PRIMA FACIE RETRIBUTIVISM: 
ON THE OBLIGATION TO ADMINISTER JUSTICE

UDC 343.2.01

Michał Peno

Faculty of Law and Administration, University of Szczecin, Poland

Abstract. The EU system for the protection of fundamental rights had been developing since 1969. The justification of punishment is a difficult problem. The paper attempts to examine retributivism in the normative perspective and to penetrate the structure of the fundamental premises and theses of retributivism. Retributivism assumes that punishment is just, in the broad understanding of the term, while in reality punishment is not just; the model of retributive punishment is contrafactual, which is evident above all in the problem of punishing the innocent. A proper modification of retributivism’s normative premises (i.e. how and why people ought to be punished, etc.) consists in seeing these premises not as unconditionally binding directives but as optimization rules, a kind of prima facie duty. These are mainly the ethical duties of the state considered from the point of view of criminal policy. In effect, it is possible to formulate a non-fundamentalist (non-idealistic) variant of retributivism - better corresponding to social reality. The core of the paper consists in outlining such a concept. The paper has been primarily inspired by the ideas of W.D. Ross and R. Alexy.

Key words: W. D. Ross, inner structure of retributivism, prima facie obligations, punishment of the innocent, criminal law theories

INTRODUCTION

Modern theories of punishment, according to the legal or philosophical tradition may be categorized as utilitarian-teleological or retributive justice-oriented. They constitute ways of justifying or explaining punishment and the practice of punishing. Justification of punishment ought to be understood as providing reasons why punishment is needed or necessary from the social or axiological point of view (why people ought to be punished), while explanation refers to describing the phenomenon of punishing and the related needs, experiences etc. It is worth noting that explaining punishment is sometimes done by

Received December 10th, 2018 / Accepted December 20th, 2018

Corresponding author: Michał Peno, Ph.D., Assistant Professor, Department of Philosophy of Law and Legal Theory, Faculty of Law and Administration, University of Szczecin (Poland), email: michal.peno@usz.edu.pl

reference to the functions which criminal law and punishment perform. In the first instance, it is a teleological function (the realization of certain social goals, punishment is a means) while, in the second instance, it is a function of the administration of justice (punishment is an ethical necessity).  

The subject under scrutiny is the problem of justifying punishment and the retributive justification. The goal of the paper is to confront the justice-based premises of retributivism with the broadly-defined problem of punishing the innocent. The first section presents the premises of retributivism or rather its main varieties in the philosophy of punishment. The second section discusses the controversies regarding punishing the innocent. The final section deals with prima facie retributivism as a response to the dilemma of administering justice as a value in an imperfect social reality, so as to render the obligations resulting from embracing retributivism more realistic (above all as a guiding principle of real criminal policy).

This article does not apply to the relation between retributivism to punishing the innocent, but such concept of a just punishment which will be based on prima facie obligation to punish the guilty. Prima facie obligation to punish the guilty should be understood in such a way that, on the one hand, there is an obligation to punish the innocent; however, on the other hand, this is not an obligation that needs to be fulfilled completely. Instead, what is meant here is the obligation to deploy best efforts, so that in the greatest number of cases a guilty person is punished, whereas an innocent person is protected against the punishment. However, things are not as classic representatives of retributivism perceive them, meaning that there exists an absolute rule (obligation) to punish the guilty, without accepting any exceptions. These exceptions result from the facts of punishing the innocent. From the retributive viewpoint, the measure applied to an innocent person cannot be called “punishment”. In the proposal of the prima facie retributivism presented in the article, there will not be such linguistic or analytical contradiction, while keeping basic attributes of retributive concept.

AN OUTLINE OF DIFFERENT VARIETIES OF RETRIBUTIVISM

Three main approaches can be currently distinguished in connection to retributivism: first – the pure theory of repayment, second – the expressive theory, and third – the fairness theory (social balance theory) (Ryberg, 2004: 43-50).

