I will criticize a proposed justification of punishment. Some theorists think that the challenge of justifying punishment can be met by emphasizing its expressive character, i.e., its capacity to express criticism. One such justification, or type of justification – call it Expressive Retributivism or ER – combines retributivist and expressivist considerations. These justifications are retributive since they do not appeal to consequences, aiming instead to show that punishment is an “intrinsically” appropriate response to offenses, i.e., something offenders deserve.\(^1\) The expressivist element in these justifications, however, is meant to correct for the widely criticized obscurity of traditional retributivism.\(^2\) Retributive arguments often rely on controversial intuitions of questionable reliability and justificatory power.\(^3\) While these intuitions are powerful, many worry that the justificatory challenge cannot be met merely by

\(^1\) In the philosophical literature on punishment, the word *intrinsic* is often meant to indicate intuitive immediacy apart from external considerations like good or bad consequences. See, e.g., Mackie 1982.

\(^2\) For examples of expressive justifications that also appeal to consequences see von Hirsch 1993 and Narayan 1993.

\(^3\) For a well-known, relatively recent statement of retributivism as well as criticism of other retributivist justifications see Moore 1993.
appealing to them. Given the fact that punishing people is to treat them in ways that are typically wrong, a plausible justification seems to require more than appeals to the prevalence and strength of these intuitions. Many voice suspicion that retributive arguments are not justifications so much as refusals to furnish justification. ER, by contrast, tries to enhance the clarity and justificatory power of retributive intuitions and concepts by appealing to the expressive character of punishment. I argue that the ER justification fails.4

I begin, in section I, by discussing the nature of punishment and outlining an expressive conception of it. I argue for a slightly modified conception on which the aim to impose suffering is an essential characteristic of punishment (in a somewhat loose sense of suffering). In section II I sketch the ER argument and identify three crucial claims, each of which I examine in turn. I round out section II with a brief discussion of the first claim, arguing that one important version of it proposed by some theorists is false. In section III I discuss the second claim. I grant its truth but offer an account of the expressive use of punishment that will serve to undermine the third claim in light of my conception of punishment. In section IV I show how this account and my conception of punishment undermine the third claim and the arguments that a number of theorists have offered in its defense.

II. THE NATURE OF PUNISHMENT

Joel Feinberg argues that punishment is more than the infliction of hard treatment on alleged offenders for their alleged offenses: it is “a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation” (400).5 This characteristic, according to Feinberg, helps distinguish punishment from other kinds of penalties and accounts for various functions that punishment performs, functions

---

4 I will discuss retributive expressive justifications throughout, but much of the discussion will bear on the acceptability of expressive justifications generally.

5 Unless otherwise noted, all references in this section are to Feinberg 1965.
that "would be impossible without it" (400, 404–408). In light of these considerations, he concludes, "both the hard treatment aspect of punishment and its reprobative function must be part of the definition of legal punishment" (400).

Feinberg’s talk of hard treatment invites some questions. What is hard treatment and how is it related to punishment? Feinberg does not offer a definition of hard treatment, but he offers some examples, e.g., imprisonment and stiff fines. Presumably, hard treatment is typically painful or unpleasant. As Feinberg seems to suggest when referring to H. L. A. Hart’s definition of punishment, hard treatment "involve[s] pain or other consequences normally considered unpleasant" (Hart 1959 [1968]: 4). But not just any involvement will suffice; painfulness and unpleasantness are not sufficient for hard treatment, even when they result from intentional actions. Feinberg, however, seems not to acknowledge this, suggesting at times that painful treatment simply is hard treatment. He claims that "[r]eprobation is itself painful, whether or not it is accompanied by further 'hard treatment'" (400, emphasis added) and that "it is social disapproval and its appropriate expression that should fit the crime, and not hard treatment (pain) as such" (423). He understands reprobation as a "stern judgment of disapproval" (403) that is painful and apparently considers it a form of hard treatment because it is painful.

Feinberg seems to think that strong criticism qualifies as hard treatment at least in part because it is typically painful or unpleasant. But it seems implausible to claim that someone who strongly criticizes someone else thereby subjects her to hard treatment. One may experience a variety of painful and unpleasant feelings like embarrassment or disappointment if someone strongly criticizes one’s scholarly work, one’s fashion sense, one’s hygiene and so on, but it does not look like one is necessarily subjected to hard treatment in such cases. Similar unpleasant feelings can result from other actions that do not seem to constitute hard treatment, including romantic rejection, medical treatments and so on.

Painfulness and unpleasantness do not seem to suffice for hard treatment – at least, they do not seem to suffice for the sort of hard treatment essential to punishment. Antony Duff claims
that punishment “aims to inflict something painful or burdensome on an offender for his offense” (Duff 1992: 49; see also, among others, Golash 2005: 45, 77–78; Kleinig 1991: 402, Primoratz 1989a: 1, 48, Sayre-McCord 2001: 2–3). According to Duff, the pains of punishment are not “mere unintended side effects” of punishment, but are among its objectives (Duff 1992: 49). J. R. Lucas makes a similar claim.

Many things are done against our will, but not in order to hurt or frustrate us. [...] Many actions are hurtful and known to be hurtful, and yet are done: but they are not done in order to be hurtful, and if there were an equally effective but less hurtful way of accomplishing the agent’s purposes, he would be perfectly ready to adopt them. [...] Punishments, by contrast, not only are unwelcome but are intended to be, and would lose their point if they were not. (Lucas 1968: 207)

