

Against Retributive Justifications of the Death Penalty

Sarah Roberts-Cady

Almost any textbook discussion of the death penalty will state that there are two common ways to defend it: by appeal to social utility or by appeal to retributive justice. The appeal to social utility is often called a “forward-looking” justification for punishment, since it attempts to justify the death penalty by appealing to the social benefits of that punishment, specifically the prevention of crimes. The appeal to retributive justice is often called a “backward-looking” justification for punishment, because it focuses on a fair response to past actions. A retributivist would argue for the death penalty based on the claim that death is what a murderer deserves. Most discussions of the death penalty move out from there, either debating the social utility of the death penalty or arguing about whether or not assigning the death penalty is giving people what they deserve. What is insufficiently debated is the relevance of retributive justice to the discussion. While retributive theories do offer some answers to questions about punishment, I will argue that they are not sufficient to establish whether or not the death penalty is a morally acceptable form of punishment.

I will not pretend to be the first to make this claim. Hugo Adam Bedau once wrote, “[R]etributive considerations may be sufficient to tell us *who* deserves to be punished—it is the guilty. But retribution fails to tell us *what* they deserve as their punishment.”¹ What I hope to accomplish in this paper is to provide a clear and systematic defense of this claim. Bedau’s assertion is based on a distinction between two kinds of questions a theory of punishment may attempt to answer: (1) Is punishment a morally justified response to wrongdoing? (2) How does one determine what particular punishments ought to be assigned to particular wrongdoers? Though one may dispute whether or not a theory of retributive justice gives a clearly defensible answer to the first question, no one disputes that it does in fact offer an answer to that question. But answering the second question is the only way to determine whether or not the death penalty is a fitting punishment for certain kinds of crimes. Bedau claims that a theory of retributive justice cannot answer this second question. Nevertheless, most of the thoughtful people involved in the public and academic debates about the death penalty seem to assume that theories of retributive justice offer an answer to the question of the moral permissibility of the death penalty. I hope to clarify why this is a mistake.²

The term “retributive justice” has been associated with a wide variety of views. In a famous article entitled “Varieties of Retribution,” John Cottingham noted that philosophical discussions of retributive justice often involve disparate

and muddled conceptions of retribution.³ In the paper, he distinguished nine different views that have been linked with retributive justice. He argues that many of these views are either implausible or improperly associated with retributive justice. As such, I will not attempt to address all nine views here. Instead, I will respond to Cottingham's call for clarity by stating that I am defining "retributive theory" (in the standard way) as those theories that accept the desert thesis. According to Cottingham, the desert thesis is the position that "punishment is justified because it is a deserved requital or reward for wrongdoing."⁴ This claim is widely accepted, and yet surprisingly difficult to justify philosophically. Many retributive theorists simply take it as a foundational principle which is grasped as true by intuition and without need of justification. I call this the "axiom" version of the desert thesis, since it takes the desert thesis as a self-evident starting point. This appears to be the position of J. D. Mabbot, Igor Primoratz, and J. L. Mackie, among many others.⁵ Alternatively, some retributivists offer theoretical arguments in support of the desert thesis. One common way of doing this is by appeal to fair play theory. This view takes the desert thesis as a secondary principle, derived from considerations of distributive justice. Fair play theorists hold that those who commit crimes gain an unfair advantage over the law-abiding citizens. Herbert Morris, one of the most influential fair play theorists, explains it this way: "A person who violates the rules has something others have—the benefits of the system—but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage. Matters are not even until this advantage is in some way erased."⁶ On this view, punishment is a way of removing that unfair advantage. In addition to Herbert Morris, philosophers George Sher, Michael Davis, and many others have endorsed versions of fair play theory.⁷ In any case, whether the desert thesis is taken as an axiom or defended as a secondary principle, these retributive theorists agree that punishment in general is justified by appeal to considerations of what the wrongdoer deserves.

How do these retributive theorists answer the second question about which punishments ought to be assigned in response to particular crimes? One common answer given by retributive theorists of all kinds is that the punishment ought to be equivalent to the crime. Immanuel Kant famously endorsed this standard, calling it the "principle of equality."⁸ It is also sometimes referred to by the Latin phrase "*lex talionis*," which means "the law of retaliation."

