

# 1

# Retribution

---

## ■ Introduction

---

Retribution is the notion that punishment is imposed because it is deserved. Murderers are to be given the death penalty because that is the penalty they have earned by their offense. The philosopher Immanuel Kant wrote that retribution is grounded in respect for the autonomy of the offender. If a criminal is punished to deter others from committing a similar crime, then they are being treated as a means to an end. If a criminal is punished too little or not at all, perhaps because of some mitigating factor in his personal history, then the notion of his free will is denied. Modern theorists offer competing formulations of retribution and different bases for its grounding, but all stand in contrast to consequentialist defenses of the death penalty. The death penalty is right or wrong, regardless of whether it deters crime or what its cost. It is right even if it serves no other purpose than to give the criminal his due. As Kant famously put it, “Even if a civil society were to dissolve itself by common agreement, the last murderer remaining in prison must first be executed.” Kant’s teachings are excerpted in the Critical Documents section that follows.

This is unquestionably a powerful idea, and public support for the death penalty is—at least on the surface—largely based on notions of retribution. In most polls, “an eye for an eye” or “punishment should fit the crime” is the plurality reason offered by proponents for their support of capital punishment. In a 2001 Gallup poll, 48% of respondents cited retribution as the basis for their support, more than twice the level of support offered for any other justification.

The concern is the quality of the argument on its own terms. If retribution is the true currency of justice, does it dictate support for the death penalty? Here we ask the horizontal question, as we do throughout the book: should a retributivist support capital punishment? The retributive argument for the death penalty has considerable appeal. Murderers have committed the most serious offense imaginable; they deserve to be treated in the most severe manner.

One objection to retributive theory is that it presumes a baseline condition of equality in society that is not present in practice. An essential presupposition of retribution is that the offender is fully responsible for his crime. Yet, until recently, the United States executed juveniles and the mentally retarded. The standard for retardation is left to the state, so states continue to execute criminals with low IQs. The mentally ill are not exempt from the death penalty. Hugo Adam Bedau suggests that it is hypocritical for retributive defenders of the death penalty not to object with equal force to the execution of these less responsible individuals. The retributivist response to this claim is that even the mildly retarded and the mentally ill understand the difference between right and wrong, and so long as they do understand this difference it is just to hold them accountable for their actions.

Professor Bedau’s argument raises an important point about retribution—it has both a positive and negative component. Retribution requires that offenders be given as much punishment as they deserve but no more than they deserve. Neither abolitionists nor defenders of the death penalty consider the full implications of retributivism. Abolitionists often point to the execution of innocents as a retributive injustice and cause for abolishing the death penalty without

recognizing the retributive injustice of not giving defendants deserving the death penalty their due. Defenders of the death penalty sometimes make the same mistake and demand the death penalty on retributive grounds without acknowledging the problem of innocence and the difficulty in determining with certainty which defendants deserve to die and which do not, each a serious concern from the standpoint of retribution.

Another commonly raised objection is that although retribution offers a compelling definition of who should be punished—the guilty and no one else—it is less effective at defining the quantity of punishment that the offender should receive. Why does the murderer deserve death and not life imprisonment without the possibility of parole? It is not useful to rely on *lex talionis*—“an eye for an eye”—to resolve this question. We do not generally think that rapists deserve to be raped or that car thieves deserve to be punished by having their cars taken from them. For some crimes, it is not even possible to imagine the equivalent punishment. How would one defraud a fraud? What would be the appropriate punishment for a traitor or a kidnapper or an embezzler? Albert Camus argued that the death penalty is too much punishment for murderers:

*An execution is not simply death. It is just as different from the privation of life as a concentration camp is from prison. It adds to death a rule, a public premeditation known to the future victim, an organization which is itself a source of moral sufferings more terrible than death. Capital punishment is the most premeditated of murders, to which no criminal's deed, however calculated can be compared. For there to be an equivalency, the death penalty would have to punish a criminal who had warned his victim of the date at which he would inflict a horrible death on him and who, from that moment onward, had confined him at his mercy for months. Such a monster is not encountered in private life.*

One response to the uncertainty of the appropriate quantity of punishment might be to say that criminals deserve the punishment the law provides, a legalistic notion of retributivism. If the law says that car thieves are to be punished by 20 years in prison, and the law is known within the community, and the thief chooses to steal anyway, then this is the punishment the offender deserves. This also means that murderers deserve death in Texas, but not in Massachusetts, and that what they deserve may change from year to year as new legislators enter office. This is an unsatisfying and arbitrary notion of retributivism. Even the most devoted retributivists generally concede that the calculation of punishment requires a more nuanced formula than either legalistic retributivism or *lex talionis* provides.

The leading modern solution to this problem has been a notion of proportionate retributivism. Penalties are ranked in order of harshness; crimes are ranked in order of severity. The most severe crime is paired with the harshest punishment, the second most severe crime is paired with the second harshest punishment, and so on until all crimes have been paired with proportionate punishments. This method dictates relative levels of desert—a murderer deserves a harsher punishment than a thief—but this method does not dictate any absolute levels of punishment. If murder is the most severe crime, the murderer may justly receive a 10-year prison term if that is the harshest punishment the society deems just, and every other criminal, thieves included, should receive less severe penalties.

