

Dewey's Rejection of Retributivism and His Moral-Education Theory of Punishment

John Shook

This essay examines the two important arguments that John Dewey's social philosophy offers against the retributivist justification for punishment. The first argument is that retributivism cannot serve as an adequate explanation for the moral practice of punishment, while the second argument develops the position that a liberally democratic society would naturally prefer an alternative to the retributive model of punishment that is more consistent with democracy's concern for the social good. Although Dewey did not fully develop a systematic theory of punishment, his writings on morality, education, and democracy provide the foundations for constructing a Deweyan approach to the justification of punishment. The cornerstone is Dewey's theory of moral education for a liberal democracy. This theory points the way toward an explanation of why and how society should aim at continually developing the moral responsibility of all citizens. In brief, punishment should play an essential role in moral education by enhancing the ability of offenders to be more socially responsible in the future. Examination of Dewey's arguments against retributivism shows how he provides an alternative to both deontological and utilitarian approaches to punishment, thereby establishing a novel version of the moral-education theory of punishment.

1. Retribution and the Moral Practice of Punishment

The relationship between punishment, responsibility, and morality can be expressed in a deceptively simple manner: Punishment minimally requires the moral disapproval of a person's conduct (broadly understood), for which he or she is responsible. Dewey believed that it is essential to distinguish two divergent notions of responsibility, which in turn ground two very different approaches to morality, and thus to punishment.¹ The past-looking or "retrospective" concept of responsibility is involved when we assess the various relevant factors of the situation leading up to a person's action to determine whether that person really performed the action. The forward-looking or "prospective" concept of responsibility is involved when we attend to the wider context of the past and future behavior of the person, so that her future behavior may be adjusted by our interventions. Dewey placed great weight on this distinction; for example, in his 1932 book *Ethics* he wrote the following:

Now the commonest mistake in connection with the idea of responsibility consists in supposing that approval and reprobation have a retrospective instead of prospective bearing. The possibility of a desirable *modification* of character and the selection of the course of action which will make that possibility a reality is the central fact in responsibility. The child, for example, is at first held liable for what he has done, not because he deliberately and knowingly intended such action, but in order that *in the future* he may take into account bearings and consequences which he has failed to consider in what he *has* done. Here is where the human agent differs from a stone and inanimate thing, and indeed from animals lower in the scale.²

Dewey's claim that "responsibility" is essentially a normative term of moral judgment understandably appears false at first reading. Do we not say that apple trees are responsible for apples, and that the moon is responsible for tides? But Dewey is not arguing that all uses of the term "responsibility" are moral uses. Rather, he is arguing that when a morally normative verdict of responsibility is made on a person's behavior, this moral verdict is essentially made for the purpose of affecting future behavior. It is this future-regarding core to moral evaluation that leads Dewey toward his rejection of retributivism.

It is not absurd to declare a person responsible for an unintended action, so long as responsibility is understood in a sufficiently broad sense that does not require intent. It is this broad sense that Dewey is after to make his point regarding our treatment of children. Did the child unintentionally hurt the cat? If punishment requires responsibility, and responsibility requires intent, then we cannot punish a child for thoughtless behavior. However, as every parent well knows, a child's punishments are more often than not earned by *unintentional* conduct. Therefore, while the presence or lack of intent may have a bearing on the degree of moral disapproval, intent is not essential to the sort of responsibility connected to punishment (and the same is true for adults, a point to be pursued below). This point alone is of course insufficient to discredit retributivism; a retributivist could note that negligent behavior, an act of omission to take greater care, is just as punishable. Dewey quite agrees and wants to pursue this point further. Since punishment is deserved for failure to take greater care as well as for intentional acts, then *how* we punish should be connected to our more fundamental interest that people take greater responsibility in general for their conduct, instead of connecting merely to our interest that people be punished for intentionally responsible conduct. We punish when people are held responsible in the broad sense *and* we desire to modify their future behavior; both conditions (past *and* forward-looking) are essentially linked together, and they must be satisfied (at minimum) to justify punishment. Dewey's conclusion is that the only way to make sense of our practice of punishing for failures of responsibility in the broad sense (for both commission and omission) is to see punishment as necessarily involving future consequences.³ The practice of making morally normative judgments about behavior and applying punishment accordingly must have originated in, and continues to exist for the sake of, efforts to

modify people's habits of conduct to make those people more morally responsible in the future. Having established this conclusion to his satisfaction, Dewey goes on to argue that our efforts to enhance future responsibility can be accurately characterized as efforts to form moral character, and thus to morally educate people.⁴ For this reason, Dewey's consequentialist approach to punishment is most appropriately categorized with moral-education theories of punishment.

