Abstract: Saul Smilansky holds that there is a widespread intuition to the
effect that pre-punishment – the practice of punishing individuals for crimes
which they have not committed, but which we are in a position to know that
they are going to commit – is morally objectionable. Smilansky has argued that
this intuition can be explained by our recognition of the importance of respect-
ing the autonomy of potential criminals. (Smilansky, 1994) More recently he
has suggested that this account of the intuition only vindicates it if determinism
is false, and argues that this presents a problem for compatibilists, who, he
says, are committed to thinking that the truth of determinism makes no moral
difference (Smilansky, 2007).

In this paper I argue that the intuitions Smilansky refers to can be explained
and vindicated as consequences of the truth of a communicative conception
of punishment. Since the viability of the communicative conception does not
depend on the falsity of determinism, our intuitions about pre-punishment do
not clash with (what Smilansky calls) compatibilism. And if the communica-
tive theory of punishment is – as Duff (2001) suggests – a form of retributivism,
the account also meets New’s (1992) challenge to retributivists to explain what
is wrong with pre-punishment.

1. Introduction

In the world as it is, punishment for a crime happens, if it happens at all,
after the crime has been committed. It seems possible to imagine worlds
different from our own, in which this is not the case. Lewis Carroll (1865)
and Philip K. Dick (2002) have both produced fiction in which individuals
are punished before they have committed a crime, either on a single
occasion, or as a matter of course. Such practices could only exist, of course, in situations in which we had significant fore-knowledge of how individuals would act.

The apparent imaginability of such scenarios gives rise to a philosophical question. Does our insistence that punishment should come after, rather than before a crime simply reflect the fact that we normally do not know enough about the future actions of potential criminals to be able to say with any certainty who will commit a crime and who will not? Or would we have good reason to insist that punishment should only come after the commission of a crime even in the absence of such practical limitations?

Saul Smilansky (Smilansky, 1994, 2007) holds that there is a widespread intuition to the effect that pre-punishment – in other words, the practice of punishing individuals for crimes which they have not committed, but which we are in a position to know that they will commit – is morally objectionable. Smilansky has argued that this intuition can be explained by our recognition of the importance of respecting the autonomy of potential criminals. (Smilansky, 1994) More recently Smilansky has suggested that this account of the intuition only vindicates it if determinism is false. If determinism is true, then the intuition should be disregarded (Smilansky, 2007). Hence, says Smilansky, compatibilism is false, for compatibilism is – at a first approximation at least – the view that the truth of determinism makes no difference, morally speaking.

While I agree with Smilansky that pre-punishment is objectionable for moral reasons and not simply practical ones and that there is a widely shared intuition to this effect, his explanation of this intuition strikes me as unsatisfactory. This is at least in part because I accept his claim that his account of the intuition raises problems for (some sorts of) compatibilists. Unlike him, I take this to count against his account of our intuition, rather than against compatibilism. My aim in this paper is to put forward an account of what is wrong with pre-punishment that is acceptable to compatibilists and incompatibilists alike.

My suggestion is that the intuitions are best accounted for by a communicative conception of punishment. On a communicative conception of punishment the harsh treatment that punishment involves is justified insofar as it is a response to certain actions on the part of the punished individual. In what follows I shall argue that the viability of a communicative account of punishment does not depend on the falsity of determinism. If I am correct, the vindication of our intuitions about pre-punishment is independent of the truth of determinism, and thus their existence presents no difficulty for what Smilansky calls compatibilism.

If my proposal is correct, then it is interesting for two further reasons. One is that it appears to refute a claim made by Christopher New in an early discussion of pre-punishment (New, 1992). New’s paper, which
introduced the notion of pre-punishment to the recent philosophical literature, can be read as offering a challenge to theorists of punishment to account for the wrongness of pre-punishment. New claims that the challenge cannot be met by either retributivist- or deterrence-based theorists of punishment. Other than Smilansky (1994), few theorists of punishment have responded to this challenge. If I am right, communicative theorists of punishment can meet it and if, as Anthony Duff (Duff, 2001) has suggested, communicative theories of the sort that I shall discuss here are retributivist theories, then New turns out to be mistaken in his claim that retributivists cannot rule out pre-punishment.