A literal version of this would state that a punishment should be exactly the same as the crime committed. This is the "eye for an eye" principle. That is, if a person steals, that person ought to have an equivalent amount of personal belongings taken from them. If a person murders, that person ought to be killed. While this interpretation of the principle of equality certainly would justify the death penalty for people who commit murder, it is widely regarded as implausible. There are two main problems with it. First, as many people have pointed out, returning the exact same harm for harm is sometimes impossible. For example, how would one have punished Hitler? One cannot commit genocide against a single individual. Second, this standard seems to imply that any punishment which

is equivalent to the crime is morally acceptable. It would entail that it is appropriate to rape the rapist or torture the torturer. Unless one is willing to accept this unsavory consequence, one has reason to reject this standard.

How else might a retributive theorist interpret the principle of equality? Tom Sorrell argues that “Equality can mean similarity of crime and punishment, but it need not mean exact similarity.”⁹ He attributes to Kant the position that punishment should be “equal to or in keeping with the loss to the victims.”¹⁰ It should be an equivalence of loss—the criminal should lose as much as the victims lost. He adds that there is no general rule for determining equivalence, just “good judgment” about what is a roughly equivalent loss or when exact equivalence is appropriate.¹¹

It seems likely the person of good judgment will run into the following problem: For certain heinous crimes, either the punishment would be roughly equivalent and immoral, or it would not be roughly equivalent. Neither death by injection nor life in prison could match the losses of the victims of the most horrible types of crimes. Undoubtedly, anything that is even roughly equivalent to the losses of young children who have been brutally tortured, raped, and murdered would also be immoral to inflict. But the problem is deeper than a mere lack of morally acceptable and equivalent punishments. Unless you accept all forms of punishment, no matter how severe, you need to appeal to some other moral standard to determine when a punishment that is roughly equivalent to the crime is morally appropriate. That is, the retributive standards alone cannot justify the use of the death penalty. One needs some other moral standard to determine which roughly equivalent punishments are or are not morally acceptable. For Sorrell, this other standard is “good judgment.” But vagueness will not save him. All that this means is that the bulk of his argument for the death penalty rests on his “good judgment” that capital punishment is a punishment that is both a roughly equivalent to some crimes and morally acceptable. Retributive theories are supposed to be the way to justify capital punishment. But Sorrell’s version of this cannot do this. He has to appeal to his “good judgment” that the death penalty is morally justified before he can use it in this retributive scheme.

Like Sorrell, Igor Primoratz believes that the principle of equality can be understood as advocating some sort of rough equivalence between crime and punishment. But he adds that, although there may be rough equivalence between fines and theft, or assault and prison, the only equivalent for murder is the death penalty. This is because life is invaluable; there is no amount of money or prison that could equal it.¹² This seems to be exactly what Kant argues in the following passage,

If, however, he has committed a murder, he must die. In this case, there is no substitute that will satisfy the requirements of legal justice. There is no sameness of kind between death and remaining alive even under the most miserable conditions, and consequently there is no equality between the crime and the retribution unless the criminal is judicially condemned and put to death.¹³

One problem with this interpretation is that it seems inconsistent with Kant's and Primoratz's own views about the worth of persons. After all, being invaluable means being without price, as Kant would claim, and thus having incomparable worth. If life has unconditional, incomparable worth, then we cannot and should not compare the worth of one person's life with that of another. According to Kant's own moral theory, such comparisons are simply a confused and immoral way of looking at the value of human lives. Thus, taking the murderer's life cannot be said to be even roughly equivalent to taking the life of the victim. But there are further complications. Suppose one could compare lives in this way. Is death by injection really roughly equivalent to brutal murder? Brutally killing the murderer might be even closer to equivalent. And the principle of rough equivalence provides no reason not to brutally murder the brutal murderer. But I do not think many people would advocate such brutal murder. So one needs some other standard of morality to determine what roughly equivalent punishments are morally acceptable punishments. We are back to the same problem: Either one must include all kinds of radical punishments or appeal to a moral standard besides retributive theory to determine what punishments are acceptable.

An alternative way of understanding the principle of equality would be to think about it in terms of the currency of freedom. Fair play theorists like Herbert Morris and George Sher explain it this way. Recall that a fair play theorist holds that the lawbreaker gains an unfair advantage when she accepts the benefits of other people's self-restraint (e.g. security) without also accepting the burden of her self-restraint. This unfair advantage comes in the form of added freedom. Sher explains it this way,

... a person who acts wrongly does gain a significant measure of extra liberty; what he gains is freedom from the demands of the prohibition that he violates. Because others take that prohibition seriously, they lack a similar liberty. And as the strength of the prohibition increases so too does the freedom from it which its violation entails. Thus, even if the murderer and the tax evader do succumb to equally strong impulses, their gains in freedom are far from equal. Because the murderer evades a prohibition of far greater force ... his net gain in freedom remains greater. And for that reason, the amount of punishment he deserves seems greater as well.¹⁴

