

Consequentialist Theories of Punishment

*This is a pre-print of the following chapter: Hsin-Wen Lee, Chapter 7 Consequentialist Theories of Punishment, published in *The Palgrave Handbook on Philosophy of Punishment*, edited by Matthew Altman 2023.

Abstract

In this chapter, I consider contemporary consequentialist theories of punishment. Consequentialist theories of punishment look to the consequences of punishment to justify the institution of punishment. Two types of theories fall into this category—teleology and aggregationism. I argue that teleology is implausible as it is based on a problematic assumption about the fundamental value of criminal punishment, and that aggregationism provides a more reasonable alternative. Aggregationism holds that punishment is morally justified because it is an institution that helps society to aggregate important moral values. Several theories fall into this category, including general deterrence theories, specific deterrence theories, and preventionism. I critically evaluate these theories and argue that only one specific deterrence theory, namely, my rights-protection theory, provides the most reasonable consequentialist account of punishment.

Short Bio:

Hsin-Wen Lee is Assistant Professor of Philosophy at the University of Delaware, USA. Her areas of specialization include political philosophy and philosophy of law. Her research interests include nationalism and the philosophy of criminal punishment. Her book *Criminal Punishment and the Protection of Rights* is forthcoming.

I. Consequentialist and Deontological Theories of Punishment

Deontology holds that the moral value of an action is determined by the qualities inherent in the action itself. For example, lying is inherently wrong, and the wrongness lies in the qualities inherent in the action itself—e.g., the principle behind it cannot become a universal law or the act fails to respect people—not in factors external to it—e.g., bringing about negative short-term or long-term consequences. In contrast, consequentialism holds that an action's moral value is determined by its consequences. Thus, the moral value of lying can be positive or negative, depending on its consequences.

Similarly, to evaluate the moral value of punishment, deontology requires that we examine the qualities inherent in the acts of punishment, while consequentialism demands that we consider the consequences of punishment. Deontological theories of punishment—retributivism (Moore 1997; Kleinig 1973), desert theory (Kershnar 2000; Von Hirsch 2017), expressivism (Feinberg 1965), communicative theory (Duff 2001), etc.—hold that punishment is morally justified when an offender commits a crime without any excuse or justification. Accordingly, the morally correct response to such criminal wrongdoing is to express blame through the imposition of proportionate punishment (Strawson, 1962). An important feature of deontological theories of punishment is that they are *backward-looking*—to justify the imposition of punishment, deontological theories look to past wrongdoing to support the moral adequacy of punishment.

In contrast, a consequentialist theory of punishment is *forward-looking*; it holds that the institution of punishment can be morally justified because of the good consequences it is expected to bring about in the future. For instance, punishment discourages criminal offenses and thereby leads to the good consequence of crime deterrence, reduction, or prevention. Different forms of punishment may bring about different costs and benefits. A fine costs offenders money, but this is good because it teaches them a lesson and brings monetary profit to society. Community service takes time and physical labor from the offender but offers valuable service to the community. Imprisonment deprives an offender of some of their most essential liberties; however, it also serves to incapacitate the offender so that they are no longer at liberty to commit more crimes in the community. The death penalty takes away an offender's life away and thus the offender cannot commit a crime again. Punishment also sends actual and potential offenders a message about the price of criminal wrongdoing—if one commits the same crime, then they will receive the same punishment as the wrongdoer. Further, some forms of punishment may serve to rehabilitate and even morally reform convicted offenders. When offenders receive the education, psychological or psychiatric treatment, or job training that they need to lead a normal life, they are less likely to commit crimes again in the future. If they morally reform and learn to obey the law, they are more likely to respect the rights of others and less likely to commit a crime again in the future. Consequentialist theories appeal to these good consequences to justify the institution of punishment.

II. “Consequentialist Theory of Punishment”—Teleology or Aggregationism?

Next, I shall consider two different types of consequentialism—teleology and aggregationism. According to teleology, there is only one thing or one state of affairs that possesses the ultimate moral value, namely, the *telos*. The *telos* is our final end or ultimate objective; it tells us how we can evaluate the moral values of different actions. The moral value of an action depends on its contribution to the “*telos*.” For two actions X and Y, if X helps to bring about the *telos* more quickly or more effectively than does Y, then X has a higher moral value than Y. Without such an instrumental contribution, actions themselves are not valuable. In other words, the only way actions can have moral values is by making an instrumental contribution to bringing about the ultimate good.

While teleology assumes that there is only one state of affairs which is ultimately valuable, aggregationism does not make this assumption. Instead, it maintains that the right course of action is the one that will help to aggregate the highest values. Thus, the best consequence depends on the value system one believes in and the possibility of value maximization. For the value system, one can be a value monist or pluralist. Monists believe that there is only one type of intrinsic value, and we ought to act in ways that will maximize the aggregated, total value at the end. For example, Jeremy Bentham believes that there is only one intrinsic value, namely, pleasure, and only one intrinsic evil, namely, pain (1780). The right action is the one that leads to the highest balance of pleasure over pain. On the other hand, a value pluralist believes that there can be different types of intrinsic values. For instance, John Stuart Mill believes that there can be higher and lower pleasures; that is, values can be ranked. Accordingly, we may rank the moral values of different actions by considering which pleasures they generate at the end. So long as an action brings about the highest possible aggregated moral value, it is the right action; if an action does not do that, then it is wrong. The possibility of value maximization depends on several factors, including the feasibility of different plans, the possibility of coordination among different people, and the current conditions of things and

people (e.g., individual preferences). Thus, whether an action is the right one depends on the highest moral value it can possibly generate.