The first theory refers to a narrowly understood category of retributive justice that demands a punishment proportional to the wrong that was done (i.e to the offence). The perpetrator of an offence deserves to be punished, which constitutes repayment, if not vengeance. The offender ought to be punished because he or she deserves it (Nozick, 1981: 377ff; Zaibert , 2006: 81ff.). Philosophy of criminal law includes a dialogical element in the second theory; namely, the perpetrator deserves to be punished, but, on top of that, punishment should be exacted because of the message it carries, a condemnation of the act, and the demand of society, including the victim, for the wrong to be righted. The perpetrator deserves punishment as well as condemnation, and he or she ought to repent (the so-called secular repentance) (Feinberg, 1970: 98; Primoratz, 1989: 199). The third theory stems from the idea of social contract and social balance, and asserts that the

---

Perpetrator ought to be punished not only because he or she deserves it, but also because a certain balance of benefits and burdens must be restored. The perpetrator enjoys more freedoms and fewer burdens than his or her fellow citizens who chose not to commit an offence and remain honest members of society. Balance ought to be restored, which is possible by means of criminal punishment (Cottingham, 1979; Duff, 1986: 289; Morris, 1968: 475ff.).

All these approaches share a common core—the claim that society or the state has an obligation to repay for a wrong (an offence) manifested in punishment, which can additionally carry a message to the responsible moral subject (the perpetrator), or a means to restore a fair balance of burdens and benefits—ordo iuris. They do not exhaust all possible ways of justifying punishment, of course. The fundamentalist character of this claim does not allow for a formulation of a conditional justification—the one that would include conditionality of the duty to repay. It is so even though dogmatic institutions of criminal law themselves allow for limitations in connection to the obligation to punish, and so the impossibility of justifying punishment in certain circumstances (e.g. for political and criminal protection of offenders who testify in criminal cases) (Husak, 2008). Also, the goal of criminal proceedings in most jurisdictions is not only to punish the perpetrator but also to ensure that an innocent person is not punished (which is an act of balancing the interest of the innocent with the obligation to respond to a crime; sometimes the latter must give way before the former, which reflects the way criminal proceedings or the system of criminal justice is shaped per se) (Cf. Merryman, 1969: 132-148; Duff, 2007: 195ff.).

The next part of the article will be discussed the problem of punishing the innocent. Based on the analysis of this issue, the proposal of the fourth variant of retributivism (ie. prima facie retributivism) will be presented.

**Punishing the Innocent and Retributivism**

We are therefore confronted with the problem of punishing the innocent in connection to the retributive justification of punishment (J. Ryberg, 2004: 43-50; Gross, 1984: 65). Referring to common moral feelings or intuitions, one could say that in the eyes of most people punishing the innocent is a wrong and a mistake. The phrase „the innocent” is not unambiguous. Yet, for the sake of this discussion, we can assume that it denotes a person who should not be punished according to the accepted rules of justice. Guilt is understood as the weight of the wrong done by the offender, which allows us to disregard theoretical constructs of guilt (offence) when discussing punishment. One can easily distinguish two basic situations. The first regards punishing the innocent as a result of an error committed during trial proceedings. This would be a person who, according to the accepted rules of criminal responsibility, should not be punished. The second is related to the way rules of liability are constructed; some approaches do not consider guilt necessary for a penal response. Here one can distinguish two cases: where the penal response is based on objective premises, and where liability is attributed not only based on guilt but also based on other criteria.

Considering the second of the distinguished categories of cases where the innocent may be held responsible, it may be used to formulate anti-utilitarian arguments. Such arguments are put forth by proponents of retributivism so as to prove the axiological superiority of punishment as proportional repayment; the superiority is supposed to consist in the fact that
it is allegedly unhumanitarian and objectifying to treat perpetrators as mere means to achieve certain social goals – to maximize utility (Honderich, 2005: 36ff; Primoratz, 1978: 185–193). Utilitarians maintain that criminal law must achieve certain goals, above all related to crime prevention. From the philosophical standpoint (i.e. without reference to a real system of criminal law in a given jurisdiction), they embrace protection of the innocent as well as abstention of punishing the guilty, as long as it serves the common good (maximalist utility). The utilitarian character, though radicalized in its philosophical and anthropological premises, does include a social defense element referring to social inutility (i.e. a danger) as a basis for a quasi-penal response (Ancel, 1965: 13).

