

The Time to Punish and the Problem of Moral Luck

DANIEL STATMAN

ABSTRACT *Christopher New recently argued for the seemingly paradoxical idea that there is no moral reason not to punish someone before she commits her crime ('prepunishment'), provided that we can be sure that she will, in fact, commit the crime in the future. I argue that the air of paradox dissolves if we understand the possibility of prepunishment as relying on an anti-moral-luck position. However, New does not draw the full conclusions from such a position, which would allow prepunishment even prior to the formation of a wrongful intention, on the basis of the idea that bad character is enough to entail blameworthiness. I conclude by examining a legal system which seems to use the idea of prepunishment, i.e. the Talmudic idea that we might punish a child who has not yet sinned in order to save him from sinning in the future.*

In a very intriguing article a few years ago, Christopher New argued that there is no moral reason not to punish someone before she commits her crime ('prepunishment'), provided that we can be sure beyond reasonable doubt that she will, in fact, commit the crime in the future[1]. Neither utilitarian theories of punishment nor retributivist theories should oppose such punishment: utilitarian benefits that can be gained by punishing after crime ('postpunishment') can also be gained by prepunishment, e.g. deterring other potential criminals, and the most basic concern of retributivists is also met, i.e. the person who is punished *deserves* the punishment. The intuitive censure of prepunishment, argues New, is thus mere prejudice.

In response, Saul Smilansky has argued that prepunishment is morally unacceptable, because it violates our deepest Kantian intuitions about respecting persons as autonomous human beings. To punish someone before he commits a crime is to ignore the possibility of 'a last-minute moral improvement' [2]. It means treating him as an object, without taking his autonomy seriously. Hence, assigning importance to a specific temporal order between crime and punishment is no prejudice. It is based on solid moral grounds.

New rejects Smilansky's objection[3]. He argues that if we knew for sure, as God might, that someone were going to commit a crime, there would be no need to wait to see that he actually commits it in order to punish him. By doing so we would not be infringing upon his autonomy by preventing the possibility of a 'last-minute moral improvement' because, *ex hypothesi*, we know that this possibility will not be actualised. A further problem in Smilansky's account is the vagueness and ambiguity of the idea that we should respect people's autonomy. As New shows, this idea can be shown to fit his own (New's) account too, i.e. to endorse prepunishment[4].

In this essay I suggest a new understanding of New. I argue that his seemingly paradoxical position on prepunishment presupposes a certain stance in the debate about moral luck, a topic which has attracted a lot of discussion in the last decade. Once we acknowledge this

stance we gain a better understanding of the nature of the debate over prepunishment, and a clearer grasp of some distinctions in the area of moral luck.

I

Friends of moral luck, such as Williams, Nagel and Nussbaum, contend that luck plays a central role in determining one's moral status[5]. By 'luck' they refer mainly to what is beyond one's control, and they argue that '[t]he things for which people are morally judged are determined in more ways than we at first realise by what is beyond their control'[6]. Critics of moral luck, such as Thomson, Rescher, and Zimmerman[7], argue that this picture is wrong, that, at most, luck affects our knowledge of a person's status, not a person's real moral status. The question of moral luck can also be described as the question about the importance in morality of our *actual* actions vs. the actions we *would* carry out (or would *have* carried out) given our intentions and character. That we form particular intentions is usually taken in this context to be (to a large extent) within our control[8], but whether our intention to act is realised or not depends on factors over which we have only limited control. Suppose John intends to kill Bill and shoots at him, but to Bill's good fortune, misses. The formation of the intention is supposedly immune to luck, but not the actual action. Should we regard John as a murderer because it was only pure luck that prevented the actual death of the intended victim, or should we avoid treating him as a murderer because, as things turned out, he did not kill Bill? Proponents of moral luck would argue that John should not be regarded as a murderer (either morally or legally), simply because he did not carry out an actual act of murder. By contrast, critics of moral luck would argue that a clear intention to commit murder ('not as a mere wish but as the summoning of all the means in our power') is both necessary and sufficient for the agent to deserve the status of a murderer; at least morally speaking, luck cannot exempt people from what they ultimately deserve, nor entitle them to what they do not really deserve.

In a similar way, critics of moral luck argue that people might be blameworthy for what they would have done given their corrupt character, even if in fact they did nothing wrong. Consider the following example by Judith Thomson:

Here are Judge Actual and Judge Counterfactual. Both are corrupt; both would accept a bribe if a bribe were offered and it was large enough . . . The defendant in a suit before Actual now offers Actual a large bribe, and Actual happily accepts it. If the defendant's suit were being tried before Counterfactual, he would have offered Counterfactual the same bribe, and Counterfactual would have accepted it equally happily, but by luck of the courthouse draw, the suit is tried before Actual, and nobody offers Counterfactual a bribe[9].