Accordingly, when one speaks of “a consequentialist theory of punishment,” one may have in mind either a teleological or an aggregationist theory of punishment. Next, I shall briefly consider the teleological theory of punishment. Presumably, a teleological theory of punishment is a theory that takes the punishment of criminal offenders to be the ultimate good. We are to evaluate the moral values of laws and policies by examining their contribution to this final purpose. The view is implausible because, to achieve the ultimate good state of affairs, some crime must be committed first. Without a criminal offense, it is impossible to achieve the telos. I am not aware of any philosopher who defends this view.

There is another closely related view about punishment. Some philosophers argue that it is inherently morally good that offenders are punished (Moore 1997: 157; Kleinig 1973: 67). Accordingly, when offenders commit crimes, they ought to be punished, unless there are countervailing reasons not to do so. This view does not assume that the punishment of offenders is the ultimate good; neither does it require that any crime be committed. All it demands is that we punish offenders when they commit crimes. According to this view, the value of punishment lies in itself and not its consequences. Thus, it is a deontological theory of punishment rather than a consequentialist one. It suggests only that we have a *prima facie* duty to punish those who commit crimes. In determining the moral value of punishment, the theory looks to the past (backward-looking), and not to the future (forward-looking). It does not specify the ultimate good state of affairs; nor does it require that we aggregate any particular value. It only tells us the kind of duty we have.

Therefore, it seems that the teleological theory of punishment is implausible. Next, I shall consider other possible accounts of consequentialist theories of punishment. How should we evaluate the consequences of punishment? How do we know whether and how punishment brings about good consequences? In the literature, several different theories are categorized as consequentialist theories of punishment—general deterrence theories, specific deterrence theories, and preventionism. General deterrence theories hold that punishment is justified because it helps to deter potential offenders. Specific deterrence theories hold that punishment is justified because it helps to deter specific offenders, namely, convicted offenders themselves. Preventionism holds that extended punishment can be morally justified because it helps to prevent crime; that is, the good consequence of having the institution of punishment is that crimes will be deterred or prevented.

In the following sections, I shall consider four general deterrence theories, two specific deterrence theories, and preventionism. I shall argue that punishment does not help to achieve the consequences that preventionism intends, and that general deterrence theories fail to take effective measures to deter crime. Only one specific deterrence theory makes a genuine effort to deter crime. I shall begin with general deterrence theories of punishment.

III. General Deterrence Theories

General deterrence theories of punishment hold that the institution of punishment is morally justified because it serves to deter potential criminal offenses. The moral value of punishment is not intrinsic but instrumental; its value is in its contribution to the aim of crime deterrence. Having the institution of punishment means having an institution that will punish offenders whenever crimes are committed. Knowing that the institution of punishment is legally valid and effective, people know that there is an institution that will impose punishment on them if they

commit crimes. Given that most people do not want to be punished, they will avoid committing crimes so that they will not be punished. In other words, having the institution of punishment gives people a prudential reason not to commit crimes, thereby deterring criminal offenses.

Next, I shall consider four general deterrence theories of punishment.

i. Warren Quinn—Automatic Retaliation Device

Warren Quinn defends a general deterrence theory of punishment. According to Quinn, punishment is morally justified because we have a right to defend ourselves against aggressions. To protect ourselves against possible aggressions, we may threaten to punish those who harm us. The right of self-defense entails a right to threaten to punish those who will harm us. Accordingly, we also have the right to punish offenders. While many believe that we have a right to threaten to punish because we have the right to punish, Quinn argues that we have a right to punish because we have a right to threaten (Quinn 1985:336-7). The right of self-defense justifies the right to threaten to punish, and the right to threaten to punish justifies the right to punish.

In a similar manner, we have a right to activate an automatic retaliation machine (1985:337). The automatic retaliation machine that Quinn has in mind is one which automatically identifies criminal wrongdoers and imposes on them the threatened punishment. Because people do not want to be punished, activating such a device will serve to deter crime and protect society. Accordingly, our right of self-defense gives us a right to activate the machine to punish actual offenders in self-defense. If our right of self-defense gives us a right to activate such a machine, then it equally gives us the right to create the institution of punishment, which serves the same function and purpose.

While it may seem clear that we may activate an automatic retaliation machine to defend ourselves, it is less clear what kind of punishment program should be enforced by such a machine. First, from the point of view of crime deterrence, the threat of a disproportionate punishment may be more effective than the threat of proportionate punishment. Does this mean that we may threaten to impose disproportionate punishment on offenders, and does it also mean that we may actually impose the threatened disproportionate punishment on them? (Alexander 1980: 209-11; Tadros 2011: 269). In the literature, proportionate punishment is a standard requirement; however, there seems to be some tension between Quinn's view and this requirement.