In the debate that follows, Claire Finkelstein argues that even proportionate retributivism does not dictate support for the death penalty. Finkelstein notes that proportionate retributivists do not believe that all forms of punishment are acceptable—virtually everyone agrees that torture is unacceptable. But there is no objective theory by which to determine which acts are morally acceptable and which are not. If torture is off limits, Finkelstein argues, then surely the death penalty is too. Finkelstein posits that the death penalty is fundamentally at odds with the retributivist's recognition of the physical integrity of human beings. Michael Davis responds by arguing that it may be possible to construct a common standard of humanness: that we can determine

with certainty which punishments are morally tolerable within a particular society. Davis argues that in the contemporary United States, early death—which is what capital punishment inflicts—is not so rare as to be inhumane.

---

## ■ Critical Documents

---

### **Immanuel Kant, “The Right of Punishing” from *The Metaphysics of Morals***

---

Judicial or juridical punishment is to be distinguished from natural punishment, in which crime as vice punishes itself and does not as such come within the cognizance of the legislator. Juridical punishment can never be administered merely as a means for promoting another good either with regard to the criminal himself or to civil society, but must in all cases be imposed only because the individual on whom it is inflicted has committed a crime. For one man ought never to be dealt with merely as a means subservient to the purpose of another, nor be mixed up with the subjects of real right. The penal law is a categorical imperative; and woe to him who creeps through the serpent-windings of utilitarianism to discover some advantage that may discharge him from the justice of punishment.

But what is the mode and measure of punishment which public justice takes as its principle and standard? It is just the principle of equality, by which the pointer of the scale of justice is made to incline no more to the one side than the other. It may be rendered by saying that the undeserved evil which any one commits on another is to be regarded as perpetrated on himself. Hence it may be said: “If you slander another, you slander yourself; if you steal from another, you steal from yourself; if you strike another, you strike yourself; if you kill another, you kill yourself.” This is the right of retaliation (*jus talionis*); and, properly understood, it is the only principle which in regulating a public court, as distinguished from mere private judgment, can definitely assign both the quality and the quantity of a just penalty. All other standards are wavering and uncertain; and on account of other considerations involved in them, they contain no principle conformable to the sentence of pure and strict justice.

It may appear, however, that difference of social status would not admit the application of the principle of retaliation, which is that of “like with like.” But although the application may not in all cases be possible according to the letter, yet as regards the effect it may always be attained in practice, by due regard being given to the disposition and sentiment of the parties in the higher social sphere. Thus a pecuniary penalty on account of a verbal injury may have no direct proportion to the injustice of slander; for one who is wealthy may be able to indulge himself in this offence for his own gratification. Yet the attack committed on the honour of the party aggrieved may have its equivalent in the pain inflicted upon the pride of the aggressor, especially if he is condemned by the judgment of the court, not only to retract and apologize, but to submit to some meaner ordeal, as kissing the hand of the injured person.

But whoever has committed murder, must die. There is, in this case, no juridical substitute or surrogate, that can be given or taken for the satisfaction of justice. There is no likeness or proportion between life, however painful, and death; and therefore there is no equality between the crime of murder and the retaliation of it but what is judicially accomplished by the execution of the criminal. His death, however, must be kept free from all maltreatment that would make the humanity suffering in his person loathsome or abominable. Even if a civil society resolved to dissolve itself with the consent of all its members—as might be supposed in the case of a people inhabiting an island resolving to separate and scatter themselves throughout the whole world—the last murderer lying in the prison ought to be executed before the resolution was carried out. This ought to be done in order that every one may realize the desert of his deeds,

and that blood-guiltiness may not remain upon the people; for otherwise they might all be regarded as participators in the murder as a public violation of justice.

---

## ■ Perspectives

---

### Issue—Does Retribution Demand the Death Penalty?

---

#### Claire Finkelstein, “Death and Retribution”

---

21 *Criminal Justice Ethics* 12 (2002)

#### I. Introduction

It is often supposed that a theory of punishment predicated on desert lends support to the death penalty. What leads to this assumption is a prior thought about the appropriate punishment for murder: If we are to punish murderers as they deserve, we will inflict on them what they inflicted on their victims, namely death. This association between a desert-based theory of punishment, known as retributivism, and the death penalty appears not only in academic writings on the subject, but in popular views of punishment as well. Public rhetoric in support of the death penalty, for example, is nearly always retributivist. Politicians urging its use in a particular case will more readily speak of justice and desert than of future dangerousness or setting an example for others. They evidently think the retributivist argument for death more appealing than the utilitarian arguments that might be made in its favor.

In my view, however, the faith that death penalty proponents place in the retributivist theory of punishment is misplaced. In this essay I argue that retributivism fails to justify the use of death as punishment, and, moreover, that a desert-based theory of punishment is particularly ill-suited to such a task. I shall not argue against retributivism as a theory of punishment per se. My more limited suggestion is that even if retributivism succeeds in justifying the practice of punishment overall, it cannot provide a compelling reason for including the penalty of death in that practice.

