Communication, Expression, and the Justification of Punishment

By Andy Engen[*]

Some philosophers (Duff, Hampton) conceive of punishment as a way of communicating a message to the punished and argue that this communicative function justifies the harm of punishment. I object to communicative theories because punishment seems intuitively justified in cases in which it fails as a method of communication. Punishment fails as communication when the punished ignores the intended message or fails to understand it. Among those most likely to ignore or fail to understand the message of punishment are the most hardened criminals, whom we typically think are appropriate targets of punishment. I suggest that an alternative justificatory strategy, one that focuses not on the successful receipt of the message of punishment by the wrongdoer but on the expression of condemnation by the community, is not subject to the same worry. The norms of successful expression are more easily met than those of communication, so that expressing condemnation toward a criminal might be justified even if he is unreceptive in the face of the expression.

Introduction

In his influential article "The Expressive Function of Punishment," Joel Feinberg argues that the definition of punishment must include not only its characteristic hard treatment but also its expressive character. Through the act of punishment society expresses both a judgment of disapproval and an emotional response of resentment toward the punished. Feinberg labels this fusion of moral judgment and emotion "condemnation." Prior to Feinberg's article, theories of punishment had focused on justifying the hard treatment of punishment, for example, the way that it deprives a prisoner of freedom of bodily movement. In the article, Feinberg maintains that because both the characteristic hard treatment and condemnation of punishment harm the punished both aspects call for philosophical justification.

Subsequent to Feinberg's article, some philosophers have explored the possibility that the justifications of these two aspects of punishment are related. One group of subsequent theories, which I will call communicative theories, emphasizes that the expressive function of punishment allows the state to communicate to the punished a message about the wrongness of his action, the understanding of which will have some desired effect on him. According to

[*]Assistant Professor, Illinois Wesleyan University, USA.

these theories this *communicative* function justifies the hard treatment of punishment. By depriving serious criminals of liberty we send them important messages about the wrongness of their actions that they need to receive.

In this paper I argue the communicative function cannot provide an encompassing justification of punishment. Commonsense morality holds that punishment is justified in cases in which its message is not understood or accepted by the punished. I will contend that communicative theories do not plausibly account for the justification of punishment in these cases. I will go on to suggest that we might be able to pull from Feinberg a more modest expressive justification of punishment, one that focuses not on the successful receipt of the message of punishment by the wrongdoer but on the successful expression of the condemnation by the community.

Communicative Theories of Punishment and the Problem of the Unreceptive Wrongdoer

In order to give a more concrete picture of what communicative theories of punishment are, I will briefly sketch a couple theories of this sort proposed by Jean Hampton and R. A. Duff. In "The Moral Education Theory of Punishment," Hampton argues that punishment can be justified as an attempt to communicate to the punished and the community at large the wrongness of his action. This purpose of punishment, Hampton maintains, is not peculiar to state punishment, but can be seen in other contexts, such as parental discipline. Hampton writes, "[T]he goal of punishment, whether carried out by the state on criminals or by parents on children, is the offender's (as well as other potential offenders') realization of an action's wrongness" (1984: 214). The hard treatment aspect of punishment makes it a medium well-suited for communicating the wrongness of his act to the criminal, according to Hampton. Punishment disrupts the criminal's pursuits of his own interests and, in harming him, prompts reflection on the harm that he inflicted on his victim (1984: 224-228). The moral education theory offers a powerful response to the challenge of justifying punishment, insofar as it is thought of as the problem of justifying the serious *harm* that hard treatment does to wrongdoers. For on Hampton's account the justifying aim of punishment is not to harm the wrongdoer but to *benefit* him. The moral education theory "attempts to justify punishment as a way to benefit the person who will experience it…. [T]he moral good which punishment attempts to accomplish within the wrongdoer makes it something which is done *for* him, not *to* him" (1984: 214). When punishment successfully educates the wrongdoer, if it makes sense to talk of harm at all, we can justify that harm by the net benefit to the wrongdoer.

R. A. Duff proposes an alternative justification of punishment grounded in the notion that it communicates a message of condemnation to the punished. He rejects moral education theories because he maintains that those who commit serious crimes typically already *know* that what they are doing is morally wrong or thought to be wrong by the community (2000: 91).

