals suffer from a great evil; "their souls are in jeopardy as their victims' are not" (Morris, 1981). Shafer-Landau rejects this idea, indicating that it would be nonsensical and cruel to say that, for example, a rapist suffers more than his victim. Furthermore, to say that a criminal is suffering from a great evil is to undermine the MGT's own premise that criminals have had the autonomy to do what they do, which is why we may punish them. After all, we offer different punishments, or none at all, to those who are coerced or insane.



In order to incorporate this objection, the SGT therefore must do two things: show that we have an obligation to benefit criminals, and show that we should benefit criminals at the expense of people who have *not* committed crimes.

As for the first statement, both Hampton's MGT and my SGT as outlined in sections 1 and 2 have accepted as the grounding premise that we have an obligation to do no harm. This is certainly contentious, but for the sake of defending this theory, I have accepted the premise as it is. Nonetheless, there are ways to defend the idea that we cannot bring net harm to criminals. If we start inflicting damage on criminals because they have done something bad, we are merely carrying out revenge. In addition, imagine that we felt as though we had no obligations to protect criminals at all and a desire to spend no resources on them; this would amount to simply executing all criminals on sight. This notion would strike us as abhorrent, leading us to believe that we do have at least some obligations to protect criminals. Still, we still would like to retain an institution of punishment. Suppose that we eliminate all punishment whatsoever. Presumably, chaos would ensue in society, and we have failed our promise not to cause our contractors harm. My aim here is to see if punishment can still be justified if we cannot harm anyone; the only way to justify it is to show that punishment provides a net *benefit* to offenders or at least have no effect on them.

Though both the MGT and the SGT are based on this premise of benefitting criminals, the SGT's proposed benefits are perhaps less offensive. The moral good theorist proposes direct benefits to criminals by reforming them into morally good people, begging the question of why we should put such focus on them when they've done nothing to deserve it. The social good theorist wants to make society a better place to live, and the benefits of punishment come from allowing criminals to reenter. Thus, the benefits to a criminal are derivative.

So much for the first obligation. Let us now discuss Shafer-Landau's second objection, that the state should not benefit criminals instead of innocent people. First, I respond that these are not mutually exclusive goals. Suppose that the government seriously invests in building up infrastructure, public education, health care, social security, and the like. We can fathom that this will help lower crime rates. Additionally, this will benefit lawful contractors and criminals alike, the former directly, and the latter by giving an incentive to undergo punishment and rejoin the society whose rules they have flouted. Our limited resources are used to make society overall a better place, which benefits everyone, criminals or not. Beyond that, as I have shown, we have an obligation not to hurt criminals, equal to an obligation to benefit them. Thus, in spending any resources on our institution of punishment, we are simply fulfilling an obligation; we do not have a *choice*, on a moral level, to spend these resources in this way – we *must*.

Furthermore, if we choose to benefit criminals, we will eventually benefit ordinary citizens as well. If our prisons are humane, dignified places which respect human rights and strive to educate and reform criminals, all contractors will be better off when those criminals not serving life sentences ultimately rejoin society. We need not view improving our system of punishment as a sacrifice of resources that could have been spent on innocents instead; everyone belongs to the same society, in the end.

Shafer-Landau also suggests that it is morally objectionable to try and benefit the criminal as opposed to the victim. There are several ways to respond to this. First, some crimes have no victims, or it is unclear who the victim may be. Say someone robs a bank. Whom are we repaying through punishment: the clerks in the stickup? The CEO of the bank? Even if there is a victim, his or her feelings are too unreliable a thing on which to predicate an entire theory of punishment. Different victims of the same crime may have entirely different opinions on how to

best “repay” the criminal’s debt. Second, criminal law in general is unconcerned with victims: crimes are offenses against the state, not individuals. It is society’s rules we have broken, and thus to society we owe a repayment.

*C. The need for indeterminate sentencing*

If an MGT proponent is truly concerned with criminals being morally educated and reformed, then she may be committed to a practice of indeterminate sentencing, or punishment until the criminal has been properly “educated,” which may be more or less time than the original sentence. Hampton anticipates this objection and claims that indeterminate sentencing would not be a way to respect autonomy. We cannot force anyone to swallow a moral lesson, and so we must teach our lesson as best as we can and “whether or not the criminal will listen and accept it is up to the criminal himself” (Hampton, 1984). Committing a criminal to prison until he shows himself reformed would encourage him to act, however falsely or manipulatively, to convince others that he has learned his lesson. Morris, too, does not believe his MGT variation would allow for indeterminate sentencing, because “the goal is not repentance at all costs, if that has meaning, but repentance freely arrived at” (Morris, 1981).