#### II. The Basic Retributivist Argument for the Death Penalty

Retributivism is the theory of punishment that asserts that punishment is justified because, and only to the extent that, the criminal deserves to be punished in virtue of the wrongfulness of his act. Traditionally, the core of the retributivist’s argument for any specific penalty is the doctrine of *lex talionis*, the idea that a person deserves to experience the suffering or moral evil he has inflicted on his victim. Taken literally, *lex talionis* is an absurd doctrine—no one thinks we should rape rapists, assault assailants, or burgle the homes of burglars. This difficulty making sense of *lex talionis* has accordingly led some retributivists to suggest that retributivism is most compelling as a general justification for the institution as a whole, without thinking of it as containing a further theory of the measure of punishment. But in the absence of its accompanying doctrine of *lex talionis*, or any other way of giving content to the notion of desert, the retributivist will remain unable to justify any specific penalty, including the death penalty. Given that retributivism is absurd if accompanied by a literal interpretation of *lex talionis*, and vacuous (for our purposes) if articulated without *lex talionis*, the retributivist must attempt to cast his defense of the death penalty in terms of a more approximate system for matching crimes with punishments, one that does not insist that the punishment exactly fit the crime.

Most retributivists who argue in favor of the death penalty attempt to do just this. That is, they abandon *lex talionis* in favor of a similar idea, namely that a criminal deserves to suffer some approximate match for what he inflicted on his victim. But this approach turns out to be

more problematic than one might have thought. Begin by considering just how approximate the doctrine must be to work. It is not only that presently we are unwilling to inflict some of the more extreme harms, like rape and torture, that criminals sometimes inflict on their victims. The prohibited list also includes more modest harms, such as forcing a member of a fraternity to imbibe too much alcohol or requiring a rogue cop to remove his clothes and walk half a mile in winter along a public road, both harms that such perpetrators have inflicted on their victims. Indeed, once one begins to consider all the deviant forms of behavior our criminal codes outlaw, it is clear that the vast majority of criminal acts are not ones we feel entitled to impose by way of punishment. There are really only several criminal acts that we regard as yielding acceptable forms of punishment: false imprisonment, theft, and in some states murder. The retributivist who wishes to match crimes with punishments must come up with a theory that would limit the deserved penalty to the three forms of criminal conduct listed above.

There are two possible strategies available to the retributivist in order to accomplish this end. The first distributes punishments proportionately, so that the worst crimes are matched with the worst penalties, and so on down the line. This method dictates only relative levels of desert, rather than requiring any particular objective measure of what criminal acts deserve what treatment. We might call this version of retributivism the “proportionate penalty” theory. One can think of this strategy as an alternative to *lex talionis* or as an interpretation of it; the label is unimportant. For the sake of clarity, let us treat *lex talionis* as the theory that calls for a strict equivalence between crime and punishment. Proportionate penalty will thus be a modification of the basic *lex talionis* doctrine. Whatever its other merits, this latter approach will not help the retributivist to argue for the death penalty—the method does not provide an argument that we ought to include a given penalty on the list of acceptable penalties. It merely insists on taking available punishments—that is, punishments we are already willing to inflict—and imposing them on perpetrators in order of severity according to the seriousness of the criminal acts performed.

The second, and more promising strategy is to attempt to establish a moral equivalence between crimes and permissible punishments. This strategy asserts that the perpetrator should suffer an amount equivalent to the harm or moral evil inflicted on the victim, but the kind of harm or moral evil involved need not match. That is, instead of either assigning the same harm or evil as punishment that the offender inflicted on his victim, or fixing penalties proportionately by making sure that the right intervals obtain between levels of punishments, we can match crimes with punishments on an absolute scale, but establish only a rough moral equivalence between the two. We would seek to inflict on the perpetrator by way of punishment the nearest morally permissible form of punishment to the act the perpetrator committed. Let us call this version of retributivism the “moral equivalence” theory of justified punishment.

Arguably, Kant suggested a moral equivalence approach to punishment on at least one occasion:

*A monetary fine on account of a verbal injury, for example, bears no relation to the actual offence, for anyone who has plenty of money could allow himself such an offence whenever he pleased. But the injured honour of one individual might well be closely matched by the wounded pride of the other, as would happen if the latter were compelled by judgment and right not only to apologize publicly, but also, let us say, to kiss the hand of the former, even though he were of lower station.*

The moral equivalence theory has no way of identifying which penalties are morally permissible and which are not. That is, whatever other lessons we should draw from Kant’s view on punishment, we should understand him as believing that a perpetrator can be treated in a way that is morally commensurate with the harm and suffering he inflicted on the victim, without having to inflict that very same punishment on him. And while Kant does not articulate the

theory in this way, the basic strategy of such views is to try to distinguish what a person deserves, in some absolute sense, from what it is permissible for society to inflict on him by way of punishment. The moral equivalence theory thus maintains that while a criminal who locks his victim in the trunk of a car before killing her may “deserve” to be locked in a trunk himself before being executed, it is not permissible for us to inflict such a punishment on him. The moral equivalence theory suggests that we eliminate forms of treatment that are impermissible from our roster of available punishments, and within that constraint, attempt to match the offender’s criminal act as closely as possible with the punishment we inflict on him.

But the moral equivalence theory is woefully incomplete. By itself it has no way of identifying which penalties are morally permissible and which are not. How do we know, for example, that locking a perpetrator in the trunk of a car and then killing him is impermissible, but that simply executing him is not? The moral equivalence theory needs to be supplemented by another moral theory, one that tells us which penalties are morally permissible and which are not.