Many acknowledge that punishment is intentionally imposed and that it usually involves painfulness or unpleasantness (e.g., Hart 1959). But theorists like Duff and Lucas also claim that the intentions behind punishment have a common content that is essential to punishment. To punish an offender, one must aim to hurt or harm her in some way; the treatment to which she is subjected must at least in part be used as a means of hurting or harming her – to make her suffer, in a rather loose sense of the term that I will use – if it is to constitute punishment. The claim that the aim to impose suffering is essential to punishment has a certain intuitive appeal, but there are other considerations that speak in its favor. The aim invariably influences the way punishments are applied and so helps to account for significant differences between punitive and non-punitive treatment, e.g., differences between imprisonment and other kinds of confinement. Medical quarantine, involuntary psychiatric commitment and protective confinement all cause suffering. Ideally, however, steps are taken to minimize the suffering they cause. This is not the case with imprisonment. Offenders are imprisoned, at least in part, in order to make them suffer and prison conditions are designed in service to this aim. While the critical symbolism of imprisonment may, as Feinberg notes, contribute to the suffering prisoners experience, (418) it does not look like this symbolism fully accounts for the differences between the way imprisonment and these other
forms of confinement are applied. In fact, the symbolism can plausibly be said to rest on these differences. Imprisonment is symbolically critical, not simply because it is a kind of confinement, but because it is imposed and applied in order to make offenders suffer for what they have done, whereas non-punitive forms of confinement are not imposed or applied with this aim. What I have said about imprisonment, it seems, can be said of punishment generally.

This aim may also help Feinberg make the distinction he wants to make between mere penalties and punishments. While he clearly thinks that appealing to punishment’s critical character can help make the distinction, he stops short of saying that it can do all the work: “Punishment, in short, has a symbolic significance largely missing from other kinds of penalties” (400, emphasis added and omitted). The hesitance here may stem from the fact that some mere penalties and some punishments involve similar forms of hard treatment in Feinberg’s sense of hard treatment. Non-punitive monetary penalties and punitive fines can be of comparable monetary value, for example. Since critical character, on Feinberg’s view, attaches to types of penalties, (402) such comparisons pose difficulties for making the distinction on the basis of critical character alone. Appealing to the aim to impose suffering can help with such cases (which is not to say that it can eliminate all the difficulties with making the distinction). Mere penalties, it might be argued, may cause suffering, but the suffering they cause is incidental and steps may be taken to reduce it. In contrast, some of the suffering caused by punishment cannot plausibly be said to be incidental in this way, but is rather one of the things aimed at in punishing.

The aim to impose suffering should, then, also be counted among the essential characteristics of punishment. The characteristic is intuitively compelling, serves to account for important differences between different types of treatment and helps make what some theorists, including Feinberg, take to be important distinctions. In what follows I will assume that the aim to impose suffering is essential to punishment and will
argue that this characteristic, in conjunction with other considerations, undermines ER.

III. CONCEPTUAL CLAIMS AND THE ER ARGUMENT

In this section I sketch the ER argument, isolate three crucial claims, briefly discuss the first claim and argue that one important version of it is false.

The ER argument can be roughly sketched as follows. Wrongdoers (or offenders – I will use the words and their variants interchangeably) must be criticized for their wrongdoing. Criticism must be proportional to wrongfulness, i.e., the greater the wrong, the harsher the criticism called for. Punishment is a particularly emphatic way of expressing criticism and some wrongs call for a degree of criticism that can only be expressed by punishment. Hence, some wrongs can be adequately criticized only by means of punishment. Adequate criticism is preferable to inadequate criticism. Hence, punishing some wrongs is preferable to not punishing them.

This sketch is probably insufficient in a number of ways, but it serves my purposes for the moment. It captures some important claims common among ER arguments: wrongdoers must be criticized for their wrongdoing, punishment expresses criticism and only punishment can express the degree of criticism called for by some wrongs. If some of these claims are false, or at least misleading, then ER fails. I will examine each claim in turn. One important version of the first claim is false, the second claim is often misleading in the context of ER arguments, and the third claim is false.

Consider the first claim: wrongdoers must be criticized for their wrongdoing. There are at least two ways to take this claim: as a normative claim or as a conceptual claim. A number of theorists make the claim in the second sense. Duff thinks if I declare firmly that a certain kind of conduct … is seriously wrong, this … commits me to certain judgments on the conduct of others – to the judgment that those who engage in such [conduct] act wrongly. But it also commits me to expressing those judgments in certain situations. If I find that
[someone] has engaged in such [conduct], I must express my view that [he] acted wrongly by criticizing or censuring [his] action. To remain silent, to let the action pass without criticism, necessarily casts doubt on the sincerity of my declaration that such conduct is seriously wrong. (Duff 2001: 28, emphasis added and omitted)

In less explicit terms, Igor Primoratz seems to be making similar claims about rule violations as well as wrongs.

Rules that state standards of behavior and command categorically imply that actions violating them are wrong, and that such actions are to be condemned, denounced, repudiated. Expressions of this condemnation and repudiation are the index of the validity of the rules and of the acceptance of the conviction that their breaches are wrong in society. If actions of a certain kind can be done without bringing about such a response from society, this indicates that no rule prohibiting such actions is accepted as a valid and binding standard of behavior. (Primoratz 1989b: 196).

Both Duff and Primoratz are making similar claims and proposing similar arguments. On their view, our responses to wrongdoing and rule breaking express attitudes. If I claim that a certain sort of act is wrong, (or that it violates a valid rule) then I will criticize such acts. If I do not do so, then, ceteris paribus, I cannot really think it wrong (or think that it violates a valid rule) – I must be confused or insincere and cannot really mean what I say.

I will reserve extended discussion for ER’s other two claims.

I want to make one point, however. The conceptual claim seems patently false. Contra Duff and Primoratz, it is possible that a person or society could react indifferently or even positively to a particular wrong or a particular wrongdoer without being insincere or confused about the relevant concepts. The same can be said of particular rule violations and even of wrongdoing and rule-breaking generally. There is no apparent absurdity in reacting positively to wrongdoing or rule breaking qua wrongdoing or rule breaking. There are all sorts of things that can be said about a person or a society that reacts like this, e.g., that they are morally perverse, but it does not look like they must be inconsistent or confused.6

---

6 Whether such reactions are likely to result in bad consequences is a different matter and irrelevant to the logical point here.
I do not want to place too much weight on this worry. I mention it because some theorists have taken pains to press a conceptual claim that they think is obvious and significant but which seems false (and perhaps not relevant to their arguments; cf. Adler 2000: 1426). Such theorists may be running conceptual and normative claims together or they may be trying to strengthen the latter by appealing to the former. I cannot diagnose the underlying errors and motivations here. At any rate, ER can probably get by without the conceptual claim. A more cautious argument can get off the ground with the simple normative claim I mentioned. I will not challenge such a claim, so I will move on to the others.