The first approach, on the other hand, raises further questions regarding retributivism, which focus on the sui generis non-universality of criminal justice. “Non-universality” denotes the impossibility to ensure that all offenders in the same situation are punished in the exact same way. There are three reasons why. First, not every offender is held criminally liable. Second, punishment is not determined according to some simple algorithm; as there is no algorithm that allows one to calculate the weight of the punishment as related to the weight of the offense, there is no certainty that offender O₁ will be punished in the same way for crime C as offender O₂ (as far as identity of offences is concerned, one may refer to formal types of offences) (See: Posner, 1985: 1193ff.). Even if there are directives determining the type and severity of punishment, there is no jurisdiction in which they would be purely based on retribution. Every criminal trial must include teleological elements, limiting its purely retributive character (e.g. limitations regarding evidence, the scope of the causal link not necessarily reflecting the entire causal chain). This can be verified empirically, but it hardly seems necessary (besides, ad casum discussions are limited to one state – one legal system – and are not universal or global). Third, not every person held to criminal responsibility is the actual perpetrator of the offence for which he or she is punished.

One can easily see that the cases discussed above are linked to the premise of retributivism which creates an especially acute problem in relation to punishing the innocent – the proportionality rule (Alexander, 2013: 309–319). Simplifying the matter somewhat, we can say that proportionality consists in the relation of the punishment to the wrong that was done. The less severe the offense, the less severe the sanction. The relation of the punishment to the act shaped by the proportionality rule cannot exist unless the act at least constitutes mala prohibit, and if the one punished is the one who deserves it.

From the point of view of retributivism, punishing the innocent constitutes a colloquial generalization resulting from the characteristics of the way truth is determined in criminal proceedings or failures in establishing the facts of the case (Cf. L. Laudan, 2011: 195–227), either resulting from a person’s failure to properly perform his or her duties (e.g. a judge’s negligence) or from circumstances over which those in charge of the proceedings have no control (e.g. an innocent person taking the responsibility for an offence, the imperfection of the methods that may be employed to arrive at the truth). Proponents of retributivism do not ascribe axiological significance to it in the sense that it does not lead them to modify their ideas regarding punishment. They do, however, formulate a certain optimization directive according to which one ought to strive not to punish the innocent and only punish those who deserve it. It is worth noting that some assert that one cannot speak of criminal punishment per se when an innocent person is punished (Quinton, 1969: 55-64). The rationale behind it is that punishment is understood as administration of justice. Punishing an innocent person will not constitute administration of justice. It is a rather marginal
opinion, however, and to some degree contrafactual, for, as a rule, legal and philosophical language does not distinguish between punishment for a crime and punishment resulting from a trial error.

One must, however, distinguish between two dimensions: the normative and the descriptive one. The first stipulates that only guilty persons should be punished, and the innocent should not be punished. The second entails an analysis of the reality of the criminal justice system. The retributive approach to punishment is normative.

Retributivism, thus, gives certain reasons for considering punishment necessary both from the social and the axiological point of view. Yet, is this statement enough? Retributivism as a theory of punishment asserts that the duty to punish the offender, and only the offender, is absolute (Dolinko, 1991: 541-542; Talbott, 1993: 151-168). It is not an idea remote from a certain philosophical and legal reality, considering that the answer Herbert L.A. Hart gives to the question “who can be punished” is: “only the perpetrator of an offence for that particular offence” (Hart, 1968: 11).

For retributivists, justice is a value in itself (Cf. Burgh, 1982: 193ff.). We also know the Pharisee maxim, in principle directed against retributivism: “it is better that one man should die than for the whole nation to be destroyed”; in turn, Immanuel Kant commented that it is better for mankind to be destroyed than for it to neglect the demands of justice. However, it seems that retributivists do not believe that this value (justice) is absolute. Retributivism is limited by the imperfection of human justice (which is why Kant had to assume there is a God and an afterlife), as well as considerations of criminal policy nature and several other factual factors discussed above. It seems that retributivists are not opposed to Voltaire’s statement that it is better for a guilty man to escape justice than to punish an innocent man (Voltaire, 1962: 20). A similar idea was voiced by the Englishman Blackstone (Blackstone, 1844: 358).