Let us suppose the moral equivalence theorist manages to supply such an account and that we accept the theory in that form. Still, it is not clear that the moral equivalence theory can be used to defend the death penalty. There are at least two problems with that view. First, even in this modified form, there clearly are some penalties we think of as morally unacceptable but which are less severe than death. And if we wish to rule out those penalties, we will be compelled to rule out death as well. Consider torture. It is difficult to see torture as off-limits on the grounds that it is unacceptably severe, because it is actually most plausible to think of torture as less severe than death. The moral equivalence theorist’s own method makes it clear why this is so: If penalties are to be the equivalent of crimes, then we should rank penalties the way we rank crimes. But we think of murder as a more heinous crime than any non-lethal assault. So torture should be a less severe penalty than death. But if torture is an unacceptable penalty, it should follow that death is unacceptable as well, given that death is a more severe penalty than torture. Let us call this argument on behalf of the abolitionist the “severity response.”

Admittedly there are some problems with the severity response. First, the death penalty proponent might dispute the claim that death is more severe than torture. Torture is more severe, he might argue, because it is more uncivilized and more brutal than death. That torture is widely regarded by everyone, on all sides of the argument, as unacceptable but death is not seems to bear out the death penalty proponent’s intuition on this point. But what about the fact that most perpetrators themselves would choose torture instead of death? The death penalty proponent might say that the criminal’s own preferences cannot be the measure of the severity of a punishment. It is perfectly possible, for example, that a given criminal would prefer to spend a night in jail than to pay a small fine, because he is very attached to money and does not particularly mind confinement. But this would not show that the night in jail is a less severe penalty than the fine. There are also those who would prefer death to life imprisonment without parole. Yet surely these preferences do not imply that life imprisonment is a more severe penalty, even for those defendants.

Against this (and on behalf of the abolitionist), it might be noted that we can speak of the pain of having a tooth pulled and compare it favorably to the pain of having a severe burn, without its being infallibly true that everyone would choose the former over the latter. We are entitled to think of incarceration as “painful” because we assess that treatment in general terms. It need not be the case that everyone who has ever been “punished” by way of incarceration found that treatment painful. There are many penalties we would readily classify as less severe than torture and death that we do not hesitate to say are impermissible to impose.

Finally, the death penalty proponent might object to the severity argument by pointing out that severity does not correlate terribly well with impermissibility. There are many penalties we would readily classify as less severe than torture and death that we do not hesitate to say are impermissible to impose. There is, for example, great resistance to the recent turn toward shame

sanctions, such as forcing a convicted sex offender to display a sign outside of his dwelling revealing his status or forcing him to affix sex offender license plates to his car. It is also questionable whether it is permissible to sterilize repeat sex offenders, or to force female adolescent offenders to subject themselves to mandatory birth control measures like Depo-Provera. If the abolitionist's severity argument were correct, however, we would have to take the lowest unacceptable penalty on the list of penalties and say that any penalty more severe than it would be morally unacceptable. This strategy would quickly rule out most sentences currently inflicted for felonies, since many objectionable shame sanctions are less severe than most terms of imprisonment.

On behalf of the abolitionist once again, and against this counter argument, we might say that perhaps we are wrong to reject shame sanctions and other minor interferences with liberty. Indeed, the above argument suggests a compelling reason to allow them, namely, that imposing them may enable us to avoid inflicting more severe penalties that involve significantly greater loss of liberty. If, for example, we have the choice between a shame sanction—like a sign or a license plate—and a period of incarceration, and if both penalties are equally effective from the standpoint of deterrence, we arguably have an obligation to inflict the shame sanction, since it is the less invasive penalty. On balance, then, the argument we offered above on the abolitionist's behalf seems to me a good one—that if we reject torture because it is too severe, we should reject death as a penalty because it is more severe. This argument, however, seems to require that we revise our intuitions about a number of lesser penalties.

Because neither proponents nor abolitionists offers an account of severity or even particularly attempts to explain its stance on death, the debate between proponent and opponent seems arbitrary. It quickly reduces to the question of whether we think death an excessively harsh penalty—and that is not a terribly nuanced ground on which to settle the matter.

But we should not be too hasty to declare the argument a draw. The result is that the death penalty proponent bears the burden of proof. The retributivist is particularly affected by this burden of proof claim, for by his lights killing a person is such an evil act that the killer incurs a tremendous moral debt, repayable only with the murderer's own life. It would seem to follow that an executioner, or society more generally, who takes a person's life must incur this same moral debt, unless his act is morally justified. In the absence of such a justification, the executioner, and society as his accomplice, is no better than a murderer.

### III. Conventional Retributivism

There is another version of the retributivist's argument we must at least briefly consider, according to which principles of desert should be understood as grounded in contractarian agreement. This view enjoys a substantial advantage over ordinary retributivism. The version of retributivism that is grounded in consent stands in an entirely different position relative to that presumption against punishment with which we began. Here the presumption, if anything, works the other way around: We seem entitled to assume, at least as an initial matter, that a treatment to which a person has consented is one it is morally acceptable to inflict on him. If it can indeed be shown that there is a consensual basis for treating the worst criminals as deserving of death, it will be much less difficult to justify inflicting it on them.

According to the view we are now considering, the reason the person who commits a particularly vicious murder deserves to die is that he has agreed to the norms that dictate this treatment. I suggest we call this "conventional retributivism," since this account makes the notion of desert, which lies at the heart of the retributivist approach, dependent on the prior agreement of those who will be affected by it. There are two versions of this kind of approach. According to the first version, we can actually think of the criminal as agreeing to his own punishment, since the criminal act itself constitutes the criminal's acceptance of the punishment that is later imposed on him. This is what I shall call the "voluntarist" theory of punishment. The central

thought of the voluntarist account is that the criminal gives his tacit consent to be punished by performing the criminal act in the first place.