IV. CONVENTIONALISM AND PUNISHMENT

ER’s second claim is that punishment expresses criticism. I grant this claim, with a caveat: the use of punishment to express criticism is conventional. In the context of ER arguments the second claim is often used in misleading ways, setting the stage for implausible arguments for the third claim, a uniqueness claim. If we can get a handle on how and why punishment is used to express criticism, we can see why the third claim is false.

Feinberg emphasizes punishment’s conventionalism – specifically the conventionalism of the hard treatment element in punishment. In societies like ours certain forms of hard treatment have become the conventional symbols of public reprobation. This is neither more nor less paradoxical than to say that certain words have become conventional vehicles in our language for the expression of certain attitudes, or that champagne is the alcoholic beverage traditionally used in celebration of great events, or that black is the color of mourning. (Feinberg 1965: 402).

This poses a problem for ER. If the use of punishment to express criticism is conventional, the role that appeals to the need for criticism can play in a justification of punishment are

---

7 I should note that Adler criticizes the justifications offered by many of the theorists I will discuss here. His treatment of their justifications is rather brief, however, and not all of his criticisms strike me as terribly clear or as decisive as the ones I offer.
limited. Showing that some particular set of conventions calls for punishment will not suffice. Those conventions themselves require justification. If criticism can be expressed in other ways, what can the ER theorist say in favor of punishment? Some think that there is nothing he can say.

Pointing out “the expressive function of punishment” helps us to understand our reactions to punishing particular kinds of people, but what role if any does it have in the justification of punishment? It seems to have no positive role. [...] Insofar as expression is our aim, we could just as well “say it with flowers” or, perhaps more appropriately, with weeds. (Scanlon 1988: 214)

Appealing to the need to express things like criticism in order to justify a convention that uses punishment to express such things risks making a question begging appeal to conventional standards. It does not look like appeals to expression and the need for criticism, then, can play a significant justificatory role.

Consider Matt Matravers’ reply to Scanlon. Matravers argues that Scanlon’s claim “is not true. We could not just as well say it with flowers because flowers typically say something else” (Matravers 2000: 257, note 19). This is true, but it seems to overlook Scanlon’s point (or, at least, what I take Scanlon’s point to be). Matravers is appealing to the standards of an existing convention, one where punishment is typically used to express one thing and flowers another. But Scanlon’s point seems to be that other conventions could be adopted. Punishment does not seem to be uniquely capable of expressing the things it expresses. Other conventions could be adopted that confer the requisite expressive powers on other sorts of treatment. Even if criticism must be expressed, the task remains to justify the use of punishment, instead of something else, to express it.

But this might be too quick. Consider an explicit denial of the conventionality claim. Anthony Skillen claims that the use of punishment – more precisely, of the hard treatment element in punishment – is

---

8 Earlier in his book, however, Matravers does seem to note the need for justifying such conventions. See Matravers 2000: 79.
hardly purely conventional. [...] Whereas black is arguably neutral in itself and only contextually and conventionally constituted as mourning wear ... it is pretty clear that losing money, years of liberty, or parts of one’s body is hardly neutral in that way. This suggests a radical inadequacy in Feinberg’s account. Feinberg vastly understates the natural appropriateness, the non-arbitrariness, of certain forms of hard treatment to be the expression or communication of moralistic and punitive attitudes. Such practices embody punitive hostility, they do not merely ‘symbolize’ it.⁹ (Skillen 1980: 517)

There may be justice in the charge that it is misleading to claim that the use of punishment to express criticism is conventional. Perhaps the claim risks overlooking facts about human nature. Given our aversion to suffering, it seems inappropriate to use penal hard treatment to express things like praise and approval. Given that we typically respond to being hurt and wronged with anger, resentment, hatred and a host of emotions that motivate the infliction of harm, it seems inappropriate to adopt conventions that would have us treat wrongdoers in ways that they find pleasant.

None of this seems to speak against Feinberg’s point, however. For one thing, Feinberg is not talking about punitive hostility, a term that Skillen substitutes for Feinberg’s term reprobation. Skillen’s claim is trivially true and in the context of a proposed justification of punishment it would be question begging. It is trivially true because punishment is obviously uniquely suited to the expression of punitive hostility (as opposed to hostility simpliciter) and to the expression of moralistic and punitive attitudes (as opposed to moralistic attitudes alone). It would be question begging unless preceded by the question of whether punitive attitudes can justifiably be expressed.

Moreover, calling the use of punishment conventional, as Feinberg does, does not imply that one thinks it purely conventional, a phrase probably meant to suggest that all possible conventions are equally acceptable. Neither Feinberg nor a critic of expressive justifications needs to say that. One could say that the use of punishment is conventional in the sense that the choice between it and a limited set of other options that could, with the right conventions, effectively convey criticism is

⁹ Primoratz endorses this claim. See Primoratz 1989b: 199.
arbitrary, at least in terms of expressive power. Among all possible conventions, some may be more naturally suitable than others, but that does not mean that punishment is not, in a sense that bears on its justifiability, just one possible way of expressing criticism.

In conjunction with the fact that the aim to impose suffering is an essential characteristic of punishment, I argue that the fact that punishment is only conventionally used to express criticism poses a significant problem for ER. In the remainder of the paper I will show how these considerations undermine ER’s third claim and arguments that have been offered in its defense.

V. THE UNIQUENESS CLAIM

I will now examine the third claim, the claim that only punishment can express the degree of criticism called for by some wrongs. ER’s justificatory task is complicated by the fact that aiming to impose suffering is essential to punishment and by the fact that the use of punishment to express criticism is conventional. Because ER theorists have tended to overlook the significance of one or both of these facts, they have underestimated their justificatory burden. I will show how these facts about punishment undermine arguments for the uniqueness claim offered by contemporary ER theorists.