The merits and problems of Dewey's moral-education theory will be considered at the conclusion of this essay. First, we must examine whether Dewey's position on retrospective and prospective views of responsibility and punishment can be effective against retributivism. The first question to be asked is, Why does Dewey use an example of punishing a child to exemplify his approach? After all, we do not regard the punishment of *children* as the primary difficulty calling for philosophical justification. Furthermore, it may be the case that our legal treatments of adults and of children are so distinct as to amount to a difference of kind, and not just degree. Perhaps Dewey is correct concerning our treatment of children but fails to appreciate that a quite different kind of punishment is applied to adults.

But where lies such an essential difference? For both children and adults, moral disapproval involves our desire to try to prevent future offenses. If we disapprove of some event (such as the cat's suffering) and discover a person's responsibility for that event, we naturally transfer our disapproval to that person's conduct as well. Our disapproval is of that *kind* of conduct, and not merely a disapproval of that *particular* action. Our view that such conduct is wrong entails, among other things, our view that such conduct should not happen (and not merely that this specific instance of conduct should not have happened). To disapprove of someone's behavior today, and yet to be indifferent to that same behavior tomorrow under identical conditions, would be to take an attitude that does not fall within the range of moral evaluation. A *moral* evaluation of a person's conduct as wrong is necessarily connected with a desire to see that person's conduct modified for the better in the future. This is as true for adults as it is for children. Because punishment follows from a disapproving moral evaluation, and such evaluations necessarily involve the desire to modify future conduct, punishment must for all people have a future-looking perspective.

To summarize so far, the foundation for Dewey's theory of punishment is his position that the assessment of past responsibility is a necessary first step for attempting to modify a person's future responsibility, which explains how the retrospective and prospective conceptions of responsibility are related. The proper function of first judging someone's past responsibility is to use that information subsequently to engage in a process of altering future responsibility (hopefully) toward more consistent moral behavior. Of course, we do make judgments about the past or "causal" responsibility of nonagents such as an apple tree and the moon. Dewey holds that insofar as we treat human beings as having *only* causal responsibility, we do not treat them as responsible *agents*, and thus we do not treat them as moral agents. Therefore, his primary contention is that when we are concerned with morality, we are

concerned with developing the future responsible agency of a person, and thus moral education is the relevant practice involved. Our scrutiny of and verdicts upon a person's causal responsibility, when made in the appropriate moral setting, are made for the sake of then attempting to enhance future responsibility. The connection between responsibility and morality rests on the fact that future responsibility is successfully enhanced when a person's habits of conduct more closely conform to moral expectations. At this early stage we can see that Dewey clearly prefers something like a deterrence stance over simple retributivism, since the future conduct of an offender is for Dewey the central consideration.

What about the connection between punishment and responsibility? Dewey, like most philosophers, holds that the judgment that a person should be punished is essentially a normative moral judgment. Given his stance that normative moral judgments are made ultimately for the sake of trying to modify persons' future conduct, he therefore finds that the essential purpose of punishing an offender is to enhance the offender's future moral responsibility. This view (let us call it the "prospective" theory of punishment) is diametrically opposed to any "retrospective" theory of punishment. A retrospective theory of punishment instead declares that the essential purpose of punishment has nothing to do with enhancing or diminishing the offender's future responsibility. Instead, the retrospective approach holds that ascertaining past responsibility for some offense is sufficient to determine whether, and in what amount, an offender should be punished. Retributivism, at least in its pure form that wants no compromise with utilitarianism,⁵ is the classic example of a retrospective theory of punishment. The concept of retribution is vague, and proponents cannot agree on its essentials, but it is most often defined as involving either (a) the attempt to remove the stigma, or cancel out the sin, of the offense, or (b) the attempt to restore some sort of balance by causing the offender to suffer in proportion to the offense. In both varieties, (pure) retributivism staunchly denies that the future consequences of punishment should determine the form or severity of punishment. Because retributivism holds that the nature of the crime itself is sufficient to decide such matters, it is a paradigm type of retrospective theory of punishment.

