

## Contents

<b>Preface. Current Issues in Ethics</b> ( <i>Magdalena Holy-Luczaj</i> ).....	3
<b>An Analysis of the Equal Freedom</b> ( <i>Andrzej Stoiński</i> ).....	5
<b>Selfishness and Cooperation: Challenge for Social Life</b> ( <i>Konrad Szocik, Stig Lindberg</i> ).....	15
<b>Selected Ethical Issues in Artificial Intelligence, Autonomous System Development and Large Data Set Processing</b> ( <i>Paweł Zgrzebnicki</i> ).....	24
<b>Artifacts and The Problem of Ethical Extensionism – Selected Issues</b> ( <i>Magdalena Holy Luczaj</i> ).....	34
<b>Salutary Meanings of Sublimation. Selected Soteriological Threads of Alienation Criticism of Religion</b> ( <i>Jowita Guja</i> ).....	44

## **Preface. Current Issues in Ethics**

*Magdalena Holy-Luczaj*

University of Information Technology  
and Management in Rzeszow,  
Poland

*e-mail:* mholy@wsiz.rzeszow.pl

### *Abstract:*

This interdisciplinary volume consist of papers on various problems in contemporary ethics. It presents the following issues: equalizing the level of positive liberty, the phenomenon of human cooperation, ethical questions related to artificial intelligence, extending ethical obligations toward artifacts, and soteriological threads of alienation criticism of religion.

*Keywords:* ethics, new trends, moral relevance, moral considerability.

The major task of ethics is to address questions related to moral dimensions of the most pressing problems in today's world. New controversies arise mainly due to the development of science and technology. We face dilemmas which past generations didn't have to deal with. Thus, we have to investigate new problems that are to be identified at the crossroad of ethics and bio, computer and industrial science, as well as technical knowledge. This requires also use the results of empirical research to lay out new arguments on the grounds of philosophy. Furthermore, our approach to ethical problems is shifting from strongly anthropocentric to posthuman and non-anthropocentric. In result, ethicists reframe the question of who belongs inside a circle of moral relevance. The novelty of research in ethics, however, does not mean that scholars turned away from the heritage of history of philosophy. On contrary, they seek inspirations in classic theories, intending to revisit them and adapt to current problems.

This broad scope of considerations in ethics as well as a variety of problems and new approaches is reflected in contributions to this volume of *Studia Humana*. The first article, "An Analysis of the Equal Freedom", investigates selected issues related to the postulates of equalizing the level of positive liberty. Its author, Andrzej Stoiński, reminds us that classic understanding of individual freedom referred to as a negative ("freedom from"), identified with a lack of compulsion, is in the opposition to the so-called positive liberty ("freedom to"). The latter is usually defined as an ability related to the concept of power. The postulate of equality in "freedom to" can justify conducting a social redistribution of goods. Thus, Stoiński examines the cases of voluntary and compulsory donation in order to present consequences of a compulsory expansion of the scope of positive liberty.

In the next paper, "Selfishness and Cooperation: Challenge for Social Life", Konrad Szocik and Stig Lindberg argue that cooperation is a great challenge for natural selection. Some scholars assume that cooperation could not evolve within the framework of natural selection. It is undeniable

that natural selection, at least at the individual level, favors selfishness and defectors. Nonetheless, this selfish tendency does not necessarily imply that cooperation could not evolve by means of natural selection. Drawing upon this assumption, Szocik and Lindberg specifically acknowledge certain basic challenges for the evolution of the human ability to cooperate at the level of large groups. They discuss topics such as the human ability for “super cooperation”, the importance of repetition and reputation, and Multilevel Selection Theory as the basic mechanisms of evolution of cooperation.