Antony Duff and Jean Hampton are among the most prominent contemporary ER theorists. They both explicitly make the uniqueness claim.  

If a communicative theory is to justify hard treatment punishments, it must show that penal hard treatment is not just a possible, but a necessary, method of communicating the censure that offenders deserve. (Duff 2001: 29)

The infliction of suffering is necessary to counter the falsehood implicit in [serious wrongdoing]. (Hampton 1992a: 12, emphasis added)

The uniqueness claim is motivated by a simple worry. Criticism can be expressed verbally. If we want to express criticism, as

10 H. L. A. Hart and Andrew von Hirsch also attribute the uniqueness claim to a number of theorists. See Hart 1963: 65 and von Hirsch 1993: 12
Duff would have us do, (he uses the word *censure*) or if we want to make statements about human worth, as Hampton proposes, then why not do so verbally? Punishment, after all, treats people in ways that are wrong under most circumstances. Just because we can express certain things with punishment does not mean that doing so is justified, especially if adequate alternative means of expression are available. The uniqueness claim gets around this worry by denying that any such alternatives are available. This makes for a stark choice: express adequate criticism (or whatever else we want to express) by means of punishment or do not express adequate criticism. The former option, it is assumed or argued, is preferable.

H. L. A. Hart questions the uniqueness claim and doubts whether its truth would justify punishment. He wonders whether the expression of criticism is valuable enough to justify punishment and he questions whether verbal criticism is insufficient to express the criticism called for by any offenses (Hart 1963: 65–66; cf. Matravers 2000: 78, Narayan 1993: 178). Andrew von Hirsch seems to have similar worries that lead him to reject the ER strategy and appeal both to consequentialist considerations and to the need for criticism (von Hirsch 1993: 12). Deirdre Golash goes a bit further, considering possible aims we might have in communicating criticism, arguing at length that many of them do not seem important enough to justify punishment and arguing that punishment cannot reliably achieve many of them anyway (Golash 2005: 117–146).

I am not going to weigh in on the importance of various values or on the efficacy of punishment in achieving certain goals. There is, I think, little benefit in such debates, at least in this context. Like Hart, Duff and many others I question the uniqueness claim. I will, however, criticize it in detail. ER theorists offer compelling examples that many think call for more than mere verbal criticism. Rather than question underlying assumptions about the value of criticism or the reliability of our intuitive reactions to such examples, and instead of brushing such examples aside with a rhetorical skeptical question or two, I will consider the arguments for the uniqueness claim and show how they fail. The arguments offered on its
behalf are flawed (in ways that highlight difficulties with arguments in favor of punishment generally – though I cannot discuss that here). The standard strategy is to compare punishment to a small set of alternatives: doing nothing, expressing criticism informally and verbally, or expressing criticism formally but by means of things like purely formal convictions and official reprimands. Such alternatives, it is argued, are insufficient to express the requisite criticism. Hence, punishment is necessary to do so.

Duff and Hampton offer sophisticated justifications of punishment, but their arguments for the uniqueness claim still fit this characterization. To demonstrate this I have to explain their views, though. Let me first, then, discuss some theorists who make this argument in simpler terms. I will then move on to Duff and Hampton.

Igor Primoratz and John Kleinig both consider using verbal criticism instead of punishment.\textsuperscript{11} They both find it unacceptable for similar reasons. [People] would surely see purely verbal condemnation of crime, however public and solemn, as half-hearted and unconvincing. The state would be seen as promulgating laws which determine some of our most important rights – to life, bodily integrity, property – and pronouncing their violations to be crimes, and then responding to their violations only by issuing verbal statements of disapproval. It would be seen as desisting from activating its apparatus of force and coercion, which is surely one of its essential, defining features. [...] In view of all this, it seems to me that both those whose rights were being violated and those violating them, and everyone else for that matter, could not fail to draw the conclusion that those rights were not valid after all, were not really recognized, at least as rights defined and guaranteed by the criminal law, in any serious manner. The notion ‘taking X seriously’ in this kind of context seems to preclude radical dissimilarity between X and whatever is done by way of responding to it. (Primoratz 1989b: 200).

It seems to be a very widely shared experience of human beings that face-to-face blame does not adequately register the seriousness of moral wrongdoing. Something more demonstrative is required. Of course, it need not take the form of physical penalties such as fines; social penalties such as

\textsuperscript{11} I should also note that Robert Nozick proposes an ER justification that I will not explicitly discuss here, partly due to reasons of space and partly because my criticisms apply mutatis mutandis to his justification. See Nozick 1981; for his statement of the uniqueness claim see 377.
ostracism may sometimes as effectively make their point. Some forms of hard treatment we may find increasingly unacceptable. And it is arguable that a more civilized society will need to make less and less use of hard treatment in responding to moral dereliction. Compensatory and restitutive penalties might be used more frequently. However, I do not think we live in a society for which hard treatment has become redundant; nor do we come near to it. (Kleinig 1991: 417–418)

Both Primoratz and Kleinig think that verbal criticism would not take some forms of offending seriously enough, that it would not be a serious enough response to some offenses. On the basis of this, Primoratz voices support for the uniqueness claim. For some offenses, he claims, “the necessary seriousness and weight can be secured only by punishment” (Primoratz 1989b: 200). Despite his careful qualifications, Kleinig’s remarks suggest a similar conclusion. His qualifications do not rule out commitment to the uniqueness claim as formulated here, i.e., that punishment is necessary for some offenses. Instead, they seem only to suggest that substantially limiting the use of hard treatment would be desirable under the right circumstances. Moreover, his reference to civilized societies seems ambiguous. It is not clear whether he is evaluating civility in terms of conventions that do not call for widespread use of severe hard treatment or in terms of the rate of wrongdoing in society (or both). Either way, he seems to be making a similar argument for a similar conclusion.