My proposal also has implications for how we should think about the pre-punishment of institutions, such as states and corporations – an issue first raised by New (1995). It is far less clear that we have an intuition that it is wrong to pre-punish such entities than that we have an intuition about the wrongness of pre-punishing individuals. Nevertheless there seems something puzzling about the practice. I suggest that pre-punishment of institutions can be morally acceptable, not because of a difference between the metaphysical natures of institutions and natural persons (although such differences do exist), but because the fact that there is no need to treat artificial persons as ends in themselves means that it is morally acceptable to inflict harsh treatment on them not as a means of communicating with them, but as a means of communicating about them. Insofar as there is a puzzle about the pre-punishment of institutions, it is because this fact establishes an important disanalogy between what would normally be involved in punishment, and what would be involved in the pre-punishment of institutions.

2. Our intuitions about pre-punishment

As I noted above, Smilansky holds that there is a widespread intuition to the effect that pre-punishment is morally wrong. I agree. But Helen Beebee (2008) has recently questioned how widespread this intuition is, citing the fact that in the UK ‘conspiracy to murder is a criminal offence as is... engaging in “any conduct in preparation for giving effect to” an intention to commit an act of terrorism’ (p. 259) and suggesting that support for such laws is indicative of support for pre-punishment of (some) would be murderers and terrorists. Smilansky’s rejoinder is that individuals punished under these statutes are not being pre-punished for murder or terrorism, but are being (standardly) punished for the distinct crimes of conspiring to murder and attempting to commit a terrorist act (Smilansky, 2008, pp. 262–3). He cites in support of this view the fact that individuals convicted of such offences typically receive different – and indeed
lighter – punishments than individuals who succeed in murdering or com-
mitting terrorist acts⁶ (Smilansky, 2008, p. 263).

While I agree with Smilansky’s ‘punishment for distinct offences’
defense against this objection, I think his argument in favor of it is incon-
clusive. For one might hold that the difference in sentences (which, if the
argument is to work, must be accompanied by an intuition to the effect
that such differential sentencing is morally appropriate – something which
he fails to demonstrate) can be explained by a difference in our epistemic
situation with respect to the crime. In the actual world, which both Beebee
and Smilansky are discussing here, we are typically in a position to be
more certain of the fact that certain individuals have committed a crime
than of the fact that they are going to. This, it might be thought, accounts
for actual sentencing policy, and our intuitions about it. If the world were
different in this respect, our intuitions might also be.

Fortunately, there is a more persuasive argument for the ‘distinct
offences’ line. It derives from our practices and intuitions about failed
conspiracies and preparatory acts. If our intuitions about the justification
for punishing such failures were based on an intuition as to the justifiability
of pre-punishing actual crimes, then it would presumably be inappropriate
to punish bungled attempts or conspiracies. For in such cases, there is ex
hypothesi, no crime distinct from the attempt or (in the case of conspiracies)
preparatory acts that one might be engaged in pre-punishing.⁷ But I doubt
that anyone feels that it is illegitimate to punish bungled conspiracies, as
they would presumably do if Beebee were right that ‘the common-sense view . . . is that such legislation is legitimate because it prevents a
crime . . . from being committed’ (Beebee, 2008, p. 260)

3. The communicative view

I’ve argued that Smilanksy is right and Beebee is wrong on the question
of whether there is a widespread intuition about the wrongness of
pre-punishment, although I’ve taken issue with Smilansky’s response to
Beebee. Now I need to say something about how these intuitions may be
vindicated on a communicative view of punishment. Before doing so, I
shall say something about what the communicative view is.

Following Duff (2001) I take a communicative view of punishment to be
one on which the moral justification of the harsh treatment which punish-
ment involves depends on the fact that this harsh treatment is a response
to an offense aimed at communicating a message to the offender. Typi-
cally, the communication is regarded as being expressed by a society or its
duly appointed representatives via the harsh treatment; the content of the
communication concerns such matters as the wrongness of the offense; the
fact that the society regards its wrongness as a serious matter; and so on.
The communication may be seen as aimed at inducing remorse on the part of the offender and encouraging his re-integration into the moral community.

