

The Utilitarian Justification of Prepunishment

Voin Milevski

Institute of Philosophy, University of Belgrade

Abstract. According to Christopher New, prepunishment is punishment for an offence before the offence is committed. I will first analyze New's argument, along with the epistemic conditions for practicing prepunishment. I will then deal with an important conceptual objection, according to which prepunishment is not a genuine kind of 'punishment'. After that, I will consider retributivism and present conclusive reasons for the claim that it cannot justify prepunishment without leading to paradoxical results. I shall then seek to establish that from the utilitarian point of view it is possible to provide a plausible justification of this practice. Finally, I shall attempt to defend the claim that the fact that utilitarianism can justify prepunishment in a satisfactory way is clearly a favourable characteristic of this ethical position.

Many philosophers have recognized that we have a strong intuition against punishing someone for a crime she has not committed.¹ On the basis of this intuition, they defend the view that punishment of the innocent is ethically unacceptable. Prepunishment is punishment for an offence before the offence is committed (New, 1992, p. 35). Because in both cases we would be punishing people who are, legally speaking, still innocent, prepunishment seems to be strikingly similar to the punishment of the innocent. And, since prepunishment is so easily identified with the punishment of the innocent, according to Saul Smilansky, it is but a logical step to conclude that prepunishment is also ethically unacceptable (Smilansky, 1994, p. 50). However, Christopher New believes that this identification is misleading and that prepunishment should not be identified with the punishment of the innocent. In New's view, only the epistemic problem of being sure beyond reasonable doubt that the agent would commit the offence prevents us from practising prepunishment. Were we sure beyond reasonable doubt that the agent would commit the offence, we would not be punishing the innocent in the relevant sense, and prepunishment would be morally justifiable (New, 1995, p. 60). As New says:

¹ This work was supported by the project "Dynamic Systems in Nature and Society: Philosophical and Empirical Aspects" (no. 179041) financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Prepunishment is a strange notion, admittedly, then, but not one which seems to be clearly excluded by epistemic or logical considerations or by either of the main theories of punishment. [...] In the case of punishment, of course, we cannot now satisfy, and may never be able to satisfy, the epistemic condition for practicing prepunishment. But the question is: is it only our lack of foreknowledge that prevents us, or is there some deeper, moral, objection to it? My suggestion is that there is no deeper moral objection (New, 1992, p. 40).

From this arises the following question: is New's argument, with which he attempts to show that it would not be morally wrong to prepunish an agent if we could be sure beyond reasonable doubt that she would commit an offence, acceptable? The answer to this question, I will argue, can be used in order to formulate a very persuasive defence of the utilitarian theory. I will first analyse New's argument, along with what he calls 'the epistemic condition for practicing prepunishment.' This analysis will bring me to the conclusion that it is possible to distinguish two versions of prepunishment, namely weak and strong. I will then deal with an important conceptual objection, according to which prepunishment is not a genuine kind of 'punishment'. After showing that ethical discussions of prepunishment are of particular importance, I will consider retributivism and present conclusive reasons for the claim that it cannot justify prepunishment without leading to serious difficulties and even paradoxical results. I shall then seek to establish that from the utilitarian point of view it is possible to provide a plausible justification of this practice. Finally, I shall attempt to defend the claim that the fact that utilitarian theory can justify prepunishment in a satisfactory way is clearly a favourable characteristic of this ethical position. That, in brief, is the aim of this paper. Let us now consider in more detail New's argument for prepunishment. The argument may be presented as follows:

- (1) If we could be sure beyond reasonable doubt that an agent intended to and would in fact commit an offence, it would not be morally wrong to prepunish her.
- (2) We might in fact sometimes be sure beyond reasonable doubt that an agent intends to and will commit an offence.
- (3) Therefore, it would not be morally wrong to prepunish an agent who would commit an offence in the future.

The argument is valid. Its logic is perfect. The only question is whether the premises are true. The first premise is a conditional; it says that if we could satisfy the epistemic condition expressed by the antecedent, there would be nothing morally wrong with prepunishment. The second tells us that in some cases we can, in fact, satisfy this epistemic condition. A great deal could be said about the second premise. For instance, we can say that, as a matter of fact, we can never know beyond reasonable doubt what any person is going to do in the future. And, certainly, if we can successfully defend the view that we possess,

say, a libertarian free will or that for any other reason our free choices cannot be foreknown beyond reasonable doubt, we must reject the second premise. However, until we are presented with a persuasive and fully elaborated defense of the view that foreknowledge beyond reasonable doubt of our future acts is impossible, I think we should grant this premise.