Nevertheless, the communicative process of punishment can have the good results of causing what Duff calls the three R's of punishment: repentance, reform, and reconciliation. His theory identifies punishment as a kind of "secular penance" (2000: 106). When the state punishes, it imposes burdens on criminals in an attempt to persuade them "of the error of their ways and to repair the damage done by their crimes to their communal relationships" (2000: 91-92). The hard treatment of punishment focuses the attention of the punished on society's condemnation of her crime in a manner conducive to bringing about her *repentance*. Further, genuine repentance will lead the criminal to *reform* her future behavior. Finally, someone who is genuinely repentant will seek *reconciliation* with her victim. These good outcomes of successfully communicated condemnation give us reason to punish serious wrongdoers. But not only does the state have reason to punish them, it has a duty to do so. The state has a duty to treat its citizens in ways that acknowledge their capacities as autonomous agents and members of the moral community. Duff writes that communicative punishment motivated by the goal of the three R's does just this. Rather than using the punished as a mere means to achieve some social good as purely consequentialist justifications prescribe, the punished are addressed as moral agents, those for whom an attempt at reconciliation is appropriate so that their membership in the moral community can be reaffirmed (2000: 111-112). The state also has a duty the to victim of crime to affirm his value by recognizing the crime as a wrong and attempting to get the one who perpetrated the crime to repent and apologize for it (2000: 113).

Though I have only roughly sketched Hampton and Duff's complex theories, I hope it is clear how they both fit my classification of communicative theories. Both hold that punishment is justified in its role of sending a message to the punished with the aim that the punished understand the message and that understanding has some desired effect on him. Much could be said for and against these theories, but I want to focus on a particular worry: that they are not able to justify punishment in those cases in which the punished are not receptive to the message conveyed by their punishment. I will use the term "The Unreceptive Wrongdoer" as a placeholder for a number of possible cases in which the punished does not respond in the desired way to the message punishment sends. When we communicate with someone while treating her as a rational agent, we must leave open the possibility that the communication might fail. One of the differences between rational communication and implantation of beliefs via brainwashing or hypnosis is that it is open to the object of communication to reject the message being communicated. The object of communication is an active participant in the exchange.[1]

At first glance, it seems that The Unreceptive Wrongdoer poses difficulties for communicative theories. A criminal who ignores the condemnatory message of punishment, or misunderstands it, or understands and rejects it will

[1]This is a point Duff is clear on in his own treatment of such cases. "[A]n exercise in forceful moral persuasion is, like any exercise in rational communication, necessarily fallible: not just because it might in fact fail to persuade him, but because it must leave that possibility open" (2000: 22).

not be morally educated by that message. The message will not cause him to repent for his crime and seek to reconcile himself with its victim. The justifying aim of theories such as Hampton and Duff's will not be achieved by punishment in such cases. Intuitively, however, punishment of The Unreceptive Wrongdoer can still be justified.[1] Among those most likely be unreceptive to the message are the most hardened criminals, whose thinking and behavior is least responsive to moral messages sent by the state or their fellow citizens. We think that imprisonment of unrepentant murderers is justified even if it does not reform them, and even if we have no legitimate expectation that it will. If someone commits a horrific crime, it does not seem that the question of whether he ought to be punished turns on whether he will be receptive to message of his punishment. While moral education and reconciliation are desirable possible outcomes of punishment, their realization is not required in order for instances of punishment to be justified.

Hampton and Duff anticipate the Unreceptive Wrongdoer in the presentation of their respective theories. They both emphasize that punishment can be justified as an *attempt* at communication, even when that attempt fails.[2] Generally, failed attempts can be justified even if they are ultimately unsuccessful. For example, I might fail the bar exam after I graduate from law school, but that failure does not make my attempt unjustified. Communicative theories can maintain that even if the state cannot guarantee punishment will lead to moral education or reconciliation, we at least ought to strive for those results by punishing criminals.

---

[1]Communicative theories could retort, "so much the worse for our intuitions." At some points both Hampton and Duff criticize current practices of punishment as being morally unjustifiable. Nevertheless, they do not take this route in responding to The Unreceptive Wrongdoer.