Shafer-Landau counters that the MGT proponent may be committed to indeterminate sentencing nonetheless: if the goal is moral education, different offenders will need various amounts of time to meet such a goal, amounts which cannot be predetermined. If we give the same sentences to similar crimes, it is very likely that criminals regardless will experience different levels of moral education and reformation. He makes the point that we can justify like sentences for like crimes, but not under any educationist justification. If the goal of the MGT is to morally educate a criminal, we must punish him until he has been educated. If the goal of the

MGT is to *attempt* to educate a criminal, then we can avoid indeterminate sentencing, but we are charged with the task of explaining why such an attempt is beneficial or valuable.

While Hampton and Morris may or may not adequately dispel this concern, the SGT remedies the problem more easily. Given that the primary goal of the SGT is social reintegration, indeterminate sentencing is senseless. We are aiming at the end of bringing a criminal back into society, so keeping him in prison indefinitely would be counterproductive. One could object that the SGT still does necessitate indeterminate sentencing, using the argument that a criminal must stay in prison until she has proven herself ready to re-enter society. However, as I have noted in section 2, readiness to re-enter society is not, or at the very least *should* not, be measured by general public opinion of each criminal. Instead, when a court of law finds a criminal guilty and sends him to be punished, we take that sentence to be the cost that a criminal must pay in order to settle his debt. Thus, when that pre-determined sentence has been met, the debt is cleared, and the criminal can rejoin society. Can the MGT use the same defense? I believe not. There is no reason to think that a predetermined sentence equals a moral education, whereas it does equal some form of debt repayment (as determined by a jury ahead of time).

#### *D. What about autonomy?*

The importance of autonomy features heavily in the MGT as well as in its criticisms. R. J. Lipkin, for instance, believes that some MGT proponents do not take autonomy seriously enough, and he creates his own version of the theory in which autonomy matters so much that criminals should be given a say in their own punishments (1988). According to most variations of the MGT, autonomy is a critical part of moral goodness, and punishment is a way to restore a

criminal to full moral autonomy. Even though a wrongdoer might lose her right to freedom, she never loses the right to autonomy.

Shafer-Landau indicates that this leads the educationist to a dilemma. Either there are situations in which autonomy may permissibly be infringed upon, or it may never be infringed upon. If it is the former, then the educationist is charged with presenting a principled and non-arbitrary distinction for cases in which autonomy may or may not be infringed upon. If it is the latter, then the educationist is obligated to explain how punishment is not a limitation on autonomy. MGT supporters seem to favor moral autonomy, but as said earlier, it is unclear as to why moral autonomy is more worthy of protection than other forms of autonomy. Furthermore, none of the theory's proponents have offered even an adequate definition of what autonomy or moral autonomy are, which makes them hard to defend.

The SGT, I believe, can deal with autonomy more elegantly than the MGT in giving it much less weight. Autonomy is not as critical to an SGT proponent as it is to the MGT, and thus need not be as ardently defended. In a social contract, autonomy is not the be-all, end-all that it is to the MGT. A social contractor understands that in order to be able to enjoy some forms of personal autonomy – safety and property rights, for instance – he must give up others, such as the autonomy to cause harm. In fact, the end result is that the ideal social contract maximizes autonomy, since it allows for the most freedom of choice while only limiting a person's agency in that he must respect others' rights.

Thus, punishment is simply one way in which autonomy is infringed upon in order to increase autonomy in the future. Criminals are deprived of at least some freedom in order to rejoin the social contract, where they theoretically have the highest amount of autonomy they could have. We can conceptualize the benefits of living in society and being punished so as to

rejoin society without needing to refer back to autonomy. As Sarah Conly (2013) suggests, autonomy is just one component of welfare, which can include physical and psychological well-being. An ideal social contract protects all of these together, and the benefit of punishment is not its esteeming of autonomy above all else but its eventual restoration of autonomy and other types of welfare once a criminal reenters society.

If the SGT does not place the same emphasis on protecting moral autonomy, critics may say that the theory would permit some extreme punishments like torture that still work as debt repayment and eventual social reintegration. I will confront this idea in section 4B.

#### **4. OBJECTIONS TO THE SGT**

##### *A. What about the heinous criminal?*

One strong objection to the SGT is that there will be certain offenders whose crimes are so severe that they will be imprisoned or punished for the remainder of their lives. For these criminals, it seems that no degree of punishment can make up for the suffering they have wrought. They cannot repay their debt to society, and so they cannot possibly benefit from their punishment.