There are some problems with this account. First, it is not at all clear that the consent that attaches to the voluntarily performed act should be thought to transfer to a consequence of that act the agent foresees but does not intend. Second, the voluntarist argument justifies far more than we are comfortable allowing by way of legitimate punishment. This is because justifications for punishment based on consent lack a principle of proportionality that would limit the level of punishment that could be imposed for a particular crime. Thus it would be perfectly acceptable to assign the death penalty for a minor traffic offense, as long as the offender was aware of the risk of receiving that penalty when he voluntarily broke the law. Although someone might be prepared to embrace this consequence of a consensual account, it seems a deeply objectionable feature of this theory, because it would put the consensual account out of sync with our prevailing practices of punishment. This is a problem that should worry the retributivist, since retributivism trades particularly heavily on intuitions about which penalties “fit” with which acts, it does not appear to be open to the retributivist to dispense with proportionality concerns.

There is, however, a second version of conventional retributivism. In this version, instead of thinking of consent as operating act-by-act, and therefore looking for a way of establishing the criminal’s consent to the actual punishment he suffers, we should understand the entity to which the criminal consents as a general institution, or set of principles, which in turn provide a justification for a particular treatment of a given offender. The consent, that is, operates at the level of what Rawls calls the “basic structure,” instead of at the level of the institution’s particular response to the performance of a given prohibited act.

The question we must now ask is whether this “consensual” account of punishment would favor the death penalty. The question is whether individuals in an antecedent position of choice would have chosen the death penalty as part of the most rational means of dealing with those who commit the worst violations of the beneficial social rules that have been adopted. The argument that they would goes as follows. Each person enters into society because he fears for his bodily security and longevity. Putative members of society would choose to include the death penalty in the schedule of available penalties for the worst crimes since, assuming the death penalty deters, they would thereby increase the expected benefits of a system of punishment. The argument for the death penalty, then, is that it increases each person’s expected security.

Against the above argument, however, one should note that every member of society must also take into account the possibility that he will be subject to the death penalty, either because he chooses to commit a crime for which the death penalty is authorized, or because he is wrongfully thought to have committed such a crime. The death penalty thus also decreases each person’s expected security. Thus the increased security provided by the deterrent benefits of having the death penalty must be balanced for each rational individual against the decreased personal security the death penalty also involves. Rational agents selecting rules for the governance of society would choose to have a death penalty if it turned out that the added deterrent benefits of having that penalty, over and above the deterrent benefits of life sentences, were greater than the decreased security to each person from the possibility of being subject to the death penalty themselves. Arguably it would be rational for contracting agents to include the death penalty in the rules governing punishment.

But what about the person who is erroneously put to death? For him, there is no net increase in expected security. As it turns out, his bodily security will have been destroyed by the existence of the death penalty. He would probably have done considerably better if he had lived with the increased threat of being the victim of a violent crime, and avoided the death penalty (even if he had to spend the rest of his life in prison anyway). And arguably since each person knows *ex ante* that he may end up in this position *ex post*, he cannot regard the death penalty as rationally motivated.



So whether rational agents in a contractarian scheme would agree to the death penalty under the circumstances is a complicated matter. If the *ex ante* expected security of the death penalty is positive, on one view this would be sufficient to make it rational for putative members of society to agree to it, even if the death penalty could end up leaving any given individual worse off than if he had not agreed to it. On another view, however, for putative members of society to agree to the death penalty, each person must believe that no matter how things turn out, he will be better off under the death penalty regime than he would be in its absence. Let us make a very implausible assumption, namely, that the deterrent benefits from the death penalty are so great as to leave even the person put to death under its rule better off than he would be in a society that did not include the death penalty in its available punishments. On this hypothesis, the difference between the two forms of the consensual account would disappear, and we have reason to think that the consensual approach to punishment would endorse the death penalty.

But even if rational agents can regard their expected security as improved by the existence of the death penalty (and indeed even if the death penalty meets the stronger condition of leaving all members of society better off than they would be in its absence), it need not be the case that rational individuals would institute it. For there are deontological norms to which the retributivist is committed, and these arguably counsel against legitimizing an agreement with a person to put him to death. To see this, consider a different, but instructive example, the case of the “Kidney Society.”

Suppose everyone were at some small risk of finding himself with two failing kidneys. In order to protect against the risk of needing a kidney and having access to none, a group of people enter into an agreement to supply one another with a kidney by lottery, should one of the members of the group end up in this situation. The terms of the agreement specify that if any member of the group finds himself needing a kidney to survive, a lottery would be held to determine who would supply that individual with the needed kidney. Once a person is chosen by lot to supply the kidney, he would have no choice but to yield, and a kidney could be removed, by force if need be, in order to supply the needed kidney to the person with dual kidney failure. In this case, we could say that there is no benefit to the person who must have his kidney removed as the upshot of the lottery. He is clearly better off with two kidneys than with one. From his perspective, the gamble has not turned out to be worth it. But there was a benefit to him in entering into the agreement with others in the first place. We can therefore hypothesize that if he thinks the danger of dual kidney failure sufficiently great, and the loss to himself of being the one to be chosen by lot to donate a kidney either sufficiently remote or sufficiently bearable, he may regard the terms of the Kidney Society’s agreement beneficial on the whole, making it fair to enforce its terms when he resists its application to himself. Like the increased net security we hypothesized the death penalty might provide, the members of the Kidney Society enjoy an increase in the net expected health benefits they experience, even counting the “costs” of having to provide a kidney.