Should the fact that Actual took while Counterfactual did not take a bribe establish any moral difference between these two judges? Thomson thinks it should not. In her colourful words, God, knowing all the relevant facts, would throw Counterfactual to no deeper circle of Hell than the one into which He would throw Actual (op. cit).

Let us return now to the problem of prepunishment. As both New and Smilansky agree, utilitarians would have no special problem with such an arrangement if it proved successful and conducive to general happiness. Thus, if prepunishment raises an interesting philosophical puzzle — as it surely does — it is a puzzle mainly for retributivists. As I

indicated earlier, my contention is that whether or not we accept the idea of prepunishment depends on the larger question of whether or not we accept the possibility of moral luck. If luck is allowed to determine moral judgments, namely, if what we *actually* do is necessary to determine our moral desert, then prepunishment is unacceptable. If an agent has not yet carried out an action, then she cannot be blameworthy for it, even if she will definitely do it, or would definitely do it unless blocked by some external factors. By contrast, if luck is excluded from the domain of morality, then the agent's desert is determined by her intentions and character, irrespective of their contingent manifestation in the world as it is [10]. And this would entail accepting the possibility of prepunishment, because once an agent forms the relevant criminal intention (again, not as a mere wish), this is sufficient to make him deserve the relevant punishment, with no need to postpone the verdict until the actual action takes place.

New rightly emphasises that due to epistemic shortcomings we will probably never be able to practise prepunishment, and a similar claim is made by critics of moral luck, for example, with regard to the idea of punishing people for what they would have done. As we lack complete knowledge, we must base the institution of punishment on what people actually do, not on what they will do, would do, or would have done. But that, contends New — together with critics of moral luck — has nothing to do with one's real desert, which is immune to the contingencies of reality and which does not depend on what one actually does.

II

Having seen that what underlies New's idea of prepunishment is a Kantian objection to moral luck, we can now notice how New stops short of deducing the full conclusion from such a position. The idea of prepunishment is that of punishing someone before he commits an offence. Now in the normal use of this expression, 'to commit an offence' is both to form the appropriate intention and to carry the intention out. Thus, we might suppose that in prepunishment we punish an agent, A, at t1, for the fact that at t2, A will do (=intend and carry out) some offence. However, New expressly separates between the intention and the action, having the intention at t1, i.e. *before* the (pre?)punishment, the punishment at t2, and the offence itself at t3, i.e. *after* the (pre)punishment. He says that there would be nothing wrong with prepunishment 'if only we could be sure . . . that he *intended to* and *would* in fact commit it' [11]; the offence, he stresses, '*was* intended and *would be committed*' [12]. Thus, while New accepts pre-act-punishment, he seems to oppose pre-intention-punishment.

The problem is that as far as the logic of New's argument goes, the distinction between these two notions of prepunishment hardly makes sense. From a utilitarian point of view, punishing a would-be violator might be justified even if she has not yet formed the required wrongful intention. After all, waiting for the intention to rise and then intervening between the intention and the action might be quite difficult. And from a retributivist point of view there would be nothing unjust in punishing A prior to forming the wrongful intention, because as A will definitely (*ex hypothesi*) do (= intend and carry out) an offence, A deserves punishment, and the requirement that A receive it only *after* the offence should also count as mere prejudice.

In terms of the debate over moral luck it seems that New limits himself to (objecting

against) one aspect of moral luck, i.e. 'resultant luck', while ignoring other relevant aspects, in particular what Nagel calls 'circumstantial luck'. Resultant luck is luck, good or bad, in the way things turn out after the agent has made a certain decision or has taken some course of action. The assumption underlying this sort of luck is that even if the decision (e.g. to go to war) or the course of action (e.g. shooting at John) are within the agent's control, their results are not. Thus, referring to the above examples, a leader should not be blamed for the results of losing a war provided that his decision to go to war was reasonable at the time of the decision, while a potential 'unlucky' murderer whose murder attempt failed as a result of pure luck should be regarded as no different from the successful (the 'lucky') murderer. Immunity from this kind of luck was Kant's main object in the famous opening passages of the *Grundlegung* where Kant praises the unconditional value of the good will that 'sparkles like a jewel' irrespective of its actual results in the empirical world.

Circumstantial luck, suggests Nagel, is luck in the kind of problems and situations one faces. For example, ordinary citizens of Nazi Germany faced a serious moral test, whether to join the regime or oppose it. It was a test to which citizens of other countries were not subjected though had they been subjected to it, no doubt many would have failed the test and behaved just like the actual Nazis. Yet we judge the actual Nazis differently than the would-be Nazis, hence, concludes Nagel, 'we judge people for what they actually do or fail to do, not just for what they would have done if circumstances had been different'[13]. Thomson's example of Judge Actual and Judge Counterfactual that I quoted above is also of this kind.