Second, proportionate or not, regardless of the amount of punishment threatened, we may wonder whether implementing the threatened punishment really serves to deter crime as the fact that a crime is committed is evidence that the threatened punishment failed to deter at least that crime. Because the threat of punishment happens before crimes are committed, there can be a discrepancy between what was threatened for self-defense and what becomes necessary for future self-defense. Carrying out the threatened punishment does not necessarily serve to deter future crimes (Farrell 1989:133-5; Ellis 2003: 339). If one is serious about crime deterrence, one will not insist on carrying out the threatened punishment, but would be willing to adjust the form and amount of punishment according to the need for future crime deterrence. Thus, it is not clear whether Quinn's argument supports punitive policies that genuinely aim to deter.

ii. Philip Montague—Principle of Distributive Justice

On the other hand, Philip Montague suggests that when offenders commit crimes, they force us to make a choice between allowing ourselves to be harmed and defending ourselves against their attack (1995:40). No one has a duty to be harmed unless one was responsible for the harm created. Crime victims have the right to act in ways that would protect themselves from the harm

created by their aggressors. Further, they are allowed to redistribute the harm that would befall them back to their aggressors. Because the harm was caused by no one else but the aggressors, they cannot complain that the harm imposed on them is unjust. This is simply a matter of distributive justice. Montague calls the principle that allow us to redistribute the harm that the aggressor created back to the aggressor *Principle J* (1995: 42). He believes that the same principle of distributive justice allows us to create the institution of criminal punishment, which allows us to redistribute the harm offenders created back to the offenders themselves. Because the institution of punishment is created to defend ourselves against unwanted harm, it is a form of self-defense and is subject to commonly accepted constraints of self-defense, including proportionality, minimization, and side effects (1995: 45-6).

While the principle of distributive justice that Montague defended is clearly a plausible one, it is less clear how this principle can be used to morally justify the institution of criminal punishment. First, the institution of punishment is primarily about the offender, not the victim; however, Principle J is primarily about the victim, not the offender. The principle supports the right of victims to defend themselves against the harm that was created by their aggressors. That is, the principle seeks to defend victims against unjustified aggression. However, it is not clear whether and how the institution of criminal punishment helps to serve this purpose. To begin with, punishment happens after a crime is committed and the harm has already been imposed on the victim. Thus, it is not clear how punishment helps to defend the victim against aggression. Further, typical cases of punishment—monetary fines, community service, imprisonment, the death penalty—are not designed to protect victims' rights or enhance their wellbeing. Depending on the conception of punishment to which one subscribes, punishment may be designed to give offenders what they deserve or to send potential offenders a warning message. It is not clear how such an institution helps to protect the victim against aggression and achieve distributive justice. Instead of punishing the offender, perhaps distributive justice should require that we make sure that the losses of victims are compensated for. Still, even when we take effective measures to protect victims and make up for their losses, it does not seem that we run out of reasons to punish offenders.

iii. Daniel Farrell—Weak Retributivism

Daniel Farrell embraces the principle of distributive justice that Philip Montague proposed. He argues that Principle J is more fundamental than the principle of self-defense. Based on this assumption, he further develops a general deterrence theory of punishment. When offenders commit crimes, they force us to make a choice between allowing victims to be harmed and harming the wrongdoers (1985:374). We are justified in choosing the latter.

Further, when offenders commit crimes, we normally feel that they should be held accountable for only the harms they created and no more. Thus, adequate punishment is determined primarily by the harms that the offender created. However, according to Farrell's account, when an offender commits a crime, the harms they created are not limited to the ones that they actually created with their offense. For instance, an offense may make the victim weaker and less able to defend themselves against a latter attacks by others, or the offense may expose the victim's weakness to other potential attackers, thereby making the victim more vulnerable than before. Thus, an offender is accountable not only for the harms they directly created with their offense, but also for the additional risk and danger that their offense indirectly brings about to the victim (1985:383-384). Accordingly, Principle J would allow us to impose additional harm on the offender when doing so is necessary for self-protection. In short, distributive justice allows us to punish wrongdoers more than they harm us. Farrell calls this

view *weak retributivism*. Farrell suggests, “wrongdoers may be punished beyond what is necessary to keep them from doing wrong again—if so punishing them can plausibly be said to be likely to deter others from doing wrong themselves” (1985: 368).

While Farrell’s argument seems plausible, it is not clear why self-defense would require that offenders be punished more than is necessary. First, it is not clear how punishing an offender more would help to deter other potential offenders from committing similar crimes. People are situated differently; they have different beliefs about facts and values. Accordingly, the same punishment may not serve to deter all of them. For various reasons, potential offenders may not be deterred even by an actual case of severe punishment. For instance, one may be confident that one will not be caught—because one is very careful about one’s criminal activities, or because one is a member of law enforcement and has some control over the criminal justice procedure, or one may lack the mental capacity to comprehend how this particular punishment gives one a reason not to commit a crime, or one may commit a crime because one is overwhelmed by strong emotion, temporarily fails to consider one’s reasons properly, and loses control of one’s actions (Lee 2017). Under those circumstances, one cannot be deterred. Consequently, punishing one to deter another is an inherently shaky undertaking. Because the target audience is not composed of one idiosyncratic group of individuals, one cannot be sure how the different audiences will perceive the reason provided.