This argument’s fate turns on the meaning of the word serious. If the word is meant to indicate something about the degree of criticism expressed, the argument is expressivist. But if the word just indicates the degree of suffering that offenders are made to endure, then the argument is not an ER argument but a regular retributivist argument – expressive considerations are not being used to clarify and support retributivist principles, rather a simple appeal is being made to retributivist principles.

An ER theorist can only make the first argument and that argument fails in light of my conception of punishment and in light of the fact that the use of punishment to express criticism is conventional. Hart is right to wonder how one might

12 For criticism of some of the idiosyncrasies of Primoratz’s justification see Davis 1991.
demonstrate, without simply appealing to retributivist principles, that sufficient criticism could not be expressed verbally (which is not to say that verbal criticism might not be deficient for many other purposes). Suppose for the sake of argument that the case can be made, though. That would not show that punishment is uniquely suited to express the criticism needed for particularly serious offenses. Keeping my conception of punishment in mind and assuming, with the ER theorist, that “human behavior is expressive,” (Hampton 1992b: 1669) there are non-punitive ways to conventionally express significant amounts of criticism without relying solely on verbal criticism.

The standard criminal process highlights many ways that could conventionally be used to express criticism. The degree of investigative effort, prioritization of certain investigations over others and the amount of resources that investigators devote to investigating different offenses can all be used to express varying degrees of criticism. Procedural standards, such as requiring trials for the most serious offenses or requiring psychological evaluations of defendants charged with or convicted of such offenses (as if to say that some offenses are so depraved that the offenders are more likely to be suffering from mental illness or incapacity) can be used to express a greater degree of criticism.\(^\text{13}\)

\(^{13}\) An important objection can be made at this point (which I owe to Jacob Beck). Since criticizing defendants prior to conviction is inappropriate, one might object that such techniques cannot legitimately be used to criticize. This objection can be met by thinking of the criticism expressed prior to conviction as criticism directed, not at a suspect or at a defendant, but at whoever committed the offense. Conviction can then serve as a formal identification, identifying this particular suspect or defendant as the offender towards whom the criticism was being directed and towards whom further criticism can be directed. There is an interesting parallel here with preventive detention. Though the criminal process prior to conviction cannot legitimately be considered to be formally criticizing particular suspects or defendants, it nevertheless stigmatizes them. In this respect it is similar to preventive detention, which, though not formally a kind of punishment, is nevertheless harmful in practically all the ways that punishment is harmful. Just as time spent in preventive detention is often credited towards sentences, and so effectively counts as punishment following conviction, so the stigma to which defendants are subjected prior to conviction can effectively count as criticism for purposes of sentencing.
condemnation from parties to the criminal process, detailed descriptions of the offenses and their character and consequences, commentary on the significance and aims of the laws broken, emphasis of aggravating factors or the lack of mitigating factors and so on.

Crucially, there are many sentences other than verbal criticism that could conventionally be used to express criticism without aiming to impose suffering (and so without punishing). Forcing offenders to pay some court costs and investigative costs and to furnish their victims with compensation could be used to express criticism (on the latter see Sayre-McCord 2001: esp. 18–20). Various incapacitative techniques could also be used to express criticism, e.g., restraining orders and even confinement. The terms and length of probation, which can subject offenders various restrictions as well as surveillance of varying degrees of invasiveness for different lengths of time, could also be used to criticize.

Many of these enforcement techniques and procedures can be applied with the aim of inflicting suffering. They can be (and often are) used as punishments or as means to punishing. But there seems no reason to think that their critical power must rest on whether or not they are applied with the aim of inflicting suffering – that is just one possible convention. They can be applied so that they cause only incidental suffering. Moreover, steps can be taken, within practical limits, to minimize that suffering. Even if applied in this way, however, it looks like they could still be used to express significant criticism given suitable conventions.

To take one example, forcing offenders to compensate their victims, as Geoffrey Sayre-McCord proposes, (though he argues for something a bit broader, which he calls reparations) can convey significant criticism (Sayre-McCord 2001). Forced compensation has no straightforward connection with aiming to impose suffering and so with punishment, yet it is not a purely verbal form of criticism either. It takes concrete steps to remedy damage done by wrongs and does not leave things as offenders have left them – and it does this without aiming to make them suffer.
My rather haphazard list of examples leaves out many alternatives and many details about how they would be applied. But it does show that the small set of options considered by some theorists, i.e., either some form of verbal condemnation or punishment, is not comprehensive. It rests on an overly broad conception of punishment that ignores an essential characteristic of punishment – the aim to impose suffering – and it does insufficient justice to the variety of non-punitive ways that criticism could be expressed. Relying on such a limited set of options makes for a justification of punishment that rests on a false dilemma. Primoratz and Kleinig fall prey to this problem. As I will show, Hampton and Duff also succumb to it.

I will discuss Hampton first and then conclude with Duff. Hampton argues that both wrongdoing and punishment are expressive (Hampton 1988, 1992a, b; cf. Kleinig 1991). She begins with an account of the expressive nature of wrongdoing.

An immoral response to a person, whether or not it produces harm, carries with it a message, in particular, an insulting message. [...] An immoral action is insulting in the sense that it sends a message which challenges the victim’s worth. (Hampton 1992a: 6; see also Hampton 1988: 124, 1992b: 1674)

In committing wrongs, wrongdoers deny the worth of their victims. Of course, wrongs are not punishable just because they express such denial, otherwise verbal denials of worth would be punishable. The way the insult is expressed is key to Hampton’s account (cf. Dolinko 1991). Serious wrongs, in a sense, try to make good on their insulting message – they try to degrade the victim and lower her worth. Hampton does not hold that victims actually lose worth, but she thinks it plausible to say that victims of serious wrongs feel as if they have lost worth. She claims that they suffer diminishment, the appearance of a loss of worth (Hampton 1992a: 6, 1992bb: 1673). To the victim, the wrongdoer and even third parties it may appear as if the victim has lost worth.