Communicative views of punishment can be distinguished from what are sometimes characterized as ‘denunciatory’ views. Advocates of denunciatory views agree with their communicative counterparts that the justification of punishment depends on the fact that it conveys a message; but disagree with them as to who the intended recipient of the message is. Unlike communicative theorists, denunciatory theorists hold that the message is expressed to the world at large, rather than the offender in particular.

Communicative and denunciatory views of punishment can be seen as belonging to a single family of theories of punishment. Call this the ‘expressivist’ family of theories. Advocates of theories from the expressivist family of theories may, but need not, hold that there is a conceptual connection between punishment and expression. (Duff seems to waver on this point.) The suggestion that there must be an expressive element to anything correctly called punishment seems to be traceable back to Feinberg (1970), who argues that in the absence of an expressive element of this sort we are unable to distinguish between regarding a particular kind of harsh treatment as a punishment and regarding it as a tax on undesirable behavior.

4. How the communicative view vindicates our intuitions and saves compatibilism

Those who accept a communicative view of punishment have good reasons to accept our intuitions about the wrongness of pre-punishment. For they are in a position to say that we cannot punish someone for a crime that they have not committed, simply because we are not warranted in responding to someone who has not yet committed a crime in the same way that we would respond to someone who has committed a crime.

It is worth distinguishing three ways in which the communicative theorist might develop this point. The first is to say that pre-punishment cannot be regarded as a response to an offense since at the time that the pre-punishment is imposed, no offense has been committed. The second is to say that there are important differences between the ways in which we are warranted in responding to communicating to those who have actually offended and the ways in which we are warranted in communicating with those who are going to offend. The third is to point out that if the harsh treatment which constitutes punishment is justified as being communication which is aimed at bringing about regret or remorse for an offense,
then any attempt at justifying harsh treatment which constitutes pre-punishment along the same lines must fail, since it is logically impossible to feel remorse for something that one has not done.

A good account of our intuitions about pre-punishment needs to incorporate all three points. If we simply focus on the idea that punishment involves communication about an offense, someone who wants to defend the consistency of pre-punishment and the communicative conception can simply point out that one can communicate – and indeed express disapproval of – events which have not yet occurred. By the same token, one cannot lean too heavily on the idea that punishment is a response to a crime in arguing that the communicative theory vindicates the intuition that pre-punishment is unacceptable. It is true that on some understandings of the word ‘response’, one cannot respond to something which has not yet occurred. But to insist on understanding the word ‘response’ in this way is apt to seem unduly stipulative. However, once we have incorporated the tie between communication and remorse or regret into our theory, the case against pre-punishment seems fairly secure: one cannot regret what has not yet occurred.

In response, one might suggest that pre-punishment should be regarded as aiming at evoking an emotion analogous to regret or remorse but differing from it in being directed at the future rather than the past. I doubt that such an emotion exists. But even if it did, it would be of no help in defending the acceptability of pre-punishment along lines similar to the communicative view. An institution of pre-punishment, conceived of as involving communication aimed at inducing a future directed analogue of regret would simply be perverse. The value of inducing regret or remorse is connected with the fact that there is a logical tie between experiencing these emotions and being motivated to make amends where this appropriate. But the best way to make amends for an offence that one has not yet committed is simply not to commit it. So the putative institution of pre-punishment would be one that aimed – in cases where it was successful – at inflicting harsh treatment on non-offenders. It would only inflict harsh treatment on offenders in cases where it failed of its aim. (By contrast, our ordinary institution of punishment often inflicts harsh treatment on non-offenders and is unsuccessful in bringing about remorse. But it is not a condition of its success that it inflict harsh treatment on the innocent.)