Yet there are reasons to suppose that forcibly removing an objector’s kidney would not be justifiable on consensual grounds, for it is not at all clear that rational agents concerned to maximize their long-term well-being would embrace the kidney scheme, or that ultimately it would be rational for individuals to ensure against kidney loss in this way. Despite the fact that each person can regard himself as better off living in a society in which he has insurance against dual kidney failure, at the risk of having to provide one of two extraneous kidneys himself, rational agents might not prefer a social scheme that provides them with such benefits. Rational agents might not want to select the “meta-regime” in which it was permissible to attack a non-consenting donor in the way necessary to enforce the Kidney Society lottery, even if the donation were one that the donor had putatively agreed to *ex ante* by signing up for the Kidney Society. One has only to reflect on the extreme discomfort we would have enforcing the kidney donation agreement if the person who had drawn the short straw in the lottery objected. Would we be willing to enforce such an agreement, solely on the grounds that the unlucky lottery participant had benefited

from the insurance such a scheme gave him up until now? It goes without saying that no court in this country would order specific performance for such an agreement. Even if organ donation agreements were not void as contrary to public policy, the most one could expect to win against a recalcitrant donor would be monetary damages for the failure to turn over his kidney.

Presumably what bothers us about enforcing a kidney lottery is that we think the person who draws the short straw in the lottery has rights to bodily integrity, rights that he did not, and could not, contract away in the initial agreement to enter the lottery. While we might imagine contracting away some substitute for one's kidney—monetary compensation, for example—we cannot quite understand one's own bodily organs as themselves subject to voluntary renunciation. Indeed, a retributivist ought to have particularly strong objections to seeing rights to bodily integrity as the subject of a contract between rational agents, for it is usually because of the retributivist's view of the robustness of such rights on the victim's part that he is so confident of his judgment that the murderer deserves to die for violating them.

#### **IV. Conclusion**

If the retributivist seeks to defend the death penalty by claiming that an offender deserves to suffer precisely the treatment he has inflicted on his victim, the theory will prove absurd and morally repugnant: We simply are not prepared to rape the rapist and assault the assailant. So he must try to find the "moral equivalent" of the offender's criminal act instead, ruling out immoral acts and restricting available penalties to morally acceptable forms of punishment. But which acts are morally acceptable? Many penalties we currently consider morally impermissible, such as torture, are actually less severe than death. And that suggests that whatever method the moral equivalence theorist uses to eliminate torture could also be used to eliminate death. So unless the retributivist is prepared to allow that criminals can be tortured, it will be hard for him to employ a moral equivalence strategy to argue for the death penalty. And in the absence of a clear justification for using death as a punishment, the death penalty is morally impermissible, given that there is a presumption against the infliction of any painful, involuntary treatment.

There is, however, another possibility for justifying the death penalty along retributivist lines, and this is to combine the basic retributivist approach with a consensual theory of punishment: Criminals would be subject to an agreed upon roster of deserved penalties, allocated according to general moral principles they endorse. On this strategy, it would not be necessary to justify the death penalty *per se*, since a treatment to which an offender has himself consented is at least presumptively morally acceptable. The most compelling version of what I have been calling conventional retributivism sees the offender as having agreed to norms that govern the institution under which he is punished.

But it is very implausible that rational agents primarily concerned with protecting and prolonging their lives would consent to an institution of punishment that contained such a terrible penalty, even if it provided them with positive net expected security. Rational agents would be unlikely to agree to equip one another with the power to put members of their own society to death, despite the benefits in expected security we are now assuming they would garner. They would thereby alienate the strong rights to bodily integrity whose protection was their primary motivation for entering into civil society, the same set of rights that arguably lies behind the retributivist's conception of desert for punishment in the first place.

#### **Michael Davis, "A Sound Retributive Argument for the Death Penalty"**

21 *Criminal Justice Ethics* 22 (2002)

Claire Finkelstein's primary conclusion is that "retributivism fails to justify the use of death as punishment." By "justify," Finkelstein seems to mean no more than "show to be morally permitted." I shall justify the death penalty in a somewhat stronger sense. I shall show statutory provision

of the death penalty to be not only sometimes morally permitted but, all else equal, something that reason sometimes recommends. I shall not, however, argue that morality (or reason) requires it. I do not believe that countries such as Canada and Mexico, or states such as Wisconsin and Michigan, are engaged in immoral conduct because they do not punish murder with death.

## I. Two Kinds of Retribution

Finkelstein begins her discussion of retributivism with an oversight that may explain how she missed the argument I am about to make. While recognizing that the literal form of *lex talionis*—the criminal deserves in punishment the same harm that the crime caused—is not a popular position even among staunch retributivists, she still seems to suppose that all forms of retributivism must proportion punishment to the harm done (at least for the major intentional crimes). The retributivist “would seek to inflict on the perpetrator by way of punishment the nearest morally permissible form of punishment to the act the perpetrator committed.” This is the “moral equivalence” (“nearest”) form of *lex talionis*.