The paper “Selected Ethical Issues in Artificial Intelligence, Autonomous System Development and Large Data Set Processing” deals with the dilemmas that due to development of information technology and its industrial adaptation, have become the real problems of the contemporary world and not only the ground for the plot of science-fiction literature and films. Paweł Zgrzebnicki shows us that, on one hand, these issues are related to an unprecedented scale on which computational algorithms are currently used as well as a level of complexity of mutual connections; on the other hand, these are linked to their autonomous behavior. States, industry, and users themselves demand formulation of understandable ethical categories and determination of transparency standards and legal norms for these algorithms’ functioning in the near future. For this reason, he focused on three areas: determining acceptable goals that can be pursued by an autonomous algorithm or a device using this algorithm, understanding the decisions made by artificial intelligence objects and making supervision over their actions possible.

The paper “Artifacts and The Problem of Ethical Extensionism – Selected Issues” addresses the problem of extending ethical obligations toward material things, which we use every day. In the first part I – honored to edit this volume of *Studia Humana* – reconstruct current approaches to the problem of the metaphysical and ethical status of artifacts. Next, I intend to unpack the suppositions that are steering the debate on moral status and expanding ethical obligations of human beings toward nonhumans that has been going on for several decades. Arguments for including different animal species, plants, other living organisms, and finally inanimate entities in the boundaries of ethics have successively appeared in numerous works dealing with this issue. Now it is worth establishing possible grounds for extending human moral obligations toward artifacts as beings, which status can be improved or deteriorated through human activity. The core part of the article analyzes ethical implications of the following issues: using artifacts, their production, purchase, and sale of artifacts, and the post-use period.

The last paper by Jowita Guja, “Salutary meanings of sublimation. Selected soteriological threads of alienation criticism of religion”, look into a few soteriological threads of alienation criticism of religion, whose feature is the creation of a new autonomous and transgressive subject. Guja has focused on the presentation of this subject using Nietzsche's philosophy perceived from a Freudian point of view. According to her, within this framework, our choices can be described as authentic if they are ultimately determined by a degree of our involvement in them and by our awareness of responsibility. The representatives of the alienation perspective more and more often write about a necessity of real, secular faith whose verification would be an explicit proclamation and in the first place acting in accordance with it. Thus, she concludes, our authenticity can only be proved by our practice.

I wish to thank all of the contributors for their outstanding papers, which, I believe, will be useful and inspiring for our readers.

## An Analysis of the Equal Freedom

Andrzej Stoiński

University of Warmia and Mazury  
in Olsztyn, Poland

*e-mail:* andrzej.stoinski@gmail.com

### *Abstract:*

The article concerns selected problems related to the postulates of equalizing the level of positive liberty. The classic understanding of individual freedom, called as negative (freedom from), identified with a *lack of compulsion*, can be in opposition to the so-called positive liberty (freedom to). The last notion is generally defined by an *ability*, which brings its relation with a concept of power. The postulate of equality in “freedom to” can be justification for conducting a social redistribution of goods. The cases of voluntary and compulsory donation are considered in the text, whose aim is to visualize consequences resulting from a compulsory expansion of the scope of positive liberty.

*Keywords:* political philosophy, ethics, freedom, positive liberty, negative liberty.

## 1. Introduction

Freedom occurs in various contexts. We can consider political, economic liberty and freedom of speech, freedom of the will, etc. However, in this text, we will concentrate our attention to liberty called as an individual. We will deal with a problem, which can occur in a case when a postulate of popularising an equality of positive liberty is processed through the state. We will make an attempt to indicate the fact that some contexts, in which “freedom to” takes place can incline us to accept the thesis that striving for materializing equality in it, in some circumstances, can result in reducing both positive and negative liberty, at least to some persons. However, it can have an impact not only in the field of economy or politics but also can have significant moral consequences.

## 2. Negative and Positive Liberty

Before we move on to main issues, we will focus our attention to present two types of individual liberty. The differentiation into positive and negative liberty is mostly related to the name of Isaiah Berlin [4]. In fact, this differentiation is older as it and originates from T. H. Green [14]. Negative liberty, which is also called “freedom from” is generally defined as the absence of external constraint or interference, whereas positive liberty (in other words “freedom to”) is rather perceived as self-mastery or self-realization [15]. Negative liberty, in compliance with its name, is thus based

on the lack of external (we can add *unjustified*) compulsion [18]. However, positive liberty is expressed in the possibility of action or also having the *ability to do* or *possession of open options*.