[2]Both philosophers propose other responses as well. About punishment that fails to educate the punished, Hampton responds, "But at least the punishment can have a deterrent effect; even if the criminal refuses to understand the state's communication about why there is a barrier to his action, at least he will understand that a barrier exists…. [E]ven if it seems likely that punishing some criminals will not effect their moral growth, and may not even deter them, the moral education of the community about the nature of their crimes can still be promoted by their punishment" (1984: 231). I contend that this response does not evade the problem of nonreceptivity. Intuitively, justified punishment can fail to effectively communicate a deterrent message to the punished or an educative message to the wider community. Duff responds that even when the punished does not actually repent, reform, and reconcile, his punishment can be understood as a symbolic reconciliation. "Although, if his punishment was simply inflicted on him, he might not even have gone through the motions of apologizing, we can extend the idea of an apologetic ritual to cover this kind of case. The offender has been subjected to what would constitute an appropriate reparative apology if he undertook it himself. His fellow citizens should therefore now treat him as if he had apologized…. He might not have paid the apologetic debt that he owed, if his punishment was simply inflicted. But something like that debt has been exacted from him, and those who exacted it should now treat him as if the debt has been paid" (2000: 123-4). I believe that this proposal is at odds with Duff's claim that we should respect the autonomy of the punished. Among other things, respecting someone's autonomy requires acknowledging the choices that he has made as being his choices and treating him as if those choices matter. Accepting symbolic reconciliation does not respect the wrongdoer's autonomy in this regard. It requires that The Unreceptive Wrongdoer be treated as if he chose other than he did.

Though initially plausible, I think that this response is problematic because of the teleological nature of attempts. The rationality of any attempt is determined in part by the chance the attempt has of succeeding. Holding other facts of the case fixed, I have greater reason to attempt to take the bar exam if I have studied law than if I have not and if the exam is written in a language that I can read than if it is not. In fact, if the exam were in a language that is completely foreign to me, that might make the likelihood of success so remote that I would not be justified in attempting to pass it. Likewise, if punishment were justified as an attempt to communicate, whether or not we have reason to punish in a particular case would seem to partly depend on whether we had reason to believe that communication would be successful in that case. We would seem not to have reason to punish the criminals most likely to be unreceptive to the message of their punishment, or at least we would seem to have much less reason to punish them than those criminals we have reason to think would be receptive. Insofar as neither of these are the case—we have reason to punish horrific crimes regardless of whether we have reason to believe the punishment will lead to, say, reform—it does not seem that punishment can be justified as an attempt at communication.

Thus far, I have focused on the teleological nature of attempts in rejecting the idea that punishment can be justified as an attempt at communication. To an extent, however, both Hampton and Duff put our reason for attempting to communicate with criminals in deontological terms: an attempt at communicative punishment is something that is *owed* to criminals or their victims. In Duff's view, the community "owes it to the victim… and to the offender as a member of the normative community, to try to get the offender to recognize [the wrong done] and to make a suitable apology for it" (2000: 214). Hampton understands the right to be punished as a right to the attempt at moral education through punishment. "[A]s Hegel puts it, punishment is a criminal's 'right' as a free person—to refuse to punish him on the grounds he has been diagnosed as hopeless is to regard him as something other than a rational being" (1984: 231). Deontological considerations can give us strong reason to attempt to do something independently of its chances at success. If I, say, promise someone I will try to do X, I have strong reason to try to do X regardless of its chances for success.

Nevertheless, I believe that the teleological nature of attempts raises difficulties for communicative theories, even when the rationality of the attempts is understood in a deontological framework. If one believes that one has no chance to succeed at some task, it is unclear that one can even attempt that task. There seems to be a practical irrationality in doing so. In fact, if someone apparently attempts to do what she occurrently believes to be impossible, it is natural to describe her action not as a genuine attempt at what she takes to be impossible but as acting *as if* attempting to do what she is certain will fail. Imagine that we walk by someone jumping up and down outside the Parthenon while saying, "I'm certain I can't do it but I'm trying to jump over the Parthenon." We are likely to interpret this as an absurd joke rather than taking her at her word.

If what we owe someone is an attempt, we cannot deliver what is owed when we believe that the attempt cannot succeed. Imagine that the science of criminal psychology has progressed to the point where scientists can predict with absolute certainty in individual cases whether any sort of punishment we can conjure up will lead to a genuine apology or the right kind of understanding in the wrongdoer. Intuitively, we would have reason to punish those who commit terrible crimes even if we knew that punishment could not be a rational attempt to communicate because the science told us beforehand any such attempt is doomed. If this is correct, the conception of punishment as an attempt to communicate cannot account for all cases of justified punishment.