The crucial difference between these two kinds of luck is that in resultant luck a causal link exists between some aspects of the agent's behaviour, which are supposedly immune to luck, and some results which are (also) a matter of luck. In other words, we are speaking of cases where the agent's behaviour is a necessary condition for the production of some results, but where luck plays a central role too, for instance, when one drives negligently and runs over a child who jumps into the road. By contrast, in circumstantial luck no such causal link exists. The German immigrant who would have become a Nazi had he stayed in Germany in the 1930s and not gone to Argentina on business did nothing to facilitate either directly or indirectly the Nazi atrocities. Yet precisely the absence of such a causal link was a matter of luck, because had he stayed in Germany he would have contributed his part to these atrocities, just as all his class-mates did. (For the sake of the argument let us assume that the truth of this counterfactual can, in principle, be established.)

Using this distinction, we can now see how the justification for pre-act-punishment draws on the rejection of *resultant* luck. In New's case, the potential driver has already formed the morally (and legally) wrong intention, an intention which is causally connected to the future crime. Since, according to opponents of luck, intentions are within the agent's control but their results are not entirely so, and since blame and praise depend on control and responsibility, we can punish agents once they have formed the wrong intentions and we need not wait to see the intentions actualised.

The possibility, however, of pre-*intention*-punishment (a possibility ignored by New), would rely on the rejection of *circumstantial* luck. Such a rejection is explicitly endorsed by opponents of moral luck, such as Thomson (see p. 130 above) Richards, Rescher and Zimmerman. According to Zimmerman, the moral blameworthiness of the actual and the would-be Nazi is the same, and if you ask what the would-be Nazi is to blame *for*, the answer is: 'he is to blame for being such that he would have made the decision to collaborate had he been in a situation that he believed to be *s* (where *s* is the situation that the collaborator

believed himself to be in'[14]. To connect this view to the idea of pre-intention-punishment: If we knew for sure that somebody would kill her husband if she found out that he was not loyal to her, why not put her in prison now? She, first of all, deserves it, since she is the sort of person who would commit murder in these circumstances, and, secondly, such prepunishment might be rather helpful; at least to her husband . . .

True, we can never know for sure (or anything close to sure) how somebody would behave in the future, or would have behaved in the past. But, as New concedes, the same holds true for pre-act-punishment too. The question is whether *in principle* prepunishment might be justified, not whether it is justified given our limited knowledge.

Let me summarise my argument: New's endorsement of prepunishment is best understood as relying on a Kantian-like opposition to moral luck. Yet such opposition should have led New to an even more extreme view than the one he recommends, to the view that prepunishment might be justified even prior to the formation of a criminal intention. It is hard to see why the wish to neutralise luck should stop at pre-act-punishment and not extend to pre-intention-punishment too. That such an extension is natural derives from the tight connection that obtains between character and blame. When we blame a person, *P*, for some behaviour, *b*, 'we are interested in this behaviour as a display of character'[15]. This explains the standard excuses — 'I was ignorant of the facts,' 'I don't know what happened to me,' etc. — which attempt to sever the connection between some particular behaviour and the agent's character. If an excuse is established, this means that though *P* actually committed *b*, *P* is nevertheless not blameworthy for it because *b* does not manifest *P*'s character. When it comes to blameworthiness and to the justification of punishment, the notion of character is thus fundamental and it is this notion that underlies the objection to luck, both resultant and circumstantial; as we blame agents for doing *b* only insofar as *b* is a display of their character, why not blame the corrupt character 'directly,' so to say, without waiting for the behavioural display. In other words, the objection to moral luck assumes that character is both a necessary and a sufficient condition for moral responsibility. Hence, *P* is to blame for reckless driving even if, fortunately, nobody was hurt, because such driving expresses bad character, and similarly *P* is to blame for the fact that *P* would have driven recklessly, because that *P* would have behaved this way demonstrates *P*'s corrupt character.

III

The possibility of prepunishment, on New's view, is a purely theoretical one which, even if accepted, would have no implications for our human institution of punishment, Owing to the limits of human knowledge, we are not in a position to know that some agent will commit some crime in the future in a way that might allow us to punish him in advance. That being so, it would be natural to expect that in legal systems assuming the existence of an omniscient legislator prepunishment could be a more workable idea. Indeed, we find an interesting instance of such reasoning in the *Talmud* in the reference to the biblical law of the stubborn and rebellious son (*Deot.* 21). The punishment inflicted on the son — stoning to death — is rather harsh and seems inappropriate given the son's crime (disobedience to his parents, gluttony and drunkenness) and given that, after all he is only a child. This moral problem is explicitly raised by the sages and the following suggestion is made:

R. Jose the Galilean said: Did the Torah decree that the rebellious son shall be brought before Beth Din [= the court] and stoned merely because he ate a *tartemar* [weight measure] of meat and drank a *log* [liquid measure] of Italian wine? *But the Torah foresaw his ultimate destiny.* For at the end, after dissipating his father's wealth, he would [still] seek to satisfy his accustomed [gluttonous] wants but being unable to do so, go forth at the cross roads and rob. Therefore the Torah said, 'Let him die while yet innocent, and let him not die guilty'[16].