### Prima Facie Retributivism

Retributivists declare that it is a duty of society or the state to repay for a wrong (offence) by means of punishment. In principle, they also oppose punishing the innocent (Tebbit, 2015:155-230). The problem is that insistence on not punishing the innocent is one of the important arguments against utilitarianism. Retributivists believe that disregarding offence and guilt manifests unjustified objectivization of the offender. The symmetrical opposite of the right to be punished is the right not to be punished. It is, however, impossible to avoid punishing the innocent, at least in the real world. Classical (retributive) criminal law, both modern and ancient, while treating punishment as vengeance (in its unrefined, primitive form) or repayment (proportional repayment), attempted to reduce the problem of punishing the innocent by not considering guilt a necessary condition of criminal liability or by justifying group responsibility or objective responsibility based on the most elementary factual connections (someone being closest to the victim of a murder, to the shore of a lake where a corpse was found etc.). The historical process of subjectification of criminal responsibility, whose core became the perpetrator’s subjective guilt, excludes such methods or mechanisms by virtue of its nature.

There is, however, a certain way to defend retributivism and make it more realistic. Importantly, it does not mean that one would have to develop a mixed theory, which would include the teleological aspect of punishment (on the one hand) and the duty to repay (on
the other hand). Considering the problem of punishing the innocent and all its implications (and causes) may lead to accepting a certain normative correction of the retributive theory that makes it more realistic. Every penal theory constitutes a model which, as an idealization, enables the study or analysis of punishment. It does not reflect reality, but may respond to its challenges. Taking the axiological standpoint, one may think of W.D. Ross’ theory of *prima facie* duties.

In W.D. Ross’s ethical system, rightness is identified with moral duty. A right act is an act that ought to be performed or morally binding (Ross, 1930: 3, 91–93). Yet, Ross introduces the concept of *prima facie* duty. He suggests the name “*prima facie* duty” or “conditional duty” as “a brief way of referring to the characteristic (quite distinct from that of being a duty proper) which an act has, in virtue of being of a certain kind (e.g. the keeping of a promise), of being an act which would be a duty proper if it were not at the same time of another kind which is morally significant.” Farther on we read: “We have to distinguish from the characteristic of being our duty that of tending to be our duty. Any act that we do contains various elements in virtue of which it falls under various categories. In virtue of being the breaking of a promise, for instance, it tends to be wrong; in virtue of being an instance of relieving distress it tends to be right” (Ross, 1930: 19; Ross, 1926-1927: 127; Ross, 1928: 95-96).

Besides characterizing *prima facie* duty, Ross proposes a list of certain fundamental *prima facie* duties. These are: 1) the duties of fidelity, 2) the duties of reparation (of a wrong), 3) the duties of gratitude (to others for services done by them to one), 3) the duties of justice, 4) the duties of beneficence, 5) the duties of self-improvement, 6) the duty of non-maleficence” (Ross, 1930: 21–27; Johnson, 1969:9). Ross says that one principle can always be abandoned for another, in the sense that some departures from the rule are permitted.

The chapter of *The Right and the Good* dealing with the relationship between the duties of the state and the rights of the citizens in the context of punishment is well-known. According to Ross the state has the duty to protect the innocent. It ought to do everything in its power in order to prevent citizens’ rights from being violated, but those who do not respect others’ right to life, freedom or possession lose (or limit) their own right to these goods. Therefore, the state does not have a *prima facie* obligation to protect offenders. The conclusions Ross draws are incompatible with retributivism; also, the fluid character of the principle of justice, as it were (its *prima facie* character), seems to contradict modern retributivist thought inspired by Kant’s ethics. Ross states that society’s interest may be great enough to justify the right to punish an innocent individual, so as to prevent the destruction of the whole nation (Ross, 1930: 56-64).

Ross’ theory, from the point of view of retributivism, is only a kind of guideline. The essence of retributivism is its clear opposition to sacrificing an individual for the common good as principle of criminal responsibility. Otherwise, retributivism would become a supplement to utilitarianism reduced to the requirement to consider guilt a premise of responsibility and moral condemnation of crimes, while accepting the possibility to make exceptions for the sake of special considerations, such as public interest.

Retributivism claims that offenders ought to be punished because of the demands of justice - because they deserve to be punished. The modification would consist in attributing the *prima facie* character not to the duty to punish the offender but to the duty to serve justice. The duty to serve justice is a *prima facie* duty, which should be understood in a very specific way.
First, it means that there is no norm (moral, legal, natural etc.) that would prohibit punishing the offender in the form of repayment for the wrong he or she had done, and that there is no norm that would obligate to accomplish any objectives through punishment.