Second, because Farrell’s argument is also based on Montague’s principle of distributive justice, it encounters the same problem. Distributive justice in forced choice situations focuses primarily on the protection of victims rather than on the punishment of wrongdoers. The principle provides moral support for policies that favors victim compensation; however, it is less clear what it entails about punishing offenders.

iv. Victor Tadros—An Enforceable Duty to Protect

Victor Tadros argues that, from the duty not to cause harm to others, we may derive a duty to protect one’s victims, which in turn supports an offender’s duty to be punished. People have rights against being harmed, and we have a duty not to harm others. If one harms another person, one violates one’s duty not to harm others. When this happens, the duty not to harm generates a duty to protect one’s victims from the harms one caused (2011: 3-4).

According to Tadros, this duty to protect is enforceable and grounds our right to punish offenders. When offenders commit crimes, they incur a duty to protect their victims. How can they fulfill this duty? Tadros believes that they can fulfill this duty by being punished. When offenders are punished, their punishment sends a warning message to potential offenders in the community and deter them from committing more crimes, thereby protecting their victims.

Tadros’s argument suffers from the same problem as other general deterrence theories, namely, the punishment of convicted offenders does not necessarily serve the purpose of crime deterrence. To be clear, punishment of offenders may deter at least some people—those who, upon learning about the punishment, are concerned about the consequence of punishment. However, offenders are different, and seeing that one is punished does not give everyone a reason against committing an offense. I have considered this point briefly in the last subsection and shall further consider this point in some detail when I consider my rights-protection theory in section IV.

v. On General Deterrence

Before moving on to consider specific deterrence theories, I should comment quickly on how the four general deterrence theories of punishment considered in this section respond to common objections against consequentialist theories of punishment.

In the literature, many consider consequentialist theories of punishment to be implausible because they are associated with problematic moral judgements. First, consequentialist theories look to the consequences of an action to determine its moral value; accordingly, as long as the consequences are good, we may punish innocent persons or impose disproportionate punishment on offenders (McCloskey 1965). Further, consequentialist theories allow us to use persons as mere means. General deterrence theories require that we punish actual offenders for the sake of warning potential offenders, and this clearly uses convicted offenders as mere means. Most people are appalled by these ideas and thus believe that we must reject consequentialist theories of punishment. Next, I shall consider how the four theories respond to these objections.

a. Punishing Innocent Persons

It does not seem that any of the general deterrence theories considered here supports the punishment of innocent persons. Quinn's automatic retaliation device is activated by actual criminal offenses only; innocent persons cannot activate this device and thus will not be punished. Montague's principle of distributive justice would require that we redistribute the harm back to the aggressor only; because we cannot "redistribute" the harm back to innocent persons, his argument would not support the punishment of innocent persons either. Lastly, both Farrell's weak retributivism and Tadros's enforceable duty view appeal similarly to Montague's Principle J; accordingly, neither would support the punishment of innocent persons.

b. Proportionate Punishment

Except for Farrell's weak retributivism, which explicitly endorses additional, disproportionate punishment, all general deterrence theories of punishment considered here require that punishment be proportionate. In Quinn's theory, the automatic retaliation device is programmed for self-defense; thus, the same principles and requirements that governs measures of self-defense also apply to this self-defensive device. The principle of proportionality is a standard requirement in self-defense. Thus, Quinn's theory has no problem accommodating the principle of proportionality.

Montague's principle of distributive justice would also support proportionate punishment. The amount of harm we may permissibly impose on an offender is the amount of harm that they imposed on their victims. We are allowed only to redistribute the harm that the offender created back to them, and we are not allowed to create additional harm. Thus, punishment must be proportionate to the crime committed.

Tadros argues that the amount of punishment that an offender ought to receive is determined by the duty that they have to protect their victims against their attack. This duty is in turn determined by the amount of harm they created via their offense. Offenders are to be punished so as to send a warning message to potential offenders. Accordingly, punishment must be proportionate to the crime committed.

On the other hand, Farrell's weak retributivism allows us to punish offenders more than is required to deter their offenses; this is because the harms generated by an offense are not limited to the harms created directly by the offense, but also the harm that are indirectly caused by it. Does this entail that we may impose disproportionately severe punishment on an offender? Farrell says yes, we may punish offenders more severely than their crimes. Disproportionate punishment can be morally justified.

c. Using Persons as Mere Means

In Quinn's theory, although the institution of punishment is created to deter criminal offenses, punishment is not implemented until some offender actually commits a crime. Punishment is imposed on actual offenders because they ignore the threat of punishment, commit crimes, and

activate the automatic retaliation process. Thus, their punishment is the result of their criminal wrongdoing and is not meant to teach anyone else a lesson. Offenders are punished because they activate the device. They are not used as a mere means (Ellis 2003).