Sher goes on to clarify that the form of the punishment is determined in the following way,

Because the wrongdoer has unfairly gained an extra measure of freedom from moral restraint, the natural way to restore a fair balance is to reduce the protection he ordinarily would have gained through moral restraints on the conduct of others. By treating the wrongdoer in what is ordinarily a forbidden way, we strip away part of the protection that moral restraints on behavior would ordinarily have afforded him. Thus, we remove precisely the sort of advantages he has gained.¹⁵

In sum, those who commit crimes gain the benefit of other people's self-restraint without the burden of a restraint on their own freedom. Punishment is designed to

remove that unfair advantage. To restore the balance, the wrongdoers lose the protection afforded by other people's restraint, taking away the advantages they have unfairly gained.

One might object to how fair play theory justifies punishment in general. Numerous important questions have been raised about this: Is it reasonable to claim that the freedom to rape and murder is an "advantage"? Is refraining from rape and murder a "burden" for most people? In what sense does a lawbreaker gain freedom that other citizens do not have? Doesn't everyone have the same freedom to break the law (and face the consequences) even if most choose not to exercise that freedom? Does fair play theory conflict with the Hegelian/Kantian view that those who act morally are freer than those who act immorally (because they are free from the inclinations that enslave the rational will)?¹⁶

Setting aside questions of the coherence of the fair play theorist's answer to how punishment is justified, let us turn to the central issue of this paper: the way fair play theorists determine which punishments ought to be assigned in response to wrongdoing. Sher's account faces the same problem as previous accounts. If a person is guilty of rape, Sher's theory implies that restoring the balance would be best achieved by allowing others (the victim or the victim's representative, perhaps) to set aside their usual restraints in order to rape the wrongdoer. This would be the most precise way of removing the unfair advantage gained by the criminal; it removes the advantage gained by the criminal from other people's restraint in precisely the way the criminal's absence of restraint created an added burden for others. It is doubtful that Sher would advocate that punishment. But his retributive considerations alone cannot explain why. He faces the same dilemma as the other theorists we have discussed: Either Sher must accept that any way of restoring the balance is morally acceptable, or he must appeal to some other moral considerations to determine which ways of restoring the balance are morally acceptable.

Although Sher does not suggest this, a fair play theorist could understand the principle of equality in an alternate way. One could argue that the wrongdoer deserves punishment which restricts her freedom in an amount equal to the amount of freedom she unfairly gained by violating the law (and perhaps also considering the freedom unfairly taken from a victim). This interpretation fits nicely with the commonly held view that prison is a form of restricting the freedom of the criminal. Using fair play theory, one could explain how theft or rape could be converted into the currency of freedom and shown to be rightly punished by a certain number of years in prison. In addition, one could argue that the death penalty is the only way to restrict the freedom of the murderer as severely as the murderer restricted the freedom of the victim.

The "equivalent restriction of freedom" interpretation of the principle of equality suffers from the same problem as the last interpretation. It seems likely that some ways of restricting freedom are not morally acceptable. For example, in fourteenth-century England, it was common to punish a person by placing him in stocks. Most people today do not think that is a morally acceptable punishment.

Yet, on this interpretation of the principle of equality, either one would have to accept that all methods of restricting freedom are morally acceptable, or one would need some other moral criteria to establish which restrictions of freedom are or are not acceptable. In other words, one would have to have already established that the death penalty is a morally appropriate way of restricting freedom on some other moral grounds before one could use it in this retributive model.

Many retributive theorists reject the principle of equivalence, suggesting instead that punishment ought to be proportional to the seriousness of the crime. This is grounded on the assumption that crimes and punishments can be ranked by degree of seriousness. This view of punishment requires that more serious crimes be punished with more serious punishments. The comparison is no longer between a single crime and a single punishment. Instead, one must look at each crime and punishment relative to other crimes and other punishments. The principle requires that if you commit a more serious crime than your neighbor, you get a more serious punishment than your neighbor. For example, if Joe Shmoe steals \$10,000, this version of the principle of equality would require that Joe get a more serious punishment than someone who stole ten dollars, and a less serious punishment than someone who murdered a person. It does not entail that Joe's punishment be equivalent (or even roughly equivalent) to his theft.