From Dewey's perspective on the proper relationship between assessing past responsibility and enhancing future responsibility, the retrospective theory of punishment effectively removes and isolates the assessment of past responsibility from its proper moral setting. That is the primary reason for Dewey's rejection of (pure) retribution: It ignores the proper moral context that gives our practices of assessing responsibility their meaning and effectively ceases to treat offenders as moral agents.⁶ On Dewey's moral theory, to treat a person as an agent necessarily involves, in part, trying where appropriate to adjust that person's future responsibility. There is an inherent relationship between making moral judgments and acting on those judgments to attempt to enhance others' moral responsibility. Dewey explicitly connects the responsibility of any person with the responsibility of others to sustain and enhance future moral conduct. Retribution stands condemned as antithetical to the responsible moral treatment of persons.

The persons who employ praise and blame, reward and punishment, are responsible for the selection of those methods which will, with the greatest probability, modify in a desirable way the future attitude and conduct of others. There is no inherent principle of retributive justice that commands and justifies the use of reward and punishment independently of their consequences in each specific case. To appeal to such a principle when punishment breeds callousness, rebellion, ingenuity in evasion, etc., is but a method of refusing to acknowledge responsibility. Now the consequence which is most important is that which occurs in personal attitude: confirmation of a good habit, change in a bad tendency. The point at which theories about responsibility go wrong is the attempt to base it upon a state of things which precedes holding a person liable, instead of upon what ensues in consequence of it. One is held responsible in order that he may become responsible, that is, responsive to the needs and claims of others, to the obligations implicit in his position. Those who hold others accountable for their conduct are themselves accountable for doing it in such a manner that this responsiveness develops.⁷

In short, Dewey convicts the retrospective theory of punishment generally and retributivism in particular for endorsing the unintelligent and irresponsible infliction of suffering on people who are not being properly treated as moral agents. It is contradictory, in Dewey's view, to morally judge a person to have behaved irresponsibly and to be deserving of punishment, and then to carry out the punishment in a manner blind to its effects on that person's future capacity for moral conduct. Even more paradoxical, from Dewey's perspective, would be any effort to resist Dewey's conclusion by challenging the key premise exposed above: that to make a moral evaluation of disapproval is to disapprove of that *kind* of conduct, and not merely a disapproval of that *particular* action. A retributivist might prefer to hold that a moral judgment is only made upon a particular instance of conduct that entails nothing about judgments on other such conduct under relevantly similar circumstances. But that sort of particularist moral theory has never been widely accepted and has more in common with emotivist and subjectivist moral theories, which actually conflict with the typical moral orientation of retributivism.

The failure of retributivism to promote morally right conduct sufficiently cannot be overcome by emphasizing retributivism's ability to satisfy some abstract principle of justice. Dewey declared that the largest obstacle to improving our system of punishment is the lingering sanctity attached to retributivist notions. If the justification of government penal systems is simply the sacred principle of justice for its own sake, then it is difficult to repress the idea that there is really no moral justification for penal systems at all:

It is the sacrosanct character thus attributed to the State's use of force which gives pungency to the Tolstoian charge that the State is the arch-criminal, the person who has recourse to violence on the largest scale. I see no way out except to say that all depends upon the efficient adaptation of means to ends. The serious charge against the State is not that it

uses force—nothing was ever accomplished without using force—but that it does not use it wisely or effectively. Our penal measures are still largely upon the level which would convince a man by knocking him down instead of by instructing him.⁸

Dewey rightly distinguished between force and violence, and between mere force and intelligently constructive uses of force. Intelligently applied punishment, by enhancing future responsibility in a morally justifiable way, could never be properly construed as mere violence, whereas retribution is nothing but violence that lacks moral justification.⁹ Our commitment to morality must be matched by our commitment to creating and maintaining morally responsible agents.