One might see a relation between this line of argument and one suggested by Kearns (2008). He argues that we can see that pre-punishment is wrong because in cases where it did not prevent crimes, we cannot be sure that it did not play a role in causing them. This would indeed be perverse; but it is a slightly different perversity from the one I am concerned with: I hold that pre-punishment that did not prevent crimes would have failed to do the thing that we typically take to justify punishment – namely produce...
remorse, and that punishment which succeeded in doing this would punishing the innocent.9

My claim that pre-punishment could not produce an emotion analogous to regret or remorse for one’s past acts, and that the institution of pre-punishment could not be justified along the same lines as our actual institution of punishment might be challenged in various ways. For example Adam Morton has suggested that one might regret being the sort of person who was going to act in a particular way.10 This is correct, but of no help to the defender of pre-punishment. An institution which aimed to induce remorse in someone for the kind of person they were would be analogous to one which punished individuals for being individuals of a particular kind – in other words, that punished them for their character. But typically we think that punishment should be directed at individuals for their acts, rather than their characters. (Totalitarian states sometimes see things differently.)

Similarly, Ulrich Steinvorth has pointed out that one might feel remorse for intending to act in a particular way.11 Again, this is true – but an institution which aimed to induce remorse for future directed intentions can be seen as ordinary punishment for an intention, not pre-punishment for the intended act. Our actual practice of punishing attempts might be interpreted along similar lines.

These suggestions involve finding a past-directed emotion of remorse or regret, directed at something other than a past act, and suggesting that pre-punishment should be seen as aimed at evoking that emotion. They fail because an institution aimed at inducing a response of this sort can simply be understood as ordinary punishment of the object of the past-directed emotion. This response would not be available if one thought that pre-punishment could be aimed at inducing thankfulness that one had been prevented for causing a crime.12 It seems unlikely that harsh treatment would be successful in inducing such a response; but even if it were, pre-punishment so conceived would have to be pre-emptive and thus end up punishing non-criminal individuals. So it would not be surprising that even if conceived of along these lines, pre-punishment would strike us as wrong.

5. Is pre-punishment really wrong?

Some philosophers might hold, as Christopher New (1992) does – that there is, in fact, nothing wrong with pre-punishment. If they are right, there must be something wrong with what I have said: I have vindicated an intuition that cannot (hence should not) be vindicated. New presents an apparently appealing case of justified pre-punishment: an individual proposes to break the speed limit in an unpopulated wilderness area, informs
the authorities of his intention, and pays the fine that would be imposed on him before breaking the speed limit. This, New argues, is a case of morally acceptable pre-punishment.

I disagree. To see why, we need to make a distinction between two things one might want to say about this case. The first is that the behavior of the authorities here is morally acceptable. The second is that what we have here is something which is a morally justified instance of pre-punishment. I claim that the first is true, but not the second. What we have here is not pre-punishment. As New introduces the notion, pre-punishment is a form of punishment. I claim that what we have here is not a form of punishment at all.

New might object that I have simply imposed a stipulation as to what counts as punishment. I have not. Instead I have said that New’s case only gives us something which is appropriately called pre-punishment if we have here an instance of harsh treatment which can be justified in a way that is analogous to the way in the harsh treatment which constitutes punishment is justified. And for reasons which I have already given, we should deny that this is the case: what we have here cannot be justified as a communicative response to an offense aimed at inducing remorse or regret for that offense.

But if the harsh treatment here cannot be justified along the same lines as it would normally be justified, how can it be justified? The communicative theorist might appeal to the point I drew from Feinberg (1970) at this point and suggest that those who feel we do have justified harsh treatment in a case such as this are thinking of the harsh treatment not as punishment but as a tax on undesirable behavior. One point that might support this way of seeing things is that the acceptability of the harsh treatment in this case seems to depend very heavily on two aspects of New’s example: first, that we have an example of a crime without obvious victims, and secondly that the harsh treatment is simply a fine, rather than involving an obvious loss of personal liberty. Once we change these features of the example it is much less clear that the case involves anything morally acceptable at all.

6. Pre-punishment and artificial persons

As I have already mentioned, New (1995) raises the issue of pre-punishing ‘artificial persons’ such as corporations. It is not clear to me whether Smilansky holds – or needs to hold – that our intuition about pre-punishment extends, or should extend to cases such as this. Nevertheless it is clear that his explanation of our intuitions about the pre-punishment of individuals could not easily be extended to cover such a case. His explanation relies heavily on the idea that we have an obligation to respect the
autonomy of individual moral agents. Few of us think that there is any parallel obligation to respect the autonomy of non-natural persons such as the Enron Corporation.