In both Montague's and Farrell's theories, the aim of punishment is to redistribute the harm back to the offender. Accordingly, neither would require offenders to be punished for reasons other than their own offense. Accordingly, when punished, offenders are not treated as mere means.

On the other hand, Tadros argues that offenders can be used as mere means. Normally, we should not use people as mere means; Tadros calls this *the means principle*. However, wrongdoers may be used as a means to protect their victims against the harms that they created (2011:193). This is because they violate their duty not to harm people, and thereby acquire a duty to protect their victims. Further, if two offenders cannot protect their own victims against the harms that they created, but are in positions that allow each of them to protect the other's victim, then Tadros argues that they have a duty to protect the other's victim, so long as their duties are similar in kind and stringency (2011: 275).

Next, I shall consider specific deterrence theories of punishment.

IV. Specific Deterrence Theories

A specific deterrence theory holds that punishment can be morally justified because punishment serves to deter specific offenders from committing crimes. According to Erin Kelly, "Specific deterrence involves the application of sanctions to a particular person to induce that person to refrain from engaging in harmful behavior. It is contrasted with general deterrence, in which we harm a person in order to deter another (2018:126)." In other words, the intended audiences are different for general and specific deterrence theories. By punishing an offender, general deterrence theories intend to deter people other than the offender, while specific deterrence theories aim to deter offenders themselves.

In the contemporary literature, only two philosophers support a specific deterrence theory—Erin Kelly and me. Next, I shall consider these two views.

i. Erin Kelly—Just Harm Reduction

The fundamental moral principle that guides Kelly's account is *fairness*. According to Kelly, the institution of criminal punishment is created to guide people's conduct. Criminal offenses are behaviors that people have no right to engage in; thus, attaching punishment to criminal offenses does not impose any unfair burden on people. Further, a competent adult enjoys fair opportunities not to commit a crime; if not, they would be excused and not be punished. One can avoid punishment simply by not committing any crimes (2018:135). The threat of punishment gives people a rational incentive not to commit crimes; thus, criminal punishment contributes to the aim of harm reduction. If one chooses to commit a crime, then the threatened punishment will be imposed on them. The institution of punishment, accordingly, is designed to convince everyone not to commit crimes. Punishment is imposed on actual offenders only.

An interesting feature of Kelly's account is that while she appeals to a specific deterrence theory to justify the institution of punishment, she adopts a general deterrence theory to select punishment policies. One might think that the specific deterrence theory of punishment would require that we make sentencing decisions by considering how effectively a form of punishment deters offenders. Because convicted offenders are different, presumably, the correct punishment for them would also be different. Further, because different offenders can be deterred by different types of punishment, the specific deterrence theory cannot support a system of standardized punishment for all offenders who commit the same type of crime.

However, this is not how Kelly thinks about sentencing policies. The principle that guides Kelly's account is fairness, and fairness requires that offenders who commit the same type of crime receive the same type of punishment. According to Kelly,

[R]easonable opportunity generalizes across persons. Fairness requires us to evaluate and respond to individual infractions with standards that extend to relevantly similar cases. We should treat like cases alike, or at least with reasonable similarity. Both positivist and non-positivist legal philosophers have recognized this principle. H.L.A. Hart refers to it as part of the minimum moral content of law, and Ronald Dworkin calls it a matter of integrity in the law (2018: 135).

How much punishment should be imposed on offenders, then? The amount of punishment is to be determined by the deterrence value of punishment, and the deterrence value of a form of punishment is determined by its *general* deterrence value, not by its impact on individual convicted offenders. That is, we are to calculate the deterrence value of punishment by considering how it deters *an average offender* (2018:136). According to Kelly, "in calculating our threats to deter people from reoffending we are, in effect, calculating the general deterrence value of the punishment (2018: 137)." After calculating how much punishment is needed to deter people from committing a particular type of crime, we may impose the estimated punishment so as to deter people from committing the same crime. Thus, according to Kelly's account, the justification for punishment is specific deterrence, but the effect is general deterrence.

Although Kelly appeals to a specific deterrence theory, her fairness view supports punishment policies that are based on a general deterrence theory. This is an interesting feature of her theory, but also a problematic one. Her argument assumes that the same standardized punishment should be used to deter different criminal offenders. However, because different criminal offenders commit the same crime for different reasons, it is not clear how imposing the same punishment on them, which is designed to deter "an average offender," can deter all of them. Some offenders commit crimes because they have false beliefs about facts or values. To deter them, we must help them to acquire true beliefs; otherwise there is no reason to expect that they will not commit a crime again. Others have problems controlling their emotions, desires, or conduct. To deter them, we must help them to control themselves; without such efforts, they are likely to commit crimes again, given their personality traits and tendencies. Still others lack the capacity to understand the impact of their actions, including the possibility that they will be punished. To deter them, we need to explore possible options and solutions (Lee 2017). Accordingly, to deter individual offenders from committing crimes, we must examine the reason why they committed crimes in the first place and then learn how to prevent them from committing crimes again in the future. This requires that we invest some time to learn the background of individual offenders—their beliefs and values, whether they can be persuaded into changing their beliefs and lifestyles, and how we may help them to lead a life that is crime free. The punishment policy that Kelly proposes—to impose punishment according to their deterrence impact on an average offender—does not support the aim of specific deterrence, as it fails to address individual offenders.

ii. Hsin-Wen Lee—The Rights-Protection Theory

I defend a societal self-defense theory of punishment, the rights-protection theory (2018).