For Dewey, punishment should be a moral practice that respects the offender as a moral agent. This stance removes from consideration two other types of consequentialist approaches, the “incapacitation” and “restitution” approaches to punishment. The incapacitation approach, in which offenders are simply prevented from offending again (using banishment, prison, etc.), is consistent with the principle that punishment should deal with future conduct. However, its application of simple brute force is inconsistent with treating offenders as moral agents. Punishment of *moral* agents would aim not just at the prevention of future wrong conduct, but rather at the encouragement of future right conduct. The restitution approach likewise involves the offender’s future conduct, by holding that punishment functions to redress the offender’s harm by forcing the offender to proportionally compensate the victim(s) of his offense. But restitution seeks to control the offender’s conduct for the short term and on a case-by-case basis; its essential aim is not to increase the offender’s capacity for responsible conduct. Dewey is seeking the consequentialist approach that best respects offenders as moral agents, and this approach clearly conflicts with retributivism.

There remain some consequentialist/deterrence approaches that are not ruled out by Dewey’s arguments. For example, a prominent consequentialist approach, rehabilitationism, can be understood as an effort to deal with an offender’s future moral responsibility. Rehabilitation’s aim of modifying a person’s character toward adequate law-abiding conduct could (depending on the methods applied) satisfy Dewey’s requirement that punishment should be a moral practice that respects the offender as a moral agent. Moral-education approaches like Dewey’s are not improperly classified under rehabilitationism’s umbrella. The conclusion of this essay further discusses the relationship between Dewey’s moral-education approach and similar theories. Dewey’s second argument against retributivism is the topic of the next section.

2. Punishment and the Social Good

Dewey’s allegiance to a prospective approach to responsibility and punishment can be further explained by setting it within the larger context of his understanding of the nature of social responsibility in a democracy.¹⁰ This understanding provides the grounds for raising Dewey’s second objection to

retributivism. Dewey is but a recent example of a philosopher committed to progressive liberal democracy who has explicitly grounded the purpose of laws, rights, and responsibilities upon the advancement of the social good. However, this emphasis on the social good should not be taken to imply that Dewey was a classical utilitarian and must not license the inference that Dewey's consequentialist approach to punishment is not really different from ordinary utilitarianism (more on this below). Unlike the typical utilitarian, who asserts that social goods are dependent on the existence of individualized private goods, Dewey doubted whether many important goods are easily individualized. With the utilitarians Dewey agreed that social norms should be judged by their ability to facilitate the production of social goods. Yet Dewey always added the proviso that many important social goods are valuable only to the extent that most or all members of society participate in, and subsequently enjoy, these goods' production. Due regard for the pervasive role of social goods in society would help to resolve the traditional utilitarian problem of gratifying one part of society by degrading another. Opponents of utilitarianism often use the concept of justice to highlight this utilitarian problem by claiming that the search for the social good must be replaced by, or at least balanced against, respect for principles of justice. But Dewey felt no need to appeal to justice conceived as independent of the social good. Justice for its own sake, he argued, has never truly been an abiding principle of democratic life. Citizens of a democracy desire justice, certainly; but they are rightfully suspicious of claims that some particular law, right, or morality transcends any and all considerations of the welfare of society. This is as true for the typical conservative protector of constitutional rights as it is for the most progressive experimenter with novel freedoms. Furthermore, citizens of a democracy do care about the impact of enforcing justice on the future behavior of the punished, and this concern is heightened when the public costs of recidivism are appraised.¹¹

From the standpoint of Dewey's philosophy of democracy, justice-centered theories attempt to ground the existence of justice in some nonsocial realm, where it is safe from social evaluation and modification. This attempt, however, only succeeds in isolating justice from human intelligence:

That something called justice is to be done even though it bring down the heavens in ruin, is the final word of an ethics which is resolutely irrelevant to the circumstances of action and the conditions of life. But the skill which invented the phrase dulls by the side of the greater cunning which disguised the fact that the phrase is an ironic summary of every moral code that leaves consequences out of account. For it is only accident if abstract ethics does not always bring wreck in its train. Yet those who recommend this adage of an unswerving morality always wear an expression of superiority though wilfully disregarding the consequences of wreaking vengeance under the plea of vindicating eternal justice.¹²

Dewey does hold that punishment must serve justice. However, for punishment to be morally responsible, the justice that punishment serves must be deeply connected to the social good. Justice-centered theories often portray

justice and its attendant values of legal duty, or rights, or righteousness as nonconventional and exempt from human deliberation. The debate between deontologists and utilitarians, from Dewey's perspective, has obscured the most significant issue: the nature and origin of social responsibility. Social responsibility is, first, the only kind of responsibility that democratic society truly cares for; and second, it is grounded in society and its structure, and not in some nonsocial realm.