Hampton builds her justification on this foundation. Punishing a wrongdoer shows that his message is false.

The retributive punisher uses the infliction of suffering to symbolize the subjugation of the subjugator, the domination of the one who dominated
the victim. And the message carried in this subjugation is ‘What you did to her, she can do to you. So you’re equal.’ The one who acted as if he were the lord of the victim is humbled to show that he isn’t lord after all. In this way, the demeaning message implicit in his action is denied. (Hampton 1992a: 13)

Punishing wrongdoers demonstrates that victims and wrongdoers are equals: victims (or their agents) can do the same sorts of thing to wrongdoers that wrongdoers do to their victims. It denies wrongdoers’ false claims about worth and eliminates the evidence for these claims, i.e., states of affairs where victims can be thought to have “lost” to wrongdoers. Punishment, in effect, annuls wrongs in Hegelian fashion (cf. Primoratz 1989b).

There are a number of worries one could bring up at this point. Hampton’s expressive account of wrongdoing, for example, is questionable. A wrongdoer need not deny his victim’s worth (Matravers 2000: 76–77). Wrongdoers, it seems, are often simply indifferent to worth and the obligations it entails. Such indifference is consistent with acknowledgement of worth. One can even imagine cases where a wrong reflects acknowledgement of worth and related obligations, e.g., a case where these are the wrongdoer’s reasons for committing the wrong. One might also worry that Hampton’s use of the word can in the statement “what you did to her, she can do to you” is ambiguous in a way that renders the statement either a claim about capabilities that is irrelevant to her argument (Hampton 1992: 1692) or a question begging claim about what is permissible. Neither makes for a plausible justification. Another worry – the familiar one voiced by Hart and others – is aptly put by David Dolinko: “it is surely not true that whatever would correct (or “nullify”) a mistaken moral claim is ipso facto morally permissible” (Dolinko 1991: 552). Even if Hampton’s expressive account of wrongdoing and punishment is correct, one might still have doubts about the justifiability of punishment.

I will leave the first two worries aside. With respect to the second worry, I suspect that Hampton could simply restate her argument in different terms to avoid it. The first worry is important, but there are clearly cases of wrongdoing that deny worth. If there are no difficulties with the rest of Hampton’s justification, she may have a strong argument for punishment in
such cases. I will also set aside the last worry – at least in the
vague form in which many theorists state it. If the uniqueness
claim is true, it can be deployed against this worry, forcing the
skeptic to confront a difficult choice. Rather than simply
voicing doubt that punishment would ever be the correct
choice, I will show how Hampton’s argument for the unique-
ness claim fails.

Hampton’s argument for the uniqueness claim fits the stan-
dard form. She offers some examples of wrongdoing – rape and
a particularly brutal hate crime – and considers the accept-
ability of not punishing such wrongs.

[A] decision not to punish wrongdoers such as the rapist is also expressive: it
communicates to the victim and to the wider society the idea that such
treatment, and the status it attributes to the victim, are appropriate, and
thus, in the case of the rape victim, reinforces the idea that women are
objects to be possessed and are “there for the taking.” Moreover, if nothing
happens to people like the [perpetrator of the hate crime] (and indeed,
nothing did), we feel a special kind of fury at the thought of what they “got
away with.” (Hampton 1992b: 1684–1685, emphasis added)

Hampton considers two possibilities: not punishing and doing
nothing. She may be on the verge of conflating the two here,
perhaps thinking that to do the former is effectively to do the
latter. Soon after, she considers responding to such offenses
verbally.

Re-establishment of the acknowledgment of the victim’s worth is normally
not accomplished by the mere verbal or written assertion of the equality of
wrongdoer and victim. For a judge or jury merely to announce [guilt and
affirm equality], after reviewing the facts of the [hate crime] ... is to
accomplish virtually nothing. [...] This representation of degradation re-
quires more than just a few idle remarks to deny. (Hampton 1992b: 1686)

Hampton uses the apparent unacceptability of these alterna-
tives to press the case for the uniqueness claim. But the argu-
ment is no more compelling when considered in light of her
theory of worth and her sophisticated expressive understanding
of wrongdoing and punishment than it was before.

The unacceptability of doing nothing or of relying only on
verbal condemnation does little to support the uniqueness claim
in light of my conception of punishment and in light of the
conventionality of the use of punishment to express criticism.
Doing nothing or expressing condemnation verbally do not exhaust the non-punitive possibilities. Many non-punitive techniques, some of which I listed earlier, can, with the right conventions in place, be used to express the criticism, beliefs, attitudes and so on that are needed to affirm equal worth on Hampton’s view. My objection against Primoratz’s and Kleinig’s arguments for the uniqueness claim also undermines Hampton’s argument for it.

Now I will move on to Duff. Of the justifications canvassed here, Duff’s is perhaps the most formidable. This formidability is belied by his apparent use of the standard argument for the uniqueness claim. We should use hard treatment punishments of certain kinds because they can serve the communicative aims of punishment more adequately than can mere convictions or symbolic punishments. (Duff 2001: 82; see also 95)

Where Duff differs is in his conception of the communicative aims of punishment. Punishment, on Duff’s view, should be used to express criticism for the purposes of correction and persuasion, more specifically at getting offenders to repent their wrongdoing, to reform themselves and to reconcile themselves with their victims (Duff 2001: 106–112). Duff’s justification portrays punishment as guided by these aims and it is intended to show why these aims necessitate punishment (Duff 2001: 92).

Unlike many advocates of the standard argument for the uniqueness claim, however, Duff considers substantive alternatives to punishment beyond mere verbal criticism – or at least he considers techniques thought by many to be non-punitive and argues that they can actually be seen as paradigmatic punishments, given his proposed aims. He considers, among other things, offender-victim mediation programs and community service orders. He argues that they qualify as punishments and that they can in fact be seen as paradigmatic punishments in light of the aims he proposes. He uses such examples, in conjunction with the insufficiency of mere verbal criticism, to press the case for punishment.