A few possible responses may be offered here. The first is that reconciliation is still possible. The criminal may never be fully reconciled in the sense that he can live as a civilian again, but he can begin to make amends by dedicating the rest of his life to this endeavor. Reconciliation need not take place in that a person rejoins society, either; his efforts to repay his debt can be recognized even as he remains in prison, and reconciliation becomes his lifelong project. In order for this to be construed as a benefit, the theory commits us to saying that the

process of reconciliation or paying the debt is inherently beneficial, even without the reward of rejoining society in the end. This is perhaps closer to Morris's and Hampton's original claim that punishment morally reforms a criminal, and regaining moral goodness is a benefit to an offender. Morris and Hampton also respond by saying that we must at least give all criminals an equal chance at a moral education; the SGT could argue similarly, though this is a more difficult task if we consider a criminal's sentence the repayment of his debt, and then make this cost impossible to fulfill, as with those who incur multiple life sentences. We do not know ahead of time if a criminal will receive a moral education from punishment; we *do* know ahead of time whether or not he will ever rejoin society.

Alternately, we could structure our systems of punishment in such a way that the more time a person is incarcerated, the more we consider her debt repaid, and the more we reward her for this. Even though she might not ever rejoin society, we could create a community of sorts within a prison, and grant her more access to and privileges within this community the longer she has been incarcerated, given her ongoing good behavior. Thus, undergoing punishment is still a benefit in letting her enjoy the advantages of a different sort of social contract.

As a final option, we can concede that the SGT does not cover all criminals. For those people who have violated the social contract so horribly we cannot imagine reintegrating them into society – the serial killers, torturers of children, and the like – punishment serves most fundamentally to restrain or remove them altogether. However, we would assume that most cases are not this extreme. Furthermore, though these may be exceptions to the rule that we must benefit criminals, we cannot base an entire theory of punishment around these rare evil-doers. Punishment is justified as an institution that holds the social contract together. If we attempted to remove every single rule breaker from the social contract altogether, our numbers would dwindle

rather quickly. We would also inspire compliance with the rules through sheer tyrannical fear-mongering as opposed to a demonstration of our commitment to the wellbeing of society.

*B. What methods can we use to reintegrate people into society?*

Having established that our goal in punishing is reintegration to society, critics naturally may question what such punishment would look like. Shafer-Landau accepted that incarceration is the only viable punishment that moral educationists can promote. However, now that we are less concerned with autonomy and moral reform, anyone could naturally suggest a number of ways to get a criminal to realign with the social contract.

First, methods such as torture could conceivably be highly effective in repaying debts incurred through wrongdoing, encouraging compliance with the social contract, and preventing recidivism. We can imagine an Orwellian society in which criminals are punished to a psychological breaking point, after which they will become submissive members of the social contract once more. However, the SGT is likely to prevent such treatment. Though it is a good thing to be part of society and thus a benefit to rejoin society, this reintegration is not unconditionally good. It is not good, for instance, for the social contract to consist in part or in whole of psychologically damaged individuals. In creating such a world, we have defeated the purpose of a social contract, which is to create a space in which people may live knowing that their basic human rights are protected. Furthermore, it is these basic human rights that protect a person against torture, not the SGT. Torture may be morally wrong outside of any theory of punishment, leading us to ban its employment regardless of what we hope to accomplish when we punish.

Second, if punishment is beneficial in that it repays the debt, we might wonder what



prohibits us from punishing in the manner of *lex talionis*. The most direct way to cancel a debt that a criminal incurs, it seems, is to inflict upon him the same exact crime that he inflicted upon someone else. This may be true, but we have stated that our goal is reintegration. If rejoining society is at all a benefit to a criminal, then society must first be a beneficial place. I imagine that the kind of society which abuses the abuser or tortures the torturer loses credibility as a benevolent place to live. Also, though we are arguing that social reintegration is the *best* good that a criminal could have – greater than any crime – it also is not priceless, and its value is not infinite. Thus, the more extreme or detrimental a punishment, the smaller the net benefit to the offender when she is reintroduced to society.

Third, let us imagine that there is a sort of coercive therapy or pill which, when administered to a wrongdoer, guarantees that she will never break a law again. If she is released back into her society, she will certainly live within the boundaries set by the social contract. Still, I argue that the SGT does not permit such a form of punishment, either. Social reintegration exists on the condition that society must accept and be willing to forgive a criminal. This acceptance will likely not be forthcoming unless the wrongdoer has done something to atone for his crimes; this debt can be repaid by hard treatment, not by use of a pill. As the moral educationists suggest, we are on the whole interested in knowing that a criminal has understood what he did wrong, which is why a pill would be inadequate. Hard treatment is the most reliable route to such understanding, which in turn makes it the method most likely to be accepted by others in society as adequate repayment of debt.