Dewey steadfastly repudiated any methodology that involves looking beyond the social sphere to discern moral or legal foundations.¹³ For example, that deontological variety of moral theory that locates moral goodness exclusively within the private space of autonomous motives has forgotten that we ascertain and judge the particular motive of a person's act only by judging that act in the wider context of that person's patterns of responsible or irresponsible social conduct. All judgments of personal character are judgments of that person's social conduct, and vice versa. A "free-will" theory of moral responsibility postulates an individual's free will as the cause for responsible, or irresponsible, moral behavior. But such a theory, instead of clearing the space for a justification of punishment, actually makes punishment of the person irrational:

It is a man, a human being in the concrete, who is held responsible. If the act does not proceed from the man, from the human being in his concrete make-up of habits, desires and purposes, why should *he* be held liable and be punished? Will appears as a force outside of the individual person as he actually is, a force which is the real ultimate cause of the act. *Its* freedom to make a choice arbitrarily thus appears no ground for holding the human person as a concrete being responsible for a choice. Whatever else is to be said or left unsaid, choice must have some closer connection with the actual make-up of disposition and character than this philosophy allows.¹⁴

We do esteem the individual's consciously willful efforts to behave morally. To this extent Dewey agreed with Kant that good will is a proper object of moral approbation. However, esteem for such moral efforts of character is justifiable only because we esteem morally responsible *conduct*, and not because willful efforts to perform the good regardless of actual consequences are morally valuable in themselves. Moral consequentialists are correct to claim that we primarily encourage people to perform moral actions and not merely to have moral character. Yet moral consequentialism fails to realize adequately that we esteem *morally responsible* conduct, and not merely isolated and thoughtless actions that just happen to meet our moral standards. Society promotes enhanced responsibility by encouraging the development of autonomously responsible character, which is identical to the development of responsible conduct. If our bad conduct does not match our good intentions, then society is willing to recognize excusing conditions. However, society also demands that in the future we take greater care in understanding existing conditions in order to execute our good intentions successfully.

Society naturally demands social responsibility, and its members must take the moral responsibility for enhancing such responsibility in all citizens. Justice, according to social-good-centered theories such as Dewey's, exists for the sake of promoting harmonious social cooperation promoting the social good. Justice therefore demands that each member of society receive the social treatment that corresponds to the degree of socially responsible behavior exhibited. Punishable actions are actions of irresponsible noncooperation that impede the production of social goods.¹⁵ Some form of intervention with respect to the offender, interrupting and altering the offender's patterns of conduct, will be sometimes necessary to modify habits of conduct (less severe offenses may require only admonishments and warnings). Combining Dewey's defense of the prospective theory of punishment with his social-good notion of justice permits us to sketch the outline of a Deweyan model of punishment: Punishment is a type of serious social intervention with respect to the offender that justly aims at enhancing the offender's future moral responsibility to cooperate in the production of social goods. This is not a complete definition of punishment, since the techniques to be used in such an intervention are not yet delineated. But at minimum the techniques must be carefully designed to serve the purpose of enhancing the offender's future moral responsibility. The techniques typically preferred under retributivist models may haphazardly and in part accomplish this same goal, but retributivism is marked by its failure to prioritize the offender's future conduct.

Our future responsibility, our capacity to engage in moral conduct more or less consistently, is the overriding moral concern for a society (especially a democratic society), according to Dewey. Punishment, because it is a practical expression of this social concern for moral conduct, must essentially aim at enhancing future responsibility. Theories of punishment based on abstract principles of justice, free-will autonomy, or calculations upon the consequences of conduct have obscured this fundamental social concern for future responsibility. Dewey's prospective approach to responsibility and punishment is a much-needed third alternative to deontological and utilitarian approaches. The social good requires social responsibility, which in turn requires social interventions in others' conduct as needed. But those interventions must satisfy justice while at the same time morally treating the offender as a moral agent.