According to this theory, the fact that we have certain core rights (to life, liberty, and property), entails that we also have derived rights to take measures to protect these core rights. That is, the fact that we have rights at all gives us a further right to defend these rights. We may take self-defensive measures against encroachment upon our rights; otherwise our rights would be empty.

This right of self-defense in turn gives us a right to create the institution of criminal punishment to sanction those who violate our rights. The institution of criminal punishment is one of the measures that we can take to protect our core rights. Because the right to punish is derived from the right of self-defense, which in turn is derived from our core rights, the same constraints that apply to the right of self-defense also apply to criminal punishment.

My rights-protection theory takes the protection of our core rights to be the fundamental value that provides the ultimate moral justification for the institution of criminal punishment. Accordingly, the institution of punishment is created to protect the core rights of individuals. Therefore, in selecting punishment policies or designing the institution of punishment, we must make sure that the policies and institutions will not undermine the fundamental value of rights-protection; further, we must prioritize those policies and institutions that serves to protect rights. In other words, the aim of rights-protection both justifies and constrains the institution of punishment.

Because the institution of punishment is created to serve the aim of rights-protection, we ought to prioritize punishment policies that contribute to this purpose. If a punishment policy or decision is likely to undermine this aim, then we ought to adopt some alternative measures to ensure that punishment does not undermine its justifying aim. This means that if forms of punishment are more likely to reduce crime and recidivism rates—e.g., job-training programs, drug-rehabilitation programs, psychological or psychiatric treatments, etc.—then we ought to adopt these forms of punishment, as they help to protect citizens’ rights. On the other hand, if a form of punishment is likely to increase crime or recidivism rates, then we ought to use it as a last resort in punishment. For instance, empirical studies show that certain forms of punishment tend to increase the recidivism rate—e.g., imprisonment tends to have a prisonization effect, which in turn makes recidivism more likely—then we ought not adopt imprisonment as the default form of punishment and should use it only when it is necessary.

Further, the constraints on the right of self-defense apply equally to the institution of criminal punishment. Some commonly accepted constraints on the right of self-defense include discrimination, proportionality, necessity, and minimum rationality.¹ Very briefly, discrimination requires that we punish the actual offender and no one else. Proportionality requires that in defending ourselves against an offense, other things being equal, we adopt self-defensive measures that are proportionate to the harm created by the offense. Necessity requires that the measures that we adopt to defend ourselves be causally connected to the ending of the aggression. Minimum rationality requires that the defensive measures that we adopt not to harm our aim of rights-protection. That is, in punishing to protect our rights, we ought not to adopt self-defensive measures that will undermine the rights we intend to protect (Lee, forthcoming).

Specific deterrence requires that we consider how the punishment of an offender increases or decreases their chance of recidivism. This requires that we analyze the type of offender one is and consider how to come up with an individualized plan to help the offender lead a normal crime-free life in the community. The imposition of proportionate harm on offenders does not necessary contribute to this aim and is not an essential aspect of punishment.

iii. On Specific Deterrence

¹ For a detailed analysis, see section I of my “Chapter Five—Self-Defense Theories of Punishment—An Examination from the Perspective of Means–Ends Rationality,” in *Criminal Punishment and Rights-Protection*, forthcoming.

Before moving on to the next section, I shall consider briefly how the two specific deterrence theories respond to common objections to consequentialist theories of punishment.

Because specific deterrence theories of punishment aim to deter actual criminal offenders, punishment is not imposed for the sake of deterring other potential offenders, but for the actual offenders themselves. Thus, specific deterrence does not use offenders as mere means. Neither does the specific theory require the punishment of innocent persons because punishment of innocent persons does not help to deter the actual offenders.

Do specific deterrence theories support disproportionate punishment? Kelly specifically argues that sentencing policies must address an average offender. Thus, her theory requires proportionate punishment. My theory, on the other hand, sees proportionate punishment policies as having expressivist value, but not ultimate value. That is, punishment ought to be proportionate as we want to express equal condemnation to the same types of offenses. However, I do not believe that this requires that offenders who commit the same type of crimes to receive the same punishment or be treated in the same way. The aim of specific crime deterrence is to deter individual offenders, not an average offender. If the same punishment deters one offender but not the other, there is no need to insist that they be punished in the same way. If the same punishment does not deter one of them, we must seek alternative forms of punishment. For instance, if a monetary fine deters one offender but not the other, then we must seek alternative forms of punishment—e.g., community service, short-term imprisonment, etc.

Further, if statistical evidence shows that a form of punishment leads to a higher recidivism rate, then, although my theory would not reject this particular form of punishment entirely, it would regard it as the last resort, not the first or the standard resort, of punishment. For instance, if there is a serious negative impact of imprisonment, as studies consistently show, then we ought no longer to use imprisonment as the standard form of punishment and ought to use it only as a last resort for punishing violent offenders. This means that even if we were to insist on proportionate punishment, we must adopt other alternative forms of punishment which could deliver proportionate punishment—open prisons, community service, etc.