All forms of the *lex talionis* use the harm the crime does as a major component when determining equivalence. All forms of the *lex talionis*, therefore, have trouble with the enormous diversity of crimes characteristic of any sophisticated legal system. Too many crimes do no harm—in any useful sense of “harm.” Any retributivism must be able to assign penalties to many “harmless” crimes, everything from attempted murder to reckless driving, from failure to place a tax stamp on a liquor bottle to conspiracy to commit embezzlement. For that reason, a number of retributivists have tried to understand punishment as taking back (or canceling) the unfair advantage the criminal gets from the crime as such (or, at least, the value of that unfair advantage) rather than as returning something like the harm he did. Here is an example of the method by which that sort of retributivism would assign penalties to crimes:

1. Prepare a list of penalties consisting of those evils (a) which no rational person would risk except for some substantial benefit and (b) which may be inflicted through the (relatively just) procedures of the criminal law.
2. Strike from the list all inhumane penalties.
3. Type the remaining penalties (by evil imposed), rank them within each type (by amount of that evil), and then combine rankings into a scale.
4. List all crimes.
5. Type the crimes (by interest protected), rank them within each type (by degree of protection), and then combine rankings into a scale.
6. Connect the greatest penalty with the greatest crime, the least penalty with the least crime, and the rest accordingly.
7. Thereafter, type and grade new (humane) penalties as in step 3 and new crimes as in step 5, and then proceed as above.

The harm a crime actually does plays no part in this assignment of (maximum) penalties. The “harm” in prospect, if the evil in question is “harm” in any interesting sense, does have a part, but one constrained by considerations of type and grade, that is, by relationships internal to a system of criminal justice.

In a society much like ours, the initial list of penalties might include: death, loss of liberty, pain, loss of property, loss of opportunities (for example, revocation of driver’s license or franchise), mutilation, and shame. These are the most persuasive penalties we can impose.

## II. Torture, Death, and Severity

Finkelstein need not object to this seven-step method. She can restate one of her arguments in its terms: Death, she might say, cannot pass the test set by step 2—or, at least, cannot if torture fails (as almost everyone now agrees it does). Death must fail if torture does because torture is not as

severe a penalty as death. Since step 2 excludes death if it excludes torture, and it does exclude torture, the death penalty cannot be justified (or, at least, is no more justifiable than torture is).

If that is the argument she would make for striking death from the list, all I need to do is provide one plausible example of a jurisdiction in which death can be justified as a penalty even though torture cannot be.

Consider, then, a jurisdiction, such as Illinois, in which the statutory penalty for armed robbery, aggravated arson, and similar serious non-lethal crimes is 6–30 years imprisonment, the statutory penalty for simple first-degree murder is 20–60 years imprisonment, and the statutory penalty for multiple murders is life imprisonment without parole. We may, I think, agree that someone who commits several first-degree murders on one occasion in an exceptionally brutal way deserves a penalty significantly more severe than the penalty for simple multiple murder (assuming the more severe penalty is both possible and otherwise permissible). That is, there is enough difference between these two categories of crime that we should (if possible and permissible) assign the more serious a substantially higher penalty. We want to give the criminal a reason not to be brutal even if he is going to murder several people, a reason that—as a rational person—he should find significant. Death is such a penalty.

But (Finkelstein might respond) you have not yet explained why we should not choose life-imprisonment-with-torture instead of death. Any rational person should consider torture (added to life imprisonment) a significant reason not to do what he might do if the penalty were only life imprisonment. If your theory justifies the use of torture, it is (presumptively) refuted, since everyone agrees torture is not justified. If, on the other hand, you cannot justify torture, how are you to justify the more severe penalty of death without begging the question of death's moral permissibility?

The reason why we cannot justify the use of torture instead of death is, of course, that torture has been struck from the list of available penalties; it is inhumane. But, in order to explain why death is not also off the list, the defender of the death penalty needs a theory of inhumane penalties (one that does not beg questions about the death penalty's moral permissibility). There is such a theory, one that explains why penalties can be morally permissible at one time and not at another, why death can be humane even if torture clearly is not, one general enough not to beg our question.

### III. Humaneness

We do not object to a penalty (such as torture) as inhumane simply because of its severity. The penalty, considered as a physical act, a certain quantity of pain or the like, never changes—though objections do. Nor do we object to a penalty as inhumane because of what we are willing to suffer. We are sometimes willing to suffer inhumane penalties ourselves (for example, a few hours of torture to avoid long imprisonment). We seem to object to a penalty as inhumane only when use of that penalty on anyone, especially someone else (against his will), shocks us; when, that is, we cannot comfortably bear its general use. Shock is neither rational nor irrational. The person who does not find shocking what everyone else does is eccentric, insensitive, or callous. She is not necessarily irrational. We think she needs to let go more, to feel more, or to live better, not to be cured or caged. Shock at this or that is not a basic evaluation all rational persons must share; nor is it the inevitable consequence of what all rational persons do share. For example, much that might turn most of us pale does not bother a surgeon. What shocks us seems to be a consequence of how we live, of what we do and do not experience.

Though I may say a penalty is inhumane because it shocks me, a penalty is not inhumane just because it shocks me. A penalty is inhumane (in a particular society) if its use shocks all or almost all; humane, if its use shocks at most a few; and neither clearly humane nor clearly inhumane, if its use shocks many but far from all. We suppress inhumane penalties (in part at least) because we do not want to be shocked by their use.