As many people have pointed out before, arranging punishment by degree of seriousness does not require including the death penalty on the scale. One could place life in prison at the top of the scale as the most serious punishment, while still fulfilling the requirement that the more serious crimes merit more serious punishments. But why not include the death penalty on the scale? It seems clear that one would not include every possible punishment on the scale of actual punishments. If the scale included every possible punishment from the mildest to the most severe possible, the top of the scale would not be death by injection, but something worse—perhaps lifelong severe torture ending in some kind of painful and humiliating death. No government with moral sensibility would allow that punishment. Some forms of punishment would be excluded from the scale on moral grounds. Thus, to justify capital punishment on the proportionality view, one would need to appeal to some additional ethical criteria for determining which punishments are or are not appropriately included on the scale of punishment. In other words, a discussion of whether or not the death penalty is an acceptable punishment would require an appeal to moral considerations outside of retributive justice.

This is in part what was argued by Claire Finkelstein in "Death and Retribution," and then admitted by her respondent Michael Davis, in "A Sound Retributive Argument for the Death Penalty."¹⁷ Davis accepts the proportionality standard for punishment and then attempts to offer a supplementary argument for why the death penalty should be included on the scale as an appropriate punishment. He does not think it weakens the retributive position to be supplemented in this way. Davis's argument, briefly, is that the standard for acceptable punishment should be determined by what is shocking or not in a particular culture. If, in a particular

society, most people find a punishment shocking, then that punishment is inhumane. Davis adds a Kantian twist to his theory by arguing that treating someone in a way that is generally found shocking is failing to treat her as a person.

Setting aside the question of whether or not Davis's standard for acceptable punishment is adequate, I will focus on the broader question of the significance of supplementing a retributivist theory with another theory in order to justify capital punishment. I agree with Davis that there is nothing wrong with adding a moral argument to the retributivist stance in order to justify the death penalty. Indeed, I think that is exactly what a retributivist must do in order to defend the death penalty successfully. But the point is that the retributivist argument does not do the work in morally justifying the death penalty; some other moral theory has done this work. The retributive arguments cannot establish the appropriateness of the death penalty (as many people regularly claim). Rather, the moral acceptability of the death penalty must first be established by means of some other moral criteria before it can even be considered as a part of the retributivist's scale of punishment. Thus, the conversation must change from a discussion of retribution to some other moral discussion. That is exactly what I am advocating happen with regard to the death penalty debate. We must move beyond a discussion of retribution in order to resolve the death penalty debate.

This article does not challenge the coherence of retributive theory nor does it challenge the consistency of a retributive theorist who supports the death penalty. I have only argued that one cannot justify the death penalty simply by establishing the claim that wrongdoers deserve punishment which fits the crime. Unless one is willing to condone all sorts of barbaric punishments, then one must appeal to additional ethical considerations to establish which equivalent (or roughly equivalent or proportional) punishments are morally acceptable. This is more than a challenge to retributive theorists to "fill out" their theories; the way in which these theories must be "filled out" requires a change of subject. Claims about the fit between crime and punishment will not settle the debate. This is because most of us believe that there are some punishments that no human deserves or, alternately, that there are some punishments that, while deserved, one ought not to inflict. To sort out how to explain this philosophically, we must discuss broader ethical issues than a fit between crime and punishment. Perhaps we should discuss what is respectful of human dignity or how punishment affects the character of those who impose it. Whatever direction the discussion of the death penalty takes, it must lead beyond a discussion of retribution in order to lead us further toward resolving this issue.

Notes

¹Hugo Adam Bedau, "The Minimal Invasion Argument Against the Death Penalty," *Criminal Justice Ethics* 21 (Summer/Fall 2002): 7.

²It is worth distinguishing the view I will be defending from other challenges to the retributive justifications of the death penalty. Thom Brooks, Daniel McDermott, and Stephen Nathanson all

argue that empirical examples of unfairness in the system (either the unfair distribution of the death penalty among the guilty or the occasional conviction of the innocent) undermine the retributive justifications of the death penalty. While I think these arguments are important and worth considering, they are very different from my own arguments. In this paper, I will be attempting to argue that even if the death penalty were distributed fairly among the guilty and with due process that protects the innocent, retributive considerations alone would not suffice to justify the death penalty. I am not just challenging whether or not retribution is being achieved by assigning the death penalty; I am challenging the relevance of retributive considerations to the death penalty discussion. For a discussion of these other challenges to the retributive justifications of the death penalty see: Thom Brooks, "Retributivist Arguments against Capital Punishment," *Journal of Social Philosophy* 35 (Summer 2004): 188–97; Daniel McDermott, "A Retributivist Argument Against Capital Punishment," *Journal of Social Philosophy* 35 (Fall 01): 317–33; Stephen Nathanson, "Does it Matter If the Death Penalty is Arbitrarily Administered?" *Philosophy and Public Affairs* 14 (Spring 1985): 149–64.