One might think that a communicative account of punishment would have particular difficulties with the idea of punishing artificial persons. For although artificial persons can be said to act, it is less obvious that they can be said to experience emotions such as remorse. For they are certainly not subjects of experience; and one might well feel that it is essential to emotions that they include an experiential component.

One response to this problem would be to suggest that a communicative theorist should think of punishment as being aimed at inducing remorse in the individuals that act for a corporation. This response is unsatisfactory: it overlooks the fact that artificial persons can continue in existence while their agents change. Some people might hold that there is a logical bar to an individual’s feeling remorse for acts committed by a collective body before they were in a position to act on behalf of it. I do not think there is any such logical bar: whether or not it is psychologically possible in any given case will depend on the make-up of the individual and in particular how strongly they identify with the collective body of which they are a part. However, I think that if the punishment of artificial persons were aimed at inducing remorse in such individuals it would be both unjust (insofar as it involved inflicting suffering on individuals who had not done anything to deserve it) and also unlikely to serve a useful function.

Instead, I suggest that the communicative theorist would do better to think along lines suggested by Margaret Gilbert (Gilbert, 2001). She suggests that although collective bodies are not centers of experience in their own right, they can be thought of as being subjects of behavioral analogues of emotion. Earlier I argued that there are logical ties between sincere regret and the desire to make amends where appropriate. There are also ties between regret and its expression. Gilbert’s suggestion exploits these logical ties: she suggests that it makes sense to think of collective bodies aiming at certain things and acting in ways that would contribute to achieving them. If she is right about that, then we can think of collectives being subject to analogues of regret and remorse and we can see punishment as being aimed at inducing these collective attitudes.

If this account of the punishment of collective bodies is correct, then it seems plausible to think that pre-punishment of artificial persons is ruled out for reasons which are analogous to the reasons for ruling out the pre-punishment of individuals. But it is worth noticing that there are other things one might say to support the intuition that some cases of things that look like pre-punishment of artificial persons are acceptable. One fairly obvious point is that preemptive taxation of undesirable activities by artificial persons is not ruled out by this line of argument, and – as I have
argued above – some things that may look like cases of justifiable pre-punishment might be admissible as cases of justifiable taxation.

A second point is more subtle. Recall the distinction I made in Section 3 between denunciatory and communicative theories of punishment. One reason for preferring communicative theories of punishment to denunciatory theories is that the former seem more obviously to respect the autonomy and agency of offenders than the latter: advocates of denunciatory theories run the risk of seeing in the bad behavior of offenders a useful opportunity, and in their lives and bodies a convenient medium for the expression of societally valuable truths. But there is no parallel reason for concern about the possibility that the harsh treatment of artificial persons for denunciatory purposes might involve treating them as mere a means.

7. Is the communicative account distinct from Smilansky’s?

I have proposed an account of our intuitions about pre-punishment which vindicates them and which I take to be distinct from Smilansky’s and superior to it. One way in which it appears distinct is that it does not seem to require a commitment to the falsity of determinism. So it is an account which someone who holds that the truth of determinism makes no moral difference ought to be comfortable with. To that extent it seems both distinct from Smilansky’s autonomy-based account of those intuitions and superior to them – at least from the point of view of those who are committed to (what Smilanksy calls) compatibilism.

But some might wonder how far the account I have offered is from Smilansky’s own account of the wrongness of pre-punishment. Smilansky holds that it is wrong to punish the (as yet) innocent because to do so fails to respect their autonomy and moral agency. But as many readers will be aware, Anthony Duff (2001) has argued that one significant point in favor of his communicative account of punishment is precisely that – unlike consequentialist deterrent based accounts – it gives an account of punishment which does not involve treating offenders as a means – in particular as a means to the deterrence of other offenders. Given the prominent role that both authors give to Kantian-style considerations of respect for autonomy and agency one might wonder how far apart they really are.