But shock as such is morally indifferent whereas humaneness seems to be morally important. What is the connection between shock and morality? The connection seems to be this: to treat a person in a way that we generally find shocking is to treat her as we ourselves do not want to be treated, as we would not treat most other persons, as we might not even be willing to treat animals. It is, in short, to degrade her, to treat her as less than a person. If morality requires us to treat each person as a person (and that, it seems to me, is relatively uncontroversial), then we do something (at least *prima facie*) morally wrong if we inflict on a person (against her will) a penalty that we find shocking.

#### **IV. How Is a Common Standard of Humaneness Possible?**

Making humaneness a function of shock in this way may seem puzzling. Shock, I said, is neither rational nor irrational. If what is inhumane depends upon most of us agreeing about what shocks us, how does it happen that there is so much agreement about what is inhumane?

The puzzle is not hard to solve. Ways of life shape our sensibilities in certain ways. A shared way of life, because it shapes a common sensibility, also forges a standard of humaneness. We are made to agree. Imprisoning someone for months or years would shock most people in a society of nomads. In such a society, the gentlest detention would be inhumane. Imprisoning does not shock us only because we are every day penned in houses, workshops, and offices and often spend years in the same city or town. Although we admit imprisonment to be a great evil, we do not think of it as an inhumane punishment.

There is then some relationship between “progress” and what is or is not inhumane. Insofar as technological progress has meant the disappearance of certain evils from daily life, it has meant as well a change of sensibility and so a change in what society should lay aside as inhumane (“cruel and unusual,” as the courts say). Public executions came to shock our humanity about the time it became rare for ordinary people to die in the street. During the last few decades, most death-penalty states switched from execution by (private) hanging, gassing, or electrocution to execution by (private) lethal injection. Why? Perhaps because death by injection is more like the hospital death we are now accustomed to (hanging, gassing, and electrocution resembling more the industrial accidents that, happily, have become relatively rare).

What then of the death penalty itself? If what I have said so far is right, the death penalty will shock enough of us only when it has become rare enough for people to die unwillingly before old age (such an early death being the evil that the death penalty imposes). How rare is “rare enough”? That is a question for social science. What is clear, I think, is that early death is not rare enough in the United States today to make the death penalty as shocking as it needs to be to be inhumane here. Certainly, death by lethal injection does not shock most of us in the way even fifty lashes with a whip does.

#### **V. The Kidney Society**

I have admitted that criminal desert is not the only constraint on how we can justifiably punish. Humaneness is another. I may seem thereby to have opened myself to Finkelstein’s final argument, the one resting on “deontological norms [apart from desert] to which the retributivist is committed.” She makes her argument using a single example, the Kidney Society. I agree with her intuition that “we” would object to the violation of bodily integrity necessary to make the Kidney Society’s version of social insurance work. We would not require specific performance even if we would enforce the contract by assessing damages for a failure to perform. What I deny is the relevance of that intuition to the death penalty. The death penalty does not involve the same violation of bodily integrity as removing a human kidney. Death by lethal injection is much closer in that respect to forced medication (something that does not shock us) than to the Kidney Society’s forced mutilation. In any case, our intuitions here rest on the same sensitivities as our intuitions of humaneness. Our intuitions of humaneness need not make sense according

to some abstract theory of what resembles what. They are what they are, and that is the end of it—until they change.

Finkelstein might, I suppose, still respond that we should reject the death penalty whenever a less severe penalty is available. Torture is a less severe penalty. Therefore, we should reject the death penalty.

To this last (desperate) argument, I would respond that torture is not available—for the reasons already given. It is inhumane (even though less severe). Is there an alternative to the death penalty, another penalty significantly more severe than life imprisonment without parole but not inhumane? Some of the other penalties on our original list, all those that significantly add to life imprisonment without parole—torture, flogging, and so on—are also inhumane (and so, are not available). The remainder of the list, though available, does not significantly add to life imprisonment without parole. Why should someone who contemplates brutally murdering a number of people care whether, in addition to life imprisonment without parole, his punishment would include loss of his driver's license or his right to vote?

If the death penalty is the only penalty Illinois has that is both available and significantly more severe than life imprisonment, Illinois is left with the death penalty as the only way to distinguish aggravated multiple murder from lesser forms of multiple murder. Since we have good reason to want to proportion punishment to criminal desert, Illinois has a retributive justification to use death as a penalty, because of the markedly greater criminal desert of aggravated multiple murder.

Of course, this argument presupposes that Illinois' statutory scheme is (relatively) just otherwise and that Illinois cannot achieve the distinction between aggravated and lesser crimes by reducing the penalty for lesser crimes. Although I believe that Illinois' statutory scheme is too severe overall, even though it is otherwise relatively just, its being so does not matter here. Even if Illinois' penalties are too severe overall, nothing in Finkelstein's argument, or in the retributive method I outlined here, rules out the possibility that a statutory scheme like Illinois' might be justified somewhere sometime. If lesser penalties will not preserve enough order, the legal system will either have to impose heavier penalties overall or, after trying to improve enforcement, subside into disorder. What scheme of penalties is sufficient to preserve order is not something to be settled by an abstract argument such as Finkelstein's. Finkelstein's argument is refuted.

### ■ Discussion Questions

1. Is the death penalty a proportionate punishment for murderers? Does the answer depend on the motivation of the murderer?
2. If we assume that death is the deserved punishment for a murder, is it deserved too for the person whom the murderer might become? In other words, does the death penalty unjustly prevent the possibility of change or growth?