The key idea here from the point of view of punishment is that the son is tried and, indeed, sentenced because of what he would ultimately become, 'his ultimate destiny,' or to translate literally, on account of 'his end' [*sof*]; the son is punished today, though he is only a child and though the crimes he has committed were not so bad, because in the future, if he grew up, he would commit much worse crimes. And this prediction is based on the special knowledge of the *Torah*, not on partial and limited human knowledge. The acceptance of the justice in prepunishment is thus essential in justifying the above law.

Note, however, the interesting difference between the justification for prepunishment that the *Talmud* suggests in the above quotation and the justification I mentioned earlier. According to the *Talmud*, the ultimate justification for killing the stubborn son is that he should die innocent and not die guilty. That is, at the time of the prepunishment the son is regarded as *innocent* and the rationale for prepunishing him is precisely to keep him in this status. Thus we impose on the son at t_1 a punishment that he does *not* deserve at t_1 , because at t_2 he would deserve it, and because we want to prevent him from becoming guilty. By contrast, on the rationale for prepunishment suggested above, we might punish a future criminal *P* at t_1 because though *P* has not yet committed the crime, *P* is nevertheless blameworthy and deserves to be punished for being such a person that she would commit the crime if she were in the appropriate circumstances.

To sum up, then, I think that New is right in holding that prepunishment is not necessarily paradoxical and I tried to show why. The central points of my argument were as follows: (a) The problem of prepunishment is just a branch of the wider problem of moral luck. (b) Viewed in this manner, pre(act)punishment is no more paradoxical than a consistent anti-moral-luck position which assigns a merely epistemic role to actual actions in determining real desert. (c) The air of paradox is further dispelled in the cases discussed by New, because the punishment comes *after* the agent forms the wrongful intention, and, in this respect, fits the common requirement that punishment follow crime. (d) A more extreme anti-luck position would allow prepunishment even prior to the formation of a wrongful intention, on the basis of the idea that bad character is enough to entail blameworthiness. (e) A further justification for prepunishment might be a desire to save the punished from sinning in the future.

Remember, however, that the idea of prepunishment is a purely theoretical one. In the world as it is we have strong epistemic and practical reasons not to use it[17]. Thus, because of well-known human limitations, post-punishment is the only decent institution of punishment available to us[18].

Daniel Statman, Department of Philosophy, Bar-Ilan University, Israel 52900.

NOTES

- [1] CHRISTOPHER NEW (1992) Time and punishment, *Analysis*, 52, pp. 35–40.
- [2] SAUL SMILANSKY (1994) The time to punish, *Analysis*, 54, p. 52.
- [3] CHRISTOPHER NEW (1995) Punishing times: reply to Smilansky, *Analysis*, 55, pp. 60–62.
- [4] Kant was aware of the difficulties in applying the principle of autonomy to practical questions and therefore for practical matters recommended the use of the first formula of the Categorical Imperative. See IMMANUEL KANT (1964) *Groundwork of the Metaphysics of Morals*, trans. H. J. Paton (New York, Harper & Row), end of ch. 2.
- [5] Their articles are included in DANIEL STATMAN (ed.) (1993), *Moral Luck* (Albany, NY, SUNY Press).
- [6] T. NAGEL, Moral luck, in STATMAN, op. cit., p. 59.
- [7] See STATMAN, op. cit.
- [8] This assumption is of course not uncontroversial. One of Nagel's objectives is to show that even with regard to intentions and character luck plays an ineradicable role.
- [9] J. THOMSON Morality and bad luck, in STATMAN, op. cit., pp. 206–207.
- [10] For a consistent and rather extreme call to exclude luck from the criminal law, see KIMBERLY D. KRESSLER (1994) The role of luck in the criminal law, *University of Pennsylvania Law Review*, 142, pp. 2183–2233.
- [11] NEW, *supra* note [3], p. 60.
- [12] NEW, *supra* note [1], p. 40, italics added.
- [13] NAGEL, in STATMAN, op. cit. p. 66.
- [14] M. ZIMMERMAN, Luck and moral responsibility, in STATMAN, op. cit., p. 228.
- [15] N. RICHARDS, Luck and desert, in STATMAN, op. cit., p. 169.
- [16] Sanhedrin 72A, italics added.
- [17] To be sure, we have good reasons to take measures in order to prevent crimes before the crimes are carried out, but such measures do not of course amount to *prepunishment*.
- [18] I wish to thank Saul Smilansky for introducing me to this topic and for helpful comments on an earlier draft.