The social good cannot simply and directly justify punishment, since that leaves the (utilitarian) problem of justifying punishment of the innocent. On Dewey's moral theory, society has the moral *duty* to punish only the truly guilty because they are guilty, thus agreeing on that point with retributivism. However, the duty to punish the guilty simply and directly *because* they are guilty cannot fully justify punishment, since that leaves the (deontological) problem of determining how punishment should proceed in the face of contextual considerations that weigh against simple "eye for an eye" or proportional punishment. "Punishment only for the guilty" may sound like a damaging concession to retrospective retributivism, but from Dewey's perspective, there is a straightforward explanation needing no retributionist principles: Only the guilty need to have their future conduct modified. On

Dewey's moral theory, society has the *moral* duty to attempt to enhance future responsibility, if indeed society is committed to any genuinely moral standards. This prospective theory is the start of an explanation as to why our practices of assigning responsibility, placing moral blame, and determining punishment can be best understood as social practices that together aim at enhancing future responsible conduct for all members of society.

3. Criticisms of Dewey's Moral-Education Approach

As a type of moral-education theory, Dewey's prospective approach to punishment must answer the usual criticisms directed toward such theories.¹⁶ A full defense against the contentions of rival theories of punishment cannot be mounted here for lack of space and the absence of a detailed Deweyan model of punishment.¹⁷ Still, some preliminary observations concluding this essay might add plausibility to Dewey's approach. To begin, the moral-education theory is sometimes characterized, by both proponents and detractors, as opening the way toward an *alternative* to punishment, and not as offering a different justification of punishment. However, such characterizations presume a retributivist essence to the nature of punishment that Dewey rightly calls into question. A variation on this point is the claim that punishment must be a harm to the offender, not a benefit, and thus a moral-education theory is really just about replacing punishment. From the Deweyan perspective, merely causing harm to the offender is insufficient to morally justify punishment. Therefore, regardless of whether the offender or society finds the moral-education approach's process of punishment to be harmful, that process remains justified through its attempt to increase the offender's future responsibility. The point of punishment is modification of future conduct, not harm for its own sake. Those unable to countenance a process of punishment that might benefit an offender in the long run should reconsider, as Dewey urges in his antiretributivist arguments, how our genuine moral duties and democratic goals suffer at the hands of retributivism.

Another potential difficulty is that a moral-education theory is often defended by appeal to principle(s) of human right or autonomy that are themselves quite controversial and need prior defense. Dewey's moral-education approach, however, does not rely on abstract principles of right or autonomy.¹⁸ His approach does rely on the assumptions that societies encourage moral responsibility and that actual punishment techniques can be designed that will (often enough) succeed in moral education. This opens the door for the objection that the specific modes of punishment that a moral-education approach can endorse are neither severe nor effective enough to prevent future offenses. Like the rehabilitation theory of punishment, the moral-education theory has been reproached for being too soft-hearted and generous toward the offender. How can society show that it takes offenses seriously, retributivists complain, unless stern suffering is imposed?

Society's reasonable need to promote and display its moral commitments can be satisfied by a moral-education theory if we understand with Dewey that our "moral commitment" is our commitment to moral agency (the doer), and not merely to moral conformity (the deed). This understanding of moral

commitment helps to answer the counterpart question of whether moral-education theory might instead be far too severe in its pursuit of moral character. It appears to set no boundaries on the invasiveness and length of punishment to achieve the required responsibility, which conflicts with the proportionality principle, our abhorrence of torture and brainwashing, and our fear of unrestrained government power. Three Deweyan responses are available. First, proportionality would be respected, since the educational effort required should match the degree of increased responsibility needed. Stealing should not be met with decades of reform to achieve moral sainthood. Second, Dewey's moral-education approach is grounded on respect for all persons, even offenders, as moral agents, so that morally unacceptable punishment techniques are ruled out. Third, even where the offense is serious and the offender has prolonged resistance to education, punishment should have a preset, proportional time limit. Without definite sentence terms, punishers might not be sufficiently motivated to maximize their efforts, and the punished would be more likely to resent and resist the punishment process. Furthermore, unrestrained government power is incompatible with the democratic component supporting Dewey's moral-education approach.