V. Preventionism

Many believe that it is more important to prevent crime than to punish offenders (Ashworth and Zedner 2014: 29). Instead of waiting for crimes to happen and then punishing offenders, we ought to take active steps to prevent crimes from being committed. In the literature, preventionism refers to a specific view about punishment which sees additional punishment as justified because it serves to prevent offenders from committing crimes.² This view is different from deterrence theories of punishment, which hold that the institution of criminal punishment is morally justified because having such an institution helps to deter criminal offenses.

² For example, Bernard Harcourt suggests that, according to consequentialism, “punishment was a central part of prevention: it was, for instance, fully justified to lengthen a sentence (to punish more) for someone who recidivated because the recidivist carried a higher likelihood of reoffending (2013: 258).” Similarly, Kevin Arthur suggests, “detention designed to protect society from predicted but unconsummated offenses does not increase the likelihood of a fair trial. Accordingly, such detention is not simply regulatory (1987: 403-404),” and “A restraint on liberty such as preventive detention is regulatory rather than punitive only if it serves a legitimate and compelling’ state purpose. The court decided that crime prevention was such a purpose (1987: 396).”

Preventionism holds that, if we know that some offenders are dangerous—for instance, they are repeated offenders who are likely to commit crimes again—then the court may impose sentences on them that were longer than the ones that they would have received if they were not likely to reoffend. According to Frederick Schauer, “Preventive detention... involves punishing people for the harmful acts in which they might engage rather than for the harmful acts in which they have already engaged (2013:3).” Consequentialists and deontologists alike generally believe that punishment must be proportionate to the crime committed. Thus, if a punishment is more severe than the crime committed, then it is disproportionate and cannot be right. However, according to preventionism, a longer sentence can be morally justified if the extended sentence serves to prevent the offender from committing more crimes. Under this view, a longer sentence can be morally justified because it serves to incapacitate the offender, thereby preventing them from committing additional crimes.

Notice that this view says nothing about the punishment of non-recidivist offenders who are not likely to reoffend. It tries to justify extended sentences for recidivists only. Thus, the theory may be combined with a retributive or a deterrence theory of punishment; they would provide us with instructions on how to punish offenders who are not likely to reoffend. Preventionism is primarily concerned with the punishment of dangerous and repeated offenders.

However, because offenders have yet to commit the crime that we suspect that they would, many are concerned with the moral soundness of preventionism. Punishment is justified for its expected preventive effect. However, there are reasons to doubt that a longer sentence serves its intended purpose. First, this argument assumes that the offender will commit a crime again, and thus needs to be prevented from doing so. However, it is not clear how this belief is justified. During an ideal trial, one sees the evidence for an offender’s crime; however, it is not clear if the evidence can provide a good reason for believing that they will commit a crime again in the future. If the evidence is sufficient to support the belief that they will commit a crime, then, very likely, the court may punish the offender for an inchoate crime—attempt, solicitation, and conspiracy. If the evidence is not sufficient to support the belief, then it is not clear why we can assume that they need to be prevented from doing so. Preventionists are likely to refer to an offender’s past recidivism to support the belief in their future re-offense. This would essentially be a type of inductive argument. Accordingly, the truth of the premise does not guarantee the truth of its conclusion. After all, it is possible that some incident happens to the offender before they have a chance to commit another crime.

Further, some are concerned that giving an offender a longer custodial sentence means offenders will receive more punishment than they deserve, or more than what is proportionate to the crime they committed (Ashworth and Zedner 2014: 151). They argue that the amount of punishment should be determined by the seriousness of the crime. However, preventionism allows preventive considerations to play a role in determining offenders’ sentences. Many are concerned that such a system of punishment is unjustified.

In addition, it is not clear why we should believe that simply having a longer sentence will prevent an offender from committing crimes again. First, it is important to acknowledge that custodial sentences do serve to incapacitate offenders. So long as an offender stays in prison, they cannot commit crimes that would harm the community (though they can still act in ways that harm other inmates or correctional officers). Thus, the community is safer and people’s rights are protected during the time of their extended stay. However, unless the offender will stay in prison until their death, there will be a day when they are out of prison. If they never change their mind, then having a longer sentence prevents them from committing a crime only during

the period of time when they are in prison; they can still commit a crime after they get out. That is, an extended sentence serves only to delay the crime, not to stop the crime.

Second, if we really care about preventing crime, we must consider how a longer sentence affects crime prevention. First, there is a risk that an offender will be subject to more bad influences in prison. While staying in prison, one must learn to live with other inmates and acclimate oneself to the prison culture. Depending on what other inmates are like, one may make friends with some of them and then learn about committing crimes more effectively. This is problematic if we are concerned with crime prevention. Second, having a longer sentence effectively means it will be more difficult for the offender to re-enter society—they lose contact with their friends and family, they are less familiar with how to live in the community, after serving time in prison it becomes more difficult for ex-inmates to find housing, jobs, a support system, etc. Longer custodial sentences effectively mean that their lives outside prison will be more difficult, not to mention that having a criminal record already makes one's life more difficult. If their re-entry back into the community is difficult, then they may be motivated to go back to their old ways and commit crimes again. This is, again, quite bad for the sake of crime prevention.