What about the first premise? This premise is a conditional: it says that if the antecedent is true, so is the consequent. In other words, it tells us that if we could satisfy the epistemic condition, prepunishment would not be morally wrong. I think this premise is plausible. But in order to see this more clearly, we need to examine what it means exactly to say that we can satisfy the epistemic condition. Although New does not say much on the subject, it seems that to satisfy this condition is the same as to have a complete description of the particular would-be crime (i.e. the predicted crime for which we want to prepunish a person). So, the next question we need to answer is this: what can we know about the would-be crime? The most straightforward answer is that we can know the identity of the would-be offender, as well as some other relevant facts about her would-be crime, e.g. the exact time and location of the crime, the motive, intentions, and other psychological states behind the crime. These facts constitute a complete description of the crime, and to know this description is the same as to fully satisfy the epistemic condition for practicing prepunishment. Obviously we cannot now satisfy, and perhaps may never be able to satisfy, the epistemic condition, but it is at least imaginable that at some point in the future we will be able to satisfy it: such a scenario seems coherent.

Furthermore, it is important to note that, based on the knowledge of the complete description of the predicted crime, we may also know what it would be necessary to do to the would-be offender, or, more precisely, how long it would be necessary to detain the would-be offender in order to prevent her from committing the crime. (For the sake of simplicity, I will continue to talk about prepunishment as involving *preventive detention*, although this is not the only way in which we can successfully prevent crimes.) Of course, it is reasonable to suppose that in some cases we simply won't know how long we need to detain a person to successfully prevent her intended crime, despite the fact that we have the complete description of the crime. In such cases it would be possible for an agent to still commit the crime after she was prepunished. In addition, in the case of offences for which the fixed penalty is a fine, prepunishment would not prevent an agent from committing the offence. For instance, let us suppose it is known that someone is going to commit an offence tomorrow, for which the fixed penalty is a fine. Obviously, prepunishing the person by fining her now would not prevent her from committing that offence tomorrow. In fact, the person may even feel entitled to commit the offence for which she had already been prepunished.

Based on these considerations, even though New talks as if there is only one kind of prepunishment, it is actually possible to distinguish two versions of this practice, weak and strong. Strong prepunishment is a practice that provides us with necessary and sufficient conditions for the successful prevention of all predicted crimes. So, if in attempting to prepunish we know the identity of the would-be offender, the exact time and location of the would-be crime, the motive, etc., as well as how long it would be necessary to detain the person to keep her from committing the crime, we have all we need to successfully prevent that crime from happening, and, accordingly, we get the strong version of prepunishment. Certainly, this kind of prepunishment would be impossible in the case of offences for which the fixed penalty were a fine, or in those cases where we could not determine how long to detain the would-be offender in order to achieve the successful prevention of predicted crimes. On the other hand, weak prepunishment could be defined as follows: it is a practice that does not preclude the possibility that a person could still commit the crime after she has been prepunished. Thus, for instance, if in attempting to prepunish we know all the relevant facts about the would-be crime, without knowing how long to detain the would-be offender to make her crime impossible, or, alternatively, if we are dealing with offences for which the fixed penalty is a fine, the most we will be able to achieve is a weak form of prepunishment, after which it would still be possible for the person to go on and commit the intended crime.

Yet both the weak and the strong version of prepunishment face a serious conceptual problem. Namely, it might be objected that although we can accept that prepunishment is a useful system of crime prevention, it is not, strictly speaking, a genuine kind of 'punishment,' but is instead some sort of preventive detention, or, to use Bradley's phrase, some sort of 'social surgery.' It is a conceptual fact that punishment, whether pre- or post-, requires that a crime be committed. In other words, a crime's having been committed is the logical presupposition of punishment. There are two lines of reasoning that we could use in order to challenge this objection. First, if we define legal punishment as the 'deprivation of goods,' as some philosophers believe we should (Gendin, 1967, p. 237), then this objection loses almost all of its plausibility; prepunishment is unquestionably a kind of deprivation and, consequently, is a kind of punishment. Second, it is important to notice that this objection represents what H. L. A. Hart called 'the definitional stop' (Hart, 1968, p. 5). According to Hart, the definitional stop is an abuse of definition that prevents us from investigating the most puzzling issues about punishment. The objection that prepunishment is not, by definition, a genuine kind of punishment and that, as a result, it has no relevance in philosophical discussions of legal punishment seems to be a clear example of such an abuse of definition. However, I don't think that arguments along these two lines of reasoning pose a serious threat to the objection that prepunishment is not a genuine kind of punishment.