Many rehabilitationists have endorsed these three responses as well. A rehabilitationist theory can be pragmatically distinguished from Dewey's moral-education approach if that theory promotes rehabilitation for some reason(s) other than respect for the offender's moral agency. To the extent that such a rehabilitation theory fails to display such respect, it will be unable to respond adequately to one or more of the three concerns raised above. However, it should not be difficult for a rehabilitationist to see the cogency of Dewey's position. A deeper problem still lurks: Perhaps rehabilitationism and moral education may stand together only to fall together. Retributivists often claim that other theories, especially rehabilitationist and moral-education theories, cannot appreciate the kind of justice required for murderers (for example), either because (a) such serious offenders could never be sufficiently rehabilitated, or (b) trying to rehabilitate them simply fails to take their crime seriously by attempting to benefit those who deserve only grave harm. A reply to (b) has been supplied already. As for (a), a moral-education approach should acknowledge that for the most serious crimes like murder, some offenders may never earn our trust in their responsibility and thus never rejoin normal society. A moral-education approach would not regard life in prison as an unacceptable outcome. Its rejection of capital punishment also has the advantage of displaying society's deep respect for human life and moral agency, while at the same time helping the truly innocent by providing for the possibility that a sentence could be overturned by new evidence.

Punishment should aim at teaching offenders and potential offenders why certain types of conduct are wrong and how to avoid choosing such conduct in the future. Another close cousin of moral-education theory, the reintegration model of punishment, also agrees on this principle and shares many essential features with Dewey's approach.¹⁹ Together with rehabilitationism, this family of consequentialist theories rejects retributivism for its inability to provide deterrence strategies that are both intelligent and morally sound. Accordingly, a retributivist's claim that only her theory can supply

sufficient justification for deterrence can no longer be convincing. Pure retributivism cares little for future consequences, and if a retributivist wishes to join forces with the consequentialists instead, then the issue must be joined on the empirical ground of careful studies comparing the actual deterrence provided by the various models. Whether Dewey's own model of moral education would perform well remains to be seen. Only empirical study of actual punishment techniques is relevant here, and the current conflicted state of legal theory on punishment indicates that alternative models and modes of punishment do deserve trial.

A early version of this essay benefited from discussion at the 29th Annual Conference on Value Inquiry in Tulsa, Oklahoma, in 2001. I am deeply indebted to my colleague Eric Reitan for many helpful suggestions along the way.

Notes

- ¹ A useful examination of Dewey's work on responsibility and freedom is that by Abraham Edel, *Ethical Theory and Social Change: The Evolution of John Dewey's Ethics, 1908–1932* (New Brunswick, N.J.: Transaction, 2001), chap. 10.
- ² John Dewey, *The Later Works of John Dewey*, vol. 7: *Ethics*, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1985), 303.
- ³ In H. L. A. Hart's terminology, a "consequentialist" justification of punishment is defined as "essentially forward-looking; it considers the future good we can do to society including the criminal." Hart, *Punishment and Responsibility* (Oxford: Oxford University Press, 1968), 159. Hart's definition is designed with a view toward grappling with utilitarianism. This definition is not used in this essay, and Dewey's theory (as I show) is not another utilitarian effort. Dewey is here arguing that the justification for punishment must involve consideration of its consequences; and so it is a consequentialist approach, but in a broader sense than Hart's.
- ⁴ See Suzanne Rice, "Dewey on Virtue, Character, and Moral Education," *Review Journal of Philosophy and Social Science* 26 (2000): 75–89.
- ⁵ An example of this unification quest is explored by Johann Klaassen, "Punishment and the Purification of Moral Taint," *Journal of Social Philosophy* 27 (1996): 51–64.
- ⁶ This argument deserves more elaboration than can be afforded here, and it is but one way to attack retributivism. For example, as Russ Shafer-Landau has argued, retribution has no resources for determining why harm, and what amount of harm, is deserved by the offender. See Shafer-Landau, "The Failure of Retributivism," *Philosophical Studies* 82 (1996): 289–316.
- ⁷ Dewey, *The Later Works of John Dewey*, vol. 7: *Ethics*, 304–5. Dewey's most fervent diatribe against retributivism (too long for quotation here) is found in his 1922 *Human Nature and Conduct*, in Dewey, *The Middle Works of John Dewey*, vol. 14, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1983), 1–230, at 17–18.
- ⁸ John Dewey, "Force and Coercion," in *The Middle Works of John Dewey*, vol. 10, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1980), 244–51, at 248.
- ⁹ See Joseph Betz, "Violence: Garver's Definition and a Deweyan Correction," *Ethics* 87 (1977): 339–51.
- ¹⁰ On Dewey's political theory see Robert Westbrook, *John Dewey and American Democracy* (Ithaca, N.Y.: Cornell University Press, 1991), 319–73, and Michael Eldridge, *Transforming Experience: John Dewey's Cultural Instrumentalism* (Nashville, Tenn.: Vanderbilt University Press, 1998).
- ¹¹ Dewey's position can also usefully illuminate the problem of past crimes that confronts emerging democracies. A similar view, that the retribution model only increases political tensions by prohibiting the growth of mutual trust, is advanced by Aleksandar