I will not challenge Duff’s proposed aims, at least not substantively; I will suggest that his formulation of the first aim is misleading and should perhaps be restated. One might worry
about the propriety of these aims or, if one thinks them appropriate, one might wonder whether they are important enough to justify punishment (on the former see Matravers 2000: 90–92; on the latter see Golash 2005: 117–146 and Narayan 1993: 178). But I will set such worries aside. Even setting them aside, Duff’s justification still has a serious problem. He overestimates the degree to which his proposed aims favor punishment because he mistakenly thinks that his proposed aims necessitate aiming to impose suffering on offenders. They do not.

Consider Duff’s discussion of offender-victim mediation – a process in which an offender is confronted by her victims, forced to face the consequences of her offending and encouraged to furnish reparations to her victims – and his explanation of why it qualifies as punishment (specifically, retributive punishment).

[C]riminal mediation is retributive in that it seeks to impose on (to induce in) the offender the suffering she deserves for her crime and is justified in those terms. She deserves to suffer censure for what she has done. Mediation aims to communicate that censure to her in such a way that she will come to understand why, as well as that, she deserves it. She deserves to suffer remorse for what she has done. Mediation aims to induce such suffering in her by bringing her to recognize the wrong she has done. She ought to make apologetic (and thus necessarily burdensome) reparation for that crime to its victim. Mediation aims to provide such reparation. By seeing criminal mediation as punishment, we can thus make clear and plausible sense of the retributivist idea that the guilty deserve to suffer, by showing what they deserve to suffer and why. (Duff 2001: 97)

My concern here is with Duff’s focus on suffering, with his claim that mediation, implemented in service to his proposed aims, aims to induce suffering and thereby punishes. Duff mistakenly thinks that the aim to impose suffering must guide such enforcement techniques and he is too quick to conclude that the suffering that is involved in those techniques renders them punitive.

Consider Duff’s proposed aim to get offenders to repent their offenses. He claims that offenders deserve censure and that when we censure offenders we want them to accept the censure as justified. We want them, in short, to recognize the
fact that they have done wrong. Duff claims both that “an authentic recognition that I did wrong must bring with it repentance of that wrong” and that repentance “is necessarily painful, since it must pain me to recognize and admit (to myself and to others) the wrong I have done” (Duff 2001: 107). In aiming to get offenders to repent their offenses, Duff claims, we thereby aim to make them suffer a particular kind of suffering that they deserve, specifically the pain of repentance and remorse.

There is an important aspect of repentance and remorse whose significance Duff apparently overlooks, however. There is nothing absurd in the idea of an offender who remorselessly and unrepentantly recognizes that he has done wrong. Contra Duff, it need not pain one to recognize that one has done wrong. More is involved in repentance and remorse than that. A crucial element seems to be the adoption of a certain sort of negative attitude towards the wrong. Having this attitude may, perhaps necessarily, make one suffer. Focusing on the suffering that such an attitude causes easily distracts from what we do when we criticize someone, however.

When we criticize someone for acting wrongly, we are trying to get him to recognize the fact that he has done wrong and we are trying to get him to adopt a negative attitude towards the wrong. Suffering that results from the adoption of such an attitude may be desirable, not qua suffering, but as evidence of the attitude. Criticizing a wrongdoer with the expectation or in the hope that he will undergo such suffering therefore seems rather different from aiming to make him suffer. The suffering that attends such attitudes need not be thought desirable either in itself or as a means to something else, but rather as a reflection of certain attitudes – its origins are important, not its intrinsic character as suffering (if the latter were the case, we could just as well inflict suffering in other ways). We can aim at getting wrongdoers to adopt the attitudes that cause this suffering without aiming to make them suffer, even if the adoption of those attitudes causes suffering. In fact, we can and often do try to minimize the suffering that accompanies these
attitudes, e.g., by comforting repentant wrongdoers (think of children or close friends). 14

Given my conception of punishment, one is not necessarily punishing someone even if one criticizes her and aims to get her to adopt these sorts of attitudes, for in aiming to do so one need not aim to make her suffer. This is simply a conceptual point about the nature of punishment and about the nature of our aims. Duff, I think, underestimates the psychological complexities attending criticism and punishment. Given these considerations, it is questionable whether he should state his first aim in terms of generating remorse and repentance – which we might crudely think of as complex states comprised at least in part of negative attitudes towards one’s wrongdoing and painful psychological states accompanying those attitudes – or simply in terms of generating certain negative attitudes towards one’s wrongdoing. The latter are what seem especially important in the context of mediation and with respect to both Duff’s communicative concerns and his other aims. These attitudes have no straightforward connection with one of punishment’s essential elements, however: the aim to impose suffering. It is partially because of Duff’s underestimation of the psychological complexities involved (and perhaps also because he is independently motivated by retributive commitments) that he concludes otherwise.

These considerations undermine Duff’s argument that mediation is a paradigmatic form of punishment, even given his aims, but they also help to undermine the argument he offers for the “further kinds of penal hard treatment involved” in

14 None of this is meant as a denial of the claim that offenders deserve to suffer remorse, but rather as a denial of the claim that their deserving that suffering speaks in favor of the aim to impose it or other sorts of suffering on them. The claim that offenders deserve to suffer remorse is compatible with the view that punishment is unjustified. Offenders may deserve to suffer remorse in the sense that it is permissible to act in ways that will result in their suffering such remorse if one does so for the right reasons, e.g., criticizing them and trying to get them both to recognize the wrongfulness of their offenses and to adopt a negative attitude towards them. Retributivism does not have a monopoly on desert claims and we must take care not to conflate the truth of some interpretations of some desert claims with the truth of retributivism.
enforcement, e.g., community service orders and reparation orders (Duff 2001: 107). Duff argues that fallible human beings sometimes need to be responded to in particularly emphatic ways in order to help them “face up to and understand” their wrongs (Duff 2001: 108).