Expression vs Communication

I have argued that The Unreceptive Wrongdoer poses a serious problem for communicative theories of punishment. It is too quick to conclude from this, however, that the expressive function of punishment cannot account for the moral reason the state has to subject criminals to hard treatment. In the remainder of paper I will suggest that the failure of communicative theories of to provide an expressive justification of punishment need not doom the prospects of all expressive theories.

In Feinberg's original piece, he did not emphasize the function punishment as *communicating* condemnation with the punished, as much as *expressing* the condemnation of the community.[1] Messages that are communicated are also expressed, but the norms of successful expression differ from those of successful communication. In particular, whether or not something is expressed is largely under the control of the person doing the expressing, whereas whether or not something is communicated also depends on the receptiveness of the person to whom it is intended to be communicated. For example, were I stranded on a deserted island, I might successfully express the fact that I need help by screaming "Help!" without communicating it to anyone. Due to this difference between the norms of successful expression and the norms of successful communication, a theory that justifies punishment as an expression of condemnation rather than as an act of communication is not subject to worries about The Unreceptive Wrongdoer. Punishment can express the community's condemnation even when that condemnation is ignored or misunderstood by the punished.

Because the punisher is in control of whether punishment expresses condemnation, the "mere" expressive function of punishment does not have the same difficulties fitting into a deontological framework as the communicative

[1]I will focus on the expressive function of punishment generally, as Feinberg understands it. Philosophers drawing on Feinberg who have argued for a noncommunicative expressive justification of punishment include Primoratz (1989), Narayan (1993), and von Hirsch (1993). If my argument of this section is convincing, it provides a reason to prefer those theories to communicative theories.

function does. If, say, what the victims of crime are owed is condemnation of the way they have been wronged, rather than attempts to get those who have wronged them to repent, reform, and reconcile with them, we can give them what they are owed even when it is unlikely that those who have wronged them will ever repent of their misdeeds.

There is reason to think that "mere" expression of condemnation is valuable to victims of crime.[1] When the community publicly condemns a crime committed toward some victim it publicly affirms the importance of the standard of conduct that has been violated and the value of the victim. Feinberg points toward something like this in his original article where he identifies "authoritative disavowal" as one of the derivative functions of the expressive role of punishment. Through expressive punishment, the community says that it will not stand for a certain sort of behavior. To show the importance of this function, Feinberg describes a law in Texas that permitted the killing of a man found in the act of adultery with ones wife. He claims that the belief that contrary to that law such killers ought to be punished may "represent the feeling that the paramour killings deserve to be *condemned*, that the law in condoning, even approving of them, speaks for all the citizens in expressing a wholly inappropriate attitude toward them" (1965: 103). T. M. Scanlon highlights the way that the expressive function of punishment affirms the value of victims of crime.

> [O]ne thing citizens may reasonably demand of a system of law is that it affirm their rights and, in particular, affirms a victim's sense of being wronged. To this I would add that a system that affirms their sense of being wronged must condemn the agent who inflicted the wrong…. (1999: 231)

Punishment that expresses condemnation, then, plays the important social role of affirming the rights of crime victims and this could be among those things that crime victims are owed by the state.

I have argued that The Unreceptive Wrongdoer gives us a reason to reject communicative, but not all expressive, theories. Punishment can succeed in expressing condemnation even when the punished is unreceptive to the message of condemnation. Thus, expressive theories can account for our conviction that those serious criminals who are unlikely to be receptive to the message of their punishment should nevertheless be punished.

This argument has implications for the capacity of expressive theories to capture other features of our moral thinking concerning punishment. For example, consider what is perhaps the most challenging objection to expressive theories, which I will call "Alternative Modes" objection.[2] The objection

[1]It also, in a sense, gives the punished something they are owed by treating them as members of the moral community, to whom moral obligations are properly addressed.