For this reason much depends here on whether the communicative account really can resist being undermined by determinism. I think it can precisely because of the way in which it builds on an asymmetry between human beings emotional responses to the present and the past. But Smilansky might object that if determinism is true there is something irrational about the asymmetry in responses. In discussing his own view about punishing the innocent, Smilansky says that if determinism is true, there is
no deep difference between the individual who is going to commit a crime and the individual who has already done so.

To this, there seems a natural response: the deep distinction is precisely that one individual has committed a crime and the other has not. What we have done is constitutive of who we currently are in a way which what we will do, but have not yet done is not. This goes some way towards explaining why it is appropriate to have different emotional responses towards an individual’s past and the future acts. This is true both of our own acts and of the acts of others. Even if I know someone is destined to be a great pianist, I cannot admire him in the same way as someone who is a great pianist, though I might be in awe of him. Equally, I cannot (and hence should not) feel remorse for my future actions in the same way as for my past actions. To say that this asymmetry was irrational would be to say either that it is something which ought not to be a part of our mental make-up; or perhaps to say that it is some limitation which we can idealize away from. I do not think either is true: the asymmetry is one that is deeply embedded in our nature, and which we cannot idealize away from without idealizing away from our nature as agents who can make plans for the future but not for the past.

The thought that there is no deep distinction, on a determinist view, between an individual who has committed a crime and an individual who will do so seems to be based on the thought that in the long run the pre-punished individual will commit the crime for which he is due to be punished. Quite so; and in the long run we are all dead. But it would still be inappropriate to schedule my wife’s funeral for tomorrow.

Department of Philosophy
Bilkent University, Ankara, Turkey

NOTES

1 I say ‘seems possible’ rather than ‘is possible’, since some might take it there is a conceptual requirement on punishment that it come after the crime. As will be seen in Section 4, I am sympathetic to this point.

2 For the purposes of this paper, I shall accept Smilansky’s characterization of compatibilism, though I suspect many might quibble with it.

3 As Sandrine Berges has pointed out to me, New is not the first philosophically inclined writer to discuss pre-punishment: see here Carroll, 1865, ch. 11

4 As Beebee notes this doesn’t quite fit New’s original (1992) conception of pre-punishment, since that conception was designed to preclude the possibility of pre-punishment preventing the corresponding crime. But both Smilansky and Beebee seem to feel that this point is tangential to their disagreement at this point, and I shall follow them in assuming that it is.

5 This point is anticipated by Beebee on p. 260 of Beebee, 2008.
6 He also adds that the existence of these offenses is established by different laws, and these laws are differently implemented from laws against terrorism and murder. Unlike the point about sentencing it is unclear to me how these points are supposed to contribute to his case.

7 Notice that for dialectical reasons it is important to focus here on conspiracies and preparatory acts whose failures don’t depend on the intervention of the justice system, since Beebee disregards the particular difficulties the preventive pre-punishment presents. For what its worth, I think she shouldn’t disregard them: it seems to be a drawback of her view that it entails not only that there is no widespread intuition against pre-punishment, but also no intuition against the punishment of those who have not and will not commit a crime.

8 Interestingly, the way in which this account of our intuitions about pre-punishment relies on the idea that punishment is a communication aimed at inducing remorse means that it is unavailable to proponents of ‘denunciatory’ versions of expressive theories of punishment. For there is no logical bar to denouncing offenses that have not yet taken place, parallel to the logical bar on feeling remorse or regret for events that have not yet taken place.

9 That apart, I differ from Kearns over two main issues. The first is that it is not as clear to me as it seems to be to him that we could not devise (fictional) cases where we could not devise (fictional) cases where we could be confident that pre-punishment did not play a role in causing the crime for which it was imposed – for example, cases in which the pre-punished individual suffered from amnesia or had their memory wiped. The second is that while it is true that if his line of thought were correct, it would show that pre-punishment is wrong, it seems too subtle to explain why the intuition of the wrongness of pre-punishment is as widespread as he, Smilansky, and I all take it to be.

10 Adam Morton – personal communication.

11 Ulrich Steinvorth in discussion in a Work-in-Progress seminar in the Philosophy Department at Bilkent University, Ankara.

12 As Irem Kurtsal Steen suggested to me in correspondence.

13 Except for those who regard taxes as forced labor.

REFERENCES