This then is one purpose of penal hard treatment, such as the reparative burden or the community service that might be imposed on an offender. It is a way of trying to focus his attention on his crime. It provides a structure within which, we hope, he will be able to think about the nature and implications of his crime, face up to it more adequately than he might otherwise (being human) do, and so arrive at a more authentic repentance. As fallible moral agents, we need such penances to assist and deepen repentance. (Duff 2001: 108)

Duff relies on the standard argument for the uniqueness claim, with a new twist: merely verbal or “purely symbolic” expressions of censure do not suffice to get offenders to understand, and look negatively on, their offenses (Duff 2001: 107). Penal hard treatment is needed. Even granting Duff the need for something more than mere censure, however, say the need for a structure that can help focus one’s attention on one’s wrongdoing so that one can adopt a negative attitude towards it, he has not made the case for punishment. Punishment does not seem uniquely capable of offering such a structure (or even more capable; see Narayan 1993: 176–177). A variety of enforcement techniques applied without the aim to impose suffering seem capable of doing this as well. One might think punishment uniquely capable if one thinks, as Duff does, that one cannot impose the proposed alternatives without aiming to impose suffering. But, as I think I have shown, no such aim need be involved.

Before concluding, one last aspect of Duff’s justification should be addressed. The third aim of punishment that he proposes is reconciliation. Duff argues that we should, in punishing wrongdoers, seek to reconcile them with those they have wronged. In cases of wrongdoing, an apology is required. Certain punishments such as community service orders and reparative burdens are, in effect, apologies for what wrongdoers have done: in undergoing them, offenders effectively apologize to and reconcile themselves with those they have
wronged. We should impose such punishments, Duff argues, in the hope that wrongdoers will come both to repent their wrongs and to accept the punishments as means of expressing the apologies they owe others. Again, Duff argues that merely verbal apologies are not enough in cases of serious wrongdoing. What is needed is an especially “forceful and weighty kind of apology” the likes of which punishment can provide (Duff 2001: 109; see also 95). The perpetrators of serious wrongs, he argues, must feel remorse and undertake burdensome tasks in order to offer adequate apologies – hence, the need for punishment, which can furnish the sorts of burdens needed for adequate apologies in such cases.

Duff’s argument here seems to suffer from the same difficulty as his arguments regarding repentance. In the case of repentance, Duff overemphasizes the suffering of repentance. He consequently underestimates the significance of the negative attitudes towards one’s own wrongdoing that give rise to the suffering and that show precisely how and why the suffering is important. This leads him to mistakenly conclude that we must aim to impose suffering on wrongdoers. Similarly, Duff thinks that the need for adequate apologies sometimes requires punishment, i.e., he concludes that we should aim to impose burdens qua burdens on some wrongdoers.

But what seems important here is not the burdensomeness of an adequate apology *per se*, i.e., the suffering it causes, but rather the attitudes expressed by it, e.g., a negative attitude towards the wrong as wrong and a desire to make amends. The importance of reparative projects and other weighty kinds of apologies seem to lie, not in their burdensomeness, but in at least two things: in the attitudes they reveal and in their contribution to repairing the damage done. The willingness to make amends and to offer a sincere, sufficiently weighty apology *in spite of* the burdensomeness of doing so is what seems important, not the burdensomeness in itself – otherwise less elaborate burdens would suffice. Burdensomeness, like the suffering that results from one’s negative attitudes towards one’s wrongdoing, seems important because it often serves as evidence of certain attitudes, not because of its intrinsic char-
acter (which seems intrinsically no different from the burdensomeness of all sorts of other activities).

It seems, then, that Duff’s aim to reconcile wrongdoers with those they have wronged does not speak in favor of aiming to make wrongdoers suffer and so does not call for punishment. If one aims at reconciliation, forcing some wrongdoers to make amends, say by means of compensation, might seem appropriate. Such compensation may be burdensome and may offer wrongdoers an opportunity to apologize, but that does not mean that, in forcing wrongdoers to furnish compensation, one thereby aims to burden them or make them suffer. Rather, one may aim to force them to fulfill an obligation incurred, i.e., to repair, at least partially, the damage done. A remorseful wrongdoer can, in spite of the burdensomeness of doing so, choose to furnish compensation in order to make up for his wrong and can thereby apologize. The burdensomeness of doing so need have no communicative value in itself, however, and need not be deliberately aimed at and varied in order to render it adequate for this purpose. Any convention that makes this the case stands in need of justification. Such a justification, it seems, eludes ER.

VI. CONCLUSION

All three of the claims I have discussed are problematic. In its conceptual variant, the claim that wrongdoers must be criticized seems false and in its normative variant it does not, by itself, favor punishment over other ways of expressing criticism. The claim that punishment expresses criticism is usually misleading in the context of ER arguments. The use of punishment to do this is conventional and there are other ways of conventionally expressing criticism. And the claim that only punishment can express the degree of criticism called for by some wrongs is false. This uniqueness claim is ER’s weakest point. Many theorists, with a diverse background of theoretical commitments, think that something like a uniqueness or necessity claim is needed to meet the justificatory challenge posed by punishment. It is difficult to see how ER can meet that challenge, however, given the considerations I have offered and
given the failure of the arguments canvassed here. There is little reason to think that the need for adequate criticism requires punishment, in the sense of a practice that aims at imposing suffering, given the possibilities for expressing criticism in non-punitive ways. The wealth of alternatives to punishment (beyond the merely verbal) that can match punishment’s expressive power without aiming to make offenders suffer undermines the uniqueness claim. If punishment is a justifiable practice, its justification is not to be found in ER.

ACKNOWLEDGEMENTS

Many thanks to Kenneth Baynes, Ishani Maitra and three anonymous referees for *Law and Philosophy* for helpful comments on earlier versions of this paper. Thanks also to participants at the 15th annual Harvard-MIT Graduate Philosophy Conference, where I presented a shorter version of the paper, and to my commentator Japa Pallikkathayil for helpful comments and questions.

REFERENCES


NATHAN HANNA


Department of Philosophy,
Syracuse University, 541 Hall of Languages,
Syracuse, NY, 13244-1170, USA
E-mail: nthanna@syr.edu