In fact, it is quite obvious that strong prepunishment cannot be called punishment, since this practice completely prevents the commission of future crimes. Evidently, whatever is done to a would-be offender that completely prevents her from committing a crime, it cannot be called punishment for that, uncommitted, crime. In contrast, the weak version of prepunishment can perhaps be called punishment, but only if the would-be offender does commit the crime after she has been prepunished. Still, despite the fact that it is possible for weak prepunishment to represent a genuine kind of punishment, I think that prepunishment should not be generally considered a practice that inflicts *punishment* on the would-be offender; rather, it is a practice that directly aims at crime prevention. So, instead of accepting New's definition of prepunishment, we should say that by prepunishing we are attempting to prevent people who intend and would commit crimes from actually committing them. The proposed definition allows us to straightforwardly reject Smilansky's objection that if we consider the punishment of the innocent to be intrinsically wrong, we should conclude that prepunishment is morally unacceptable (Smilansky, 1994, p. 52). Contrary to Smilansky's objection, we can see that if prepunishment cannot be identified with punishment, it cannot be identified with the punishment of the innocent, either.

Still, punishment is unquestionably one of the basic institutions of every political society. Hence, if prepunishment is not to be considered a genuine kind of punishment, what kind of philosophical importance should we attach to this practice? I will argue that the main importance of prepunishment is to be found in the fact that in circumstances where the epistemic condition was satisfied, we would, most certainly, modify our punishment systems so that they would function in the same or very similar way as the system of prepunishment. To make this claim as convincing as possible, let us suppose that we have satisfied the epistemic condition and that we can determine exactly how long a person needs to be detained in order to successfully prevent her from committing a crime. It would surely be a puzzling thing to say that we still have to wait until this person actually commits the crime, so that we can punish her afterwards. What would be the point of waiting? In attempting to answer this question, Smilansky stresses that in prepunishment there is still time for the would-be offender to change her mind and decide not to commit the crime. This moral opportunity for the would-be offender to change her mind needs to be acknowledged, and is the essential point of waiting for the person to commit the crime—and thus not prepunishing her before giving her a chance to make her own choice (Smilansky, 1994, p. 52). In contrast to this view, New maintains that in such cases where the epistemic condition for prepunishment is satisfied, waiting for a person to commit the crime would be completely pointless:

If we knew she was not going to change her mind, giving her an opportunity to do so would have no point; it would be waiting for something to happen which we knew was

not going to happen. There is a point in giving someone such an opportunity if we don't know whether she is going to change her mind; then, perhaps, we ought to wait and see, we ought 'to give her a... chance to remain innocent.' But if we do know, there is no point in waiting and seeing—we already know how she is going to manifest her moral personality and exercise her autonomy; there is no last-minute improvement to hope for (New 1995, pp. 61–62).

New has a good point here. If we have to wait and see that the person is going to commit the crime, it only means that we haven't fully satisfied the epistemic condition and that we don't know whether the person is going to change her mind about her intended crime. However, as we have seen, the full satisfaction of the epistemic condition presupposes that we have a complete description of the would-be crime. Since this description includes facts about the relevant psychological states of the would-be offender in the moment of offence, this means that we *know* that she will not change her mind, and in that case there is certainly no point in not preventing that crime from happening. Admittedly, this kind of 'crime prevention' would not be properly called 'punishment', but that is quite irrelevant here. We have seen that ethical discussions of prepunishment can have relevance in their own right, even though, strictly speaking, this practice does not represent a genuine kind of punishment. Hence, there are good reasons to examine whether it is possible to provide a moral justification of prepunishment, and this is what I intend to do in the remainder of this paper. I will consider two theories, utilitarianism and retributivism, and try to find out if these theories can justify prepunishment.

Let us first consider the retributive theory. As John Cottingham points out, the term 'retributive' has become so imprecise and has so many meanings and interpretations that it is doubtful whether it can serve any useful purpose (Cottingham, 1979, p. 238). For this reason, it might be difficult to establish whether any characterisation represents a genuine definition of retributivism. However, Cottingham's analysis shows that the desert theory is both historically and etymologically an authentic version of retributivism (1979, p. 239). In addition, it is arguably the most popular version of this theory. Therefore, we can accept that retributivism is the theory according to which the fact that punishment is *deserved* is sufficient to justify it.

New claims that there is nothing in this theory that prescribes that the offender should suffer after rather than before the offence and that, therefore, prepunishment can be justified by the retributive theory (New, 1992, p. 37). And, since retributivism is the application of the principle of desert to the problem of punishment, in order to show that New is quite wrong about the retributive justification of prepunishment, it is important to make the concept of desert as clear as possible. First, most desert theorists accept the view that desert is a three-place relation between a person, the grounds on which she is said to deserve (the desert base), and the treatment or good that she is said to deserve (the deserved good). Also, most desert claims have, or are supposed to have,

normative force. To say that someone deserves something is to claim that she ought, other things being equal, to get that thing, or that it would be morally better that she get it. However, beyond the identification of these very general features, desert theorists tend to disagree over what else may be said about the concept of desert (Olsaretti, 2003, p. 4). Consequently, the concept of desert is not especially well understood and there are many difficult questions surrounding this concept.