## 5. CONCLUSION

I have shown that the Social Good Theory is a stronger theory than the Moral Good Theory. Moral education and moral good are only ever an ambiguous result of punishment. Even if criminals do always receive a moral education, or the chance at a moral education, it is still difficult to argue that such a situation is inherently beneficial. That would depend on our acceptance of adherence to the moral law as the greatest good, a rather abstract concept.

Being part of society, on the other hand, is a more concrete and pragmatic way to measure good. We do not have to claim that existence within the social contract is an unqualified good, rather indicating the various benefits that a person receives: communal infrastructure, education, health care, as well as a guarantee for all contractors that they will be ensured protection of basic rights.

There is one final question: can the Social Good Theory, as it stands, actually justify punishment? To date, I have not found a theory of punishment that I have considered satisfying: not retributivism, not utilitarianism, not a mixed conception, and not an alternate theory. In my view, the biggest blow to the MGT and SGT is their counterintuitive nature. Simply put, we do not view criminals as people we should benefit. Likewise, imagine we abolished all systems of punishment overnight. Our biggest fear would not be that wrongdoers have no way to repay their debt and integrate back into society, but rather that utter chaos would result due to the lack of apparent consequences for rule-breaking. Thus, it seems to me that we have created punishment first and foremost to reinforce compliance with laws and prevent criminal mayhem: score for the utilitarians.

That said, neither the MGT nor the SGT should be dismissed offhand. These theories could act as a standard of reform, not a justification of what currently exists. Shafer-Landau was

reasonable to claim that educationist goals cannot be the *sole* justification for punishment. Yet the Moral and Social Good Theories are illuminating, in a way that no major justification for punishment has been, as to the ways in which we currently mistreat criminal populations. When we accept the premise that we cannot hurt anyone, and ask ourselves how we can still punish under that assumption, our viewpoint of current punishment practices must shift. We do value deterring crime, and the institution of punishment is unlikely to disappear overnight; the SGT (and the MGT) can be informative as to what kinds of punishment are permissible. These theories help us recognize that the ways in which we punish now, particularly in the United States, are hardly conducive to moral education or social reintegration.

Finally, the SGT is a way of marrying some utilitarian and retributivist concerns. Benefitting criminals, after all, can promote overall social utility (and vice versa), as I have previously explained in section 3B. However, the means needed to benefit a criminal via his eventual reintegration is a type of debt repayment through harsh treatment, something which assuages the retributivist camp. Though the SGT would likely ban the *lex talionis*, it would still mete out punishments in rough accordance with the degree of the crime. And the reason social reintegration would occur at all is because a criminal's previous society would view punishment as something the criminal deserved, and once she has undergone it, she has fulfilled this obligation and repaid her debt.

Therefore, though the SGT is not the perfect justification or even the sole justification of punishment in this or an ideal world, it does represent an improvement over the MGT. Though the theory would flourish best in a much more just world than our own, it serves to indicate flaws and inadequacies in current punishment systems, and remind us that social harmony and inclusion are worthy goals for which we should strive.

## **BIBLIOGRAPHY**

- Conly, S. (2013). *Against Autonomy*. Cambridge: Cambridge University Press.
- Duff, R.A. (1986). *Trials and Punishments*, Cambridge: Cambridge University Press.
- Hampton, J. (1984). The Moral Education Theory of Punishment. *Philosophy & Public Affairs*, 13(3), pp.208-238.
- Lipkin, R. J. (1988). The Moral Good Theory of Punishment. *U. Fla. L. Rev.* 40.
- Morris, M. (1981). A Paternalistic Theory of Punishment. *American Philosophical Quarterly* 18(4): 263–71.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge, Mass.: Harvard University Press.
- Reeve, C. and Miller, P. (2015). *Introductory readings in ancient Greek and Roman philosophy*. Indianapolis, IN: Hackett Pub. Co.
- Scott, S. (n.d.). *Examining the Moral Good Theory of Punishment*. Undergraduate. Pomona College.
- Shafer-Landau, R. (1991). Can punishment morally educate? *Law Philos*, 10(2), pp.189-219.
- Shook, J. (2004). Dewey's Rejection of Retributivism and His Moral-Education Theory of Punishment. *Journal of Social Philosophy*, 35(1), pp.66-78.
- Sverdlik, S. (2013). Punishment and Reform. *Criminal Law and Philosophy*, 8(3), pp.619-633.