In other words, simply having a longer sentence does not necessarily serve to prevent an offender from committing additional crimes. Setting aside the issue of whether offenders will commit crimes again, to ensure that they will not, the criminal justice system must take active steps to ensure that offenders are not likely to commit crimes again in the future. Elsewhere I argue that there are different types of offenders and that different strategies must be employed to deter them (Lee 2017). For example, some offenders are not concerned about the consequences of their action because of they lack proper education or financial support; to prevent them from committing crimes again, it is important to make sure that they understand the impact and significance of their actions and to make sure that they have the opportunity to earn an adequate income. Other offenders might require education, training, or treatment. There is no perfect way to ensure that offenders will not commit a crime—education, training, and treatment programs may or may not help. Nevertheless, simply putting offenders in prison for a longer period of time can only deter those who strongly dislike prison; it may not deter those who do not care or cannot care.

VI. Conclusion

In this chapter, I consider consequentialist theories of punishment. I begin by considering two types of consequentialist theories of punishment—teleology and aggregationism. After explaining why a teleological theory is implausible, I consider four general deterrence theories, two specific deterrence theories, and preventionism. I argue that the policies supported by general deterrence theories fail to serve the aim of crime deterrence, and preventionism does not serve the aim of crime prevention. Kelly's specific deterrence theory, namely just harm reduction, also fails to serve the purpose of specific deterrence. Only my rights-protection theory supports an institution of punishment that is designed to deter crime.

References

- Ashworth, Andrew and Lucia Zedner. 2014. *Preventive Justice*, Oxford University Press. Paperback 2015.
- Bentham, Jeremy. 1780. *An Introduction to the Principles of Morals and Legislation*. Mineola, New York: Dover Publications, Inc., 2007.
- Bentham, Jeremy. 1830. *The Rationale of Punishment*, James T. McHugh (ed.), Amherst, New York: Prometheus Books, 2009.
- Duff, R Anthony. 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press.
- Ellis, Anthony. 2003. "A Deterrence Theory of Punishment." *The Philosophical Quarterly* 53, no. 212 (July): 337-51.
- Farrell, Daniel M. 1985. "The Justification of General Deterrence." *Philosophical Review* 94, no. 3 (July): 367-94.
- Farrell, Daniel M. 1989. "On Threats and Punishments." *Social Theory and Practice* 15, no. 2: 125-54.
- Feinberg, Joel. "The Expressive Function of Punishment." *The Monist*, Vol. 49, no. 3, 1965: 397-423.
- Harcourt, Bernard E. 2013. "Punitive Preventive Justice: A Critique." in *Prevention and the Limits of the Criminal Law*, Andrew Ashworth, Lucia Zedner, and Patrick Tomlin (eds.), 252-72. Oxford: Oxford University Press.
- Kelly, Erin. I. 2018. *The Limits of Blame—Rethinking Punishment and Responsibility*. Cambridge, MA: Harvard University Press.
- Kershnar, Stephen. 2000. "A Defense of Retributivism." *International Journal of Applied Philosophy* 14, no. 1: 97-111.
- Kleinig, John. 1973. *Punishment and Desert*. The Hague: Martinus Nijhof.
- Lee, Hsin-Wen. 2017. "Taking Deterrence Seriously—The Wide-Scope Deterrence Theory of Punishment." *Criminal Justice Ethics* 36, no. 1 (April): 2-24.
- Lee, Hsin-Wen. 2018. "A New Societal Self-Defense Theory of Punishment—The Rights-Protection Theory." *Philosophia—Philosophical Quarterly of Israel* 46, Issue 2 (June): 337-53.
- McCloskey, H. J. 1965. "A Non-Utilitarian Approach to Punishment." *Inquiry* 8: 239-55.
- Mill, John Stuart. 1859. "On Liberty" in *John Stuart Mill On Liberty and Other Essays*, Oxford University Press, New York, 1991.
- Montague, Phillip. 1995. *Punishment as Societal-Defense*, Rowman & Littlefield Publishers.
- Moore, Michael S. 1997. *Placing Blame: A Theory of the Criminal Law*. Oxford: Oxford University Press.
- Quinn, Warren. 1985. "The Right to Threaten and the Right to Punish." *Philosophy and Public Affairs* 14, no. 4 (Autumn): 327-73.
- Von Hirsch, Andreas. 2017. *Deserved Criminal Sentences: An Overview*, London: Bloomsbury.
- Schauer, Frederick. 2013. "The Ubiquity of Prevention." In *Prevention and the Limits of the Criminal Law: Principles and Policies*, Andrew Ashworth, Lucia Zedner and Patrick Tomlin (eds.), 10-22. Oxford: Oxford University Press.
- Strawson, P. F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48: 1-25.
- Tadros, Victor. 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*, New York: Oxford University Press.