In order to improve our understanding of this concept, Fred Feldman analyses two general theses about desert and desert base, which are considered to be part of the received wisdom about desert (Feldman, 1995, p. 64). The first thesis links the concept of desert to the concept of responsibility: a person cannot deserve anything in virtue of an action or fact unless she is responsible for that action or fact. According to the second thesis, the desert base and desert necessarily stand in a certain temporal relation. More precisely, the desert base must always precede desert. In that sense, desert of rewards and punishments seems rooted in the past, so that if at t S deserves x in virtue of the fact that S did or suffered something at t' , then t' cannot be later than t . Feldman believes he can show that this is not the case. His example concerns soldiers who volunteer for suicidal missions. These soldiers are thought to be deserving of great honours; they may even be given medals or promotions *before* they go off to perform the actions in virtue of which they deserve to be so treated. In Feldman's view, this represents a clear example in which we can see that the desert base is rooted in the future, rather than in the past (1995, p. 71). Therefore, the desert base need not always precede desert. New accepts this view about desert. In fact, he uses a very similar example to show that prepunishment respects moral personality and autonomy:

Suppose that the commanders of kamikaze pilots in World War II had known before the pilots took off on their suicidal missions that they would carry them out successfully. They might then have wished to decorate the pilots for their gallantry before they took off, rather than only posthumously. We would surely not regard that as a failure to respect, but rather as an acknowledgement of, the pilots' moral personality and autonomy. Why, then, should we treat the case of punishing differently (New, 1995, p. 62)?

Now we are in a better position to evaluate New's claim about the retributive justification of prepunishment, according to which there is nothing in this theory that prescribes that the offender should suffer after rather than before the offence. It is important to notice that for this claim to be true, Feldman's view about the temporal relation between desert and the desert base has to be correct. More precisely, it must be possible to *prove* that in some situations we can deserve something even if the base of that desert comes later than the desert itself. However, Feldman does not succeed in proving his position. Though he calls our attention to interesting and, at least *prima facie*, acceptable cases where the desert base does not precede desert, his position still remains quite

controversial. Hence, New's retributive justification of prepunishment depends on a questionable account of desert. The traditional view, according to which the desert base must always precede desert, gives us a strong reason not to believe that prepunishment (weak or strong) could be justified by the retributive theory.

But there are other serious problems with the retributive justification of prepunishment. Namely, it is possible to show that the retributive justification of this practice is either impossible or that it leads to paradoxical results. Let us start with strong prepunishment. Retributivists cannot justify this practice, because it completely prevents the commission of future crimes, and the complete prevention of future crimes means that the base of desert (i.e. the base that constitutes the retributive justification of prepunishment) will also be prevented. Hence, the argument that shows why the retributivist theory cannot justify strong prepunishment can be put like this:

- (1) According to the retributive theory, the fact that the punishment is deserved is sufficient to justify it.
- (2) The desert base is the fact to which we appeal in order to explain someone's desert.
- (3) The offender deserves punishment because of the offence she has committed.
- (4) Therefore, the offence committed is the base of desert for punishment.
- (5) If we prepunish a would-be offender, she is not going to commit the offence for which we are going to prepunish her.
- (6) Therefore, prepunishment prevents the base of desert from happening.
- (7) But the base of desert is the retributive ground for moral justification of punishment.
- (8) Therefore, prepunishment cannot be justified by the retributive theory.

The argument appears to be logically perfect. Also, all the premises are either widely accepted definitions or claims that could be logically derived from these definitions. For this reason, contrary to New's claim, we can conclude that the retributive justification of strong prepunishment not only seems problematic, but is completely impossible. But the situation is no better with weak prepunishment. In fact, the retributive justification of this practice leads to paradoxical results. Since the offence committed is the base of desert for prepunishment, retributivists can justify this practice only if the would-be offender commits her intended crime after she is prepunished. So, in order to justify weak prepunishment, retributivists must insist on the commission of future crimes. It is easy to see why this is paradoxical: retributivists insist on the

commission of crimes in order to justify a practice that directly aims at crime prevention. To sum up: the retributive justification of prepunishment depends on the problematic and controversial theory of desert; it is also impossible in the case of strong prepunishment and it leads to paradoxical results in the weak version of this practice. Let us now see if the utilitarian theory can fare better than retributivism.