³ John Cottingham, "Varieties of Retributivism," *The Philosophical Quarterly* 29 (July 1979): 238–46.

My thanks to an anonymous reviewer who pointed me to this paper.

⁴ *Ibid.*, 239.

⁵ J. D. Mabbot, "Punishment," *Mind* 48 (April 1939): 152–67; Igor Primoratz, *Justifying Legal Punishment* (Atlantic Highlands, NJ: Humanities Press, 1989): 158–59; J. L. Mackie, "Morality and the Retributive Emotions," *Criminal Justice Ethics* 1 (Winter/Spring 1982): 3–10.

⁶ Herbert Morris, "Persons and Punishment," *The Monist* 52 (1968): 478.

⁷ George Sher, *Desert* (Princeton, NJ: Princeton University Press, 1987). Michael Davis, "Criminal Desert and Unfair Advantage: What's the Connection?" *Law and Philosophy* 12 (May 1993): 133–56.

⁸ Immanuel Kant, *Metaphysical Elements of Justice*, tr. John Ladd, 2nd ed. (Indianapolis: Hackett Publishing Company, 1999), 138. It is clear that Kant accepted the desert thesis and the principle of equality, but it is not at all clear how it fits with the rest of his ethical and political theory. There is disagreement in the literature about whether Kant accepts these claims as self-evident axioms. There is discussion about whether or not they are linked to a social contract theory. There is debate about whether or not he is even consistently or completely a retributivist. There is debate about whether or not his retributivist position is consistent with the rest of his moral theory. Indeed, there is disagreement about whether or not Kant even has a complete theory of punishment. I will not address these issues about Kantian interpretation, as they do not affect my thesis. For further discussion of these interesting issues, see Vernon Thomas Sarver, "Kant's Purported Social Contract and the Death Penalty," *Journal of Value Inquiry* 31 (December 1997): 455–72; Sharon Byrd, "Kant's Theory of Punishment: Deterrence in Its Threat, Retribution in its Execution," *Law and Philosophy* 8 (August 1989): 151–200; Thomas Hill, Jr., "Kant on Wrongdoing, Desert and Punishment," *Law and Philosophy* 18 (July 1999): 407–41; Mark Tunick, "Is Kant a Retributivist?" *History of Political Thought* 17 (Spring 1996): 60–78; Jean-Christophe Merle, "A Kantian Critique of Kant's Theory of Capital Punishment," *Law and Philosophy* 19 (May 2000): 311–38; Douglas Lind, "Kant on Capital Punishment," *Journal of Philosophical Research* 19 (1994): 61–74; Jeffrie G. Murphy, "Does Kant have a Theory of Punishment?" *Columbia Law Review* 87 (1987): 509–32; Allen Wood, *Kantian Ethics* (Cambridge: Cambridge University Press, 2008); Michael Clark, "A Non-Retributive Kantian Approach to Punishment," *Ratio* XVII (March 2004): 12–27; Thom Brooks, "Kantian Punishment and Retributivism: A Reply to Clark," *Ratio* XVIII (June 2005): 237–45; Michael Clark, "Kantian Punishment: Rejoinder to Brooks," *Ratio* XVIII (September 2005): 361–64.

⁹ Tom Sorell, *Moral Theory and Punishment* (New York: Basil Blackwell, 1987), 138.

¹⁰ *Ibid.*, 138.

¹¹ *Ibid.*, 150–51.

¹² Primoratz, *Justifying Legal Punishment*, 158–59.

¹³ Kant, *Metaphysical Elements of Justice*, 139.

¹⁴George Sher, *Desert* (Princeton, NJ: Princeton University Press, 1987), 82.

¹⁵*Ibid.*, 84.

¹⁶For a fuller discussion of the philosophical questions raised about fair play theory, see David Dolinko, "Thoughts About Retributivism," *Ethics* 101 (April 1991): 537–59. Dennis Klimchuk, "Restitution and Revenge," *Law and Philosophy* 20 (January 2001): 81–101.

¹⁷Claire Finkelstein, "Death and Retribution," *Criminal Justice Ethics* 21 (Summer/Fall 2002): 12–21; Michael Davis, "A Sound Retributive Argument for the Death Penalty," *Criminal Justice Ethics* 21 (Summer/Fall 2002): 22–26.