[2]Both Feinberg (1965: 115-116) and Scanlon (1988: 214) appeal to a version of the Alternative Modes objection in rejecting the idea that the expressive function of punishment justifies its

emphasizes that expression can take many forms, and claims that expressive theories must explain why punishment should be the mode of expression. If its message could be expressed without the characteristic hard treatment of punishment, the expressive role of punishment would not justify that hard treatment. At first glance, it may seem as though communicative theories like those offered by Duff and Hampton have a compelling response to this objection. They can maintain hard treatment is the most effective way to communicate to the punished in a way that will lead to education or reconciliation. For instance, Hampton claims the hard treatment of punishment efficaciously communicates to criminal the wrongness of their actions because it can lead them to reflect on their victims' suffering (1984: 227).

This response to the Alternative Modes objection is not available to theories that focus on punishment's role in expressing, but not necessarily communicating, condemnation. A possible response on behalf of this sort of theory could draw upon Feinberg's notion that punishment "speaks for all citizens." In order for an expression to speak on behalf of all citizens as expressing condemnation, it must be recognizable to those citizens as expressing condemnation. Some potential modes of expression, say giving criminals large amounts of money in response to their crimes, would be ruled out as unrecognizable to citizens as an expression of condemnation. Further, in response to serious wrongdoing a high degree of condemnation is called for, so that in order for an expression to speak for all citizens in response to serious crimes, that expression must be recognizable to all citizens as expressing a high degree of condemnation. Punishment unambiguously expresses a high degree of condemnation, whereas citizens might wonder if responses that did not involve hard treatment took seriously the wrongdoing. Punishment thus speaks on behalf of the entire community in condemning serious crimes.

The respective responses that I have sketched for the communicative and merely expressive theories would need to be further developed to address the Alternative Modes objection. For my purposes it is enough to emphasize the following difference. If successful, the merely expressive response would explain why punishment of The Unreceptive Wrongdoer is called for rather than alternative modes of expression, while the communicative response would not. Insofar as Hampton and Duff argue that the hard treatment of punishment is necessary to effectively communicate the message of punishment, the argument does not apply to those cases in which punishment cannot accomplish this goal. Sending a criminal an introductory ethics textbook would just as well (or poorly) discharge any duty to "attempt" to educate and achieve the 3 R's of repentance, reform, and reconciliation.[1]

I conclude that The Unreceptive Wrongdoer poses a significant and compelling objection to communicative theories, but not to all expressive theories.

---

hard treatment. More recent critics include Boonin (2008: 176-179), Brooks (2012: 117-118), Hanna (2008), and Tadros (2011: 108-109).

[1]C. L. Ten makes a similar point in objection to an earlier formulation of Duff's view (1990: 204).

Acknowledgements

I'm grateful for comments I received on versions of this paper from Alan Coffee, Niko Kolodny, Samuel Scheffler, and R. Jay Wallace, as well as audiences at the ATINER 9th Annual International Philosophy Conference, the International Conference on the Philosophy of Criminal Punishment at CUHK, and the 4th Annual Berkeley-London Conference.

References

Boonin, David, 2008, *The Problem of Punishment*. New York: Cambridge University Press.

Brooks, Thom, 2012, *Punishment*. New York: Routledge.

Duff, R. A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press.

Feinberg, Joel, 1970, "The Expressive Function of Punishment," in his *Doing and Deserving*, Princeton, N. J.: Princeton University Press.

Hampton, Jean, 1984, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs*, 13: 208-238.

Hanna, Nathan, 2008, "Say What? A Critique of Expressive Retributivism," *Law and Philosophy*, 27: 325-349.

Narayan, Uma, 1993, "Appropriate Responses and Preventive Benefits: Justifying Censure and Hard Treatment in Legal Punishment." *Oxford Journal of Legal Studies*, 13: 166–82.

Primoratz, Igor, 1989, "Punishment as Language," *Philosophy*, 64: 187–205.

Scanlon, T. M., 1988, "The Significance of Choice" in *The Tanner Lectures on Human Values, Vol. 8*. University of Utah Press.

Scanlon, T.M., 1999, "Punishment and the Rule of Law"; reprinted in his, 2003, *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge: Cambridge University Press.

Tadros, Victor, 2011, *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Press.

Ten, C.L., 1990, "Positive Retributivism," *Social Philosophy and Policy*, 7: 194-208.

von Hirsch, 1993, *Censure and Sanctions*, Oxford: Oxford University Press.