Utilitarianism is the theory according to which the moral rightness of our acts and the acceptability of our social practices are to be judged solely by their consequences. The utilitarian point of view essentially rests on the belief that human acts are to be judged by their consequences; that is to say, the consequences of our actions and practices are the sole criterion of their morality. The utilitarian theory of punishment is just the application of utilitarianism as a general ethical theory to the problem of justification of punishment. In this sense, punishment is morally justified by its good consequences, and the most important consequences of punishment are its preventive effects: its contribution to the prevention of offences, to the reduction of the rate of law-breaking behaviour (Primoratz, 1989, p. 10). Therefore, the ideal goal of the utilitarian theory of punishment is the complete prevention of future crimes.

Now, is it possible to provide a utilitarian justification of prepunishment? It seems highly unlikely that utilitarians would justify weak prepunishment, since this practice makes it possible for a person to commit a crime after she has been prepunished. In other words, weak prepunishment is not a particularly effective system of crime prevention, and as such, from the utilitarian point of view, it could be justified only if other, more effective, system of crime prevention is not obtainable. Still, this does not mean that the utilitarian justification of weak prepunishment is, in any way, impossible or paradoxical, as in the case of the retributive justification of this practice. On the other hand, when it comes to strong prepunishment, since it guarantees very effective crime prevention and, since preventive effects are the most important consequences of a system of punishment, there are obviously very good reasons providing a utilitarian justification of this practice.

Even so, the utilitarian justification of strong prepunishment is by no means without problems. The most serious problem has to do with the general objection that utilitarianism departs from the principle of proportionality. According to this principle, serious crimes merit severe punishments, while minor infractions should receive only mild punishments. The principle is widely accepted in philosophical discussions about the appropriate justification of legal punishment and to what it must conform, and any departure from it would represent a serious problem for an ethical theory. It is not, however, very difficult to show that strong prepunishment could result in violations of this principle: this practice could make it possible to detain someone for a very long period of time in order to prevent a relatively minor offence. This is a serious problem for utilitarianism;

a problem for which I, unfortunately, cannot offer any satisfactory solution here. Nevertheless, it seems that, overall, utilitarians can provide a much better justification of prepunishment (both weak and strong) than retributivism. First, unlike in the case of retributivism, the utilitarian justification of weak prepunishment would not lead to paradoxical results: utilitarians do not insist on the commission of crimes in order to justify a practice that directly aims at crime prevention. Also, despite the serious objection that utilitarianism might violate the principle of proportionality, a plausible case for the utilitarian justification of strong prepunishment can still be made, while the retributive justification of this practice was proved to be impossible.

But the main question still remains: why is it a good thing that the utilitarian theory can provide moral justification of strong prepunishment? Without a clear answer to this question it would be possible for retributivists to object that the fact that retributivism cannot justify strong prepunishment could be taken as proof that there is something wrong with this practice, rather than that there is something wrong with their position. However, this objection can be rejected without difficulty. Namely, it is true that we cannot now satisfy, and perhaps may never be able to satisfy, the epistemic condition in such a way as to justify strong prepunishment. But it is at least imaginable that at some time in the future we will be able to satisfy it. As has been shown, in these circumstances we would surely modify our punishment systems so that they would function in the same or very similar way as strong prepunishment. And certainly, if crime prevention and reduction of the rate of law-breaking are thought to be our main aim, as it might be quite reasonable to suppose, then the strong version of prepunishment would be a perfect system of crime prevention. Based on this point, we can conclude that the fact that the utilitarian theory provides a plausible justification of strong prepunishment is a favourable characteristic of this ethical position.

References

- Cottingham, J. (1979). Varieties of Retribution. *Philosophical Quarterly* 29, 238-246.
- Feldman, F. (1995). Desert: Reconsideration of Some Received Wisdom. *Mind* 104, 63-77.
- Gendin, S. (1967). The Meaning of "Punishment." *Philosophy and Phenomenological Research* 28, 235-240.
- Hart, H. L. A. (1968). *Punishment and Responsibility*. Oxford: Clarendon Press.
- New, Ch. (1992). Time and Punishment. *Analysis* 52, 35-40.
- New, Ch. (1995). Punishing Times: Reply to Smilansky. *Analysis* 55, 60-62.
- Olsaretti, S. (2003). *Desert and Justice*. Oxford: Oxford University Press.
- Primoratz, I. (1989). *Justifying Legal Punishment*. London: Humanities Press International.

Smilansky, S. (1994). The Time to Punish. *Analysis* 54, 50-53.