Second, the state has the duty to protect the innocent. This is, however, an optimization norm, which can be fulfilled only to some extent. It does not mean that the state may punish the innocent in the interest of society (i.e. there is no norm that prohibits punishing the innocent as a means to a goal or, all the more so, a duty to do so); but, it ought to do everything in its power not to punish the innocent. If perceived as a gradual achievement, this goal cannot be fully achieved or not achieved at all. The duty to punish the guilty also has a *prima facie* character.

Third, proportionality of the punishment to the guilt (severity of the offence) is also an optimization norm (directive), a certain principle of punishment.

These three elements, taken together, characterize retributivism as a *prima facie* duty to administer justice. The first one is fundamental, however, as it defines retributivism in a negative way. They outline the core of retributivism in relation to utilitarianism or penal abolitionism. The claim that there is no norm prohibiting criminal punishment as repayment for a wrong, nor a norm requiring that instrumental results be obtained through punishment, is in opposition to the programme of penal utilitarianism, which sees punishment as a means to certain social or personal ends, as prevention, resocialization etc. (Kaufman, 1960: 49-53). It is also a conclusion incompatible with the abolitionist or minimalist approach, which sees punishment as conditionally permitted or prohibited for moral reasons (Christie, 1977:1-15; Peno, 2016: 28-38). They do not provide an answer to the question about the reason for punishing; yet, if we eliminate the utilitarian value, justice considerations stand out from among other possibilities. Supplementing the acceptance of punishment as repayment for an offence are three farther duties, which are optimization norms, meaning that they do not have a fundamental character. Rather, they must be performed in the highest possible degree. Violating them is not permissible as a matter of principle.

The vision of *prima facie* duty depends on interpretation. *Prima facie* duty is only an apparent but not a real duty; or a real duty that can be outweighed by more stringent considerations but continues to survive even when outweighed. The second interpretation seems to be better suited to the retributivism idea of punishment. There is a real obligation or duty to punish, but there is not the obligation that can be either fulfilled or not. It can be assumed that the society or the state has the obligation to do as much as possible to punish only the offenders and to protect the innocent.

One can hardly miss the link between the outline of criminal justice presented above and Robert Alexy’s understanding of principles. According to Alexy, principles are optimization requirements. They can be fulfilled only to some extent (See: Alexy, 2000:294-304). This is also how a principle of retributivism can be formulated. It would assert that the state ought to realize justice in the highest possible degree. It seems that this approach is not only compatible with the essence of retributivism but also remains in opposition to penal utilitarianism.
CONCLUSION

How does the outlined theory solve the problem of punishing the innocent? Accidental punishing of an innocent individual results in improving the criminal justice system as a mechanism functioning in a certain state or society. The argument that just repayment does not exist, because in reality punishment is an outcome of many factors differentiating the situation of individuals in the same position, does not weaken the outlined retributive concept because the only thing it entails is that the rule to punish only the guilty should be respected as much as possible, and mechanisms helping the pursuit of this goal should be created. Thus, the presence of injustice in the relationship between the state and the citizen becomes apparent. Yet, It is not an argument against retributivism per se.

The retributive character of the outlined concept is manifested by the three following characteristics (Cf. Radzik, 2017: 164): first, the assertion that punishment can only constitute repayment (which ought to be understood in a normative sense); second, justice remains a value that must be pursued in the highest possible degree; a just repayment is thus the value and the goal that ought to be pursued; third, the assertion that punishment should not constitute (only) a means to achieving certain goals and should not be inappropriate to the degree of guilt and the severity of the offence.

It seems that retributivism, having undergone various changes, has gradually adopted a mixed form, essentially combining utilitarianism with certain justice-oriented considerations. It is therefore worthwhile to consider such approaches to retributivism which, while modifying the classical, formal and fundamental (and thus contrafactual) idea of justice, remain in acute opposition to the utilitarian programme in criminal law. The conception outlined here embraces this idea of just punishment.

REFERENCES

Christie, N. (1977). Conflicts as Property” British Journal of Criminology 17(1)

**PRIMA FACIE RETRIBUTIVIZAM: O DUŽNOSTI SPROVODENJA PRAVDE**


Ključne reči: V.D. Ros, unutrašnja struktura retributivizma, prima facie dužnosti, kažnjavanje nevinih, krivičnoprawne teorije.