- Fatić, "Retribution in Democracy," in *Political Dialogue: Theories and Practices*, ed. S. Esquith (Amsterdam: Rodopi, 1996), 335–55.
- ¹²John Dewey, "Fiat Justitia, Ruat Coelum," in *The Middle Works of John Dewey*, vol. 10, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1980), 281–84, at 281.
- ¹³Dewey never found persuasive any appeal to transsocial rights. He was particularly skeptical about that sort of alleged rights that have the curious dual properties of being sufficiently general to be possessed by all people simply because they are human, yet particular enough that we can detect when someone's behavior has caused one or more of those rights to evaporate. Daniel McDermott relies on these suspicious dual properties in his attempt to explain why we can harm people simply because a "human" right has somehow vanished. McDermott, "The Permissibility of Punishment," *Law and Philosophy* 20 (2001): 403–32.
- ¹⁴John Dewey, "Philosophies of Freedom," in *The Later Works of John Dewey*, vol. 3, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1984), 92–114, at 94.
- ¹⁵The necessity of punishment for developing stable group cooperation is supported by game-theoretic analyses of reciprocity. Robert Boyd and Peter Richerson conclude that if sufficiently heavy punishments are enforced, moralistic strategies (those that cooperate, punish noncooperators, and punish those who do not also punish noncooperators) can be evolutionarily stable strategies. Boyd and Richerson, "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups," *Ethology and Sociobiology* 13 (1992): 171–95.
- ¹⁶For a liberal formulation of a moral-education theory and some responses to criticism, see Jean Hampton, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13 (1984): 208–38. See also Jeffrie Murphy, "Retributivism, Moral Education and the Liberal State," *Criminal Justice Ethics* 4 (1985): 3–10; and Ido Weijers, "Punishment and Upbringing: Considerations for an Educative Justification of Punishment," *Journal of Moral Education* 29 (2000): 61–73. Moral-education theory and reintegration theory have many connections with the "penance" model of punishment and also to "restorative" justice. See R. A. Duff, *Trials and Punishments* (Cambridge: Cambridge University Press, 1986); Robert Lipkin, "Punishment, Penance and Respect for Autonomy," *Social Theory and Practice* 14 (1988): 87–104; Wesley Cragg, *The Practice of Punishment: Towards a Theory of Restorative Justice* (London: Routledge, 1992); Aleksandar Fatić, *Punishment and Restorative Crime-Handling: A Social Theory of Trust* (Aldershot, England: Avebury-Ashgate, 1995); Francis Schweigert, "Learning the Common Good: Principles of Community-Based Moral Education in Restorative Justice," *Journal of Moral Education* 28 (1999): 163–83; and Kent Roach, "Changing Punishment at the Turn of the Century: Restorative Justice on the Rise," *Canadian Journal of Criminology* 42 (2000): 249–80.
- ¹⁷Such a model is developed in a separate paper that delineates the necessary steps in the intervention process of moral education and defends this process against further objections.
- ¹⁸See John Dewey, *A Common Faith*, in *The Later Works of John Dewey*, vol. 9, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1986), 21–39, and "Democracy and Human Nature," in *The Later Works of John Dewey*, vol. 13, ed. Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1988), 136–55. Dewey's rejection of autonomy for its own sake circumvents Russ Shafer-Landau's criticism that the theory of punishment as moral education is unstably grounded on respect for moral autonomy. Shafer-Landau, "Can Punishment Morally Educate?" *Law and Philosophy* 10 (1991): 189–219.
- ¹⁹See John Braithwaite, *Crime, Shame and Reintegration* (Cambridge: Cambridge University Press, 1989); John Braithwaite and Philip Pettit, *Not Just Deserts* (Oxford: Oxford University Press, 1990); John Braithwaite, "Shame and Criminal Justice," *Canadian Journal of Criminology* 42 (2000): 281–98; and Eric Reitan, "Punishment and Community: The Reintegrative Theory of Punishment," *Canadian Journal of Philosophy* 26 (1996): 57–81.