We can have the impression at first glance that negative liberty constitutes the basis of positive liberty that it is its necessary condition [19].<sup>1</sup> However, it is not this way. A basic phenomenon, around which the sense of a notion of negative liberty is established, is *lack of*. It is the lack of a compulsion, which is not justified by protecting against this constraint. *Ability* is such a phenomenon in the case of positive liberty. However, one is not a necessary condition of the other. The *ability to do* can be fulfilled irrespective of the fact, there is or there isn't external pressure placed on us. The absence of negative liberty does not mean the absence of ability at the same time and the lack of ability does not mean lack of negative liberty at the same time. In order to illustrate this statement, it is sufficient to indicate the fact that our *possibility* to vote for the parliament does not depend on the *lack of external compulsion*. "Freedom to" vote will be fulfilled in the equally good way in the case when we will not be forced by some external factor and also if we are pressured to vote by force or a threat of its use.

While negative liberty does not need material fulfillment to be processed, then positive liberty depends on possessed some kind of resources to some extent. It is not only dependent on a pay of a given person, however, we can also take it as gradable, because these persons, who dispose of greater resources, also have greater possibilities. Therefore, this means that they have a greater scope of positive liberty. Economic prosperity is an aspect of possessing resources, which is taken into account in the majority of cases. However, such approach has significant disadvantages. The term *resources* can be understood in many ways. It can be understood not only as a material pay of a given entity, but also physical, psychic or other properties, which decide about his or her possibilities [24, pp. 169–221].

When looking from this perspective, out of two persons of the same sex and age, the same health condition, the same height and the same physical fitness, this person who has greater knowledge, experience, intelligence or beauty, will have considerably greater capabilities than the other one. The problem will occur in the case when we will make an attempt to compare various categories, in which these properties occur. What does constitute greater capabilities of two comparable persons: advantage in the form of intelligence or beauty, admired features of character or extensiveness of knowledge? As long as comparable persons are not in a definite situation in a specific environment and saying in other words, do not find themselves in a definite context, their capabilities seem to be incomparable. Even so, as it could seem, a clear advantage, which is a higher level of prosperity of one person, can in some specific and extreme circumstances indicate a weakness rather than an *ability*.<sup>2</sup> When drawing conclusions from this situation, we should refer to the notion of *contextual capabilities* or *contextual advantage*. When concluding this topic we can then say that positive liberty (the ability of action) depends on the context of resources [1],<sup>3</sup> which gain values due to a specific situation, in which a given entity is placed.

### 3. Equality in Freedom

The differences between both concepts of liberty in the perspective of our discussion, will be clearer when we compare them to the postulate of spread out equality in freedom [13, pp. 169–171], [9, p. 45–47].<sup>4</sup> Then it turns out that the consequences of making common liberty from compulsion will be different than popularising liberty as the *ability to do*. Fulfilment of the requirement of the lack of compulsion in the social sphere, will bring the same state of the lack of compulsion of each person in relation to all other persons [27], [17, p. 231]. The situation will look differently, if analogical actions will be concentrated on equalizing abilities to do. The cause of such a state of affairs is the fact that persons have been equipped with two possibilities with various level: *ability to do* or *possession of open options*. As we have mentioned earlier, these differences are in general based on disproportions in possessing properties, which are the source of the *ability to do*. The contrasts between persons can come from the differences in disposing of broadly understood

resources, such as even: intelligence, talents, and features of character, physical fitness, skills and even beauty. Significant is not only the number of these properties, but also their indivisibility.

If a superior aim of action will be indicated as aspiration to equalize “freedom to”, then we can imagine acting of a state, which will be directed to equalise these *abilities to do* by taking the capabilities away from these persons, who have more of them than the other ones and giving them to these, who do not have resources in such a way that *abilities to do* of everybody, resulting from the fact of possessing resources, will have a tendency to go in the direction of equality. The second manner of operating is processed by applying a derivative division of goods of economical character as an intention to change them for other capabilities, within the framework of the society treated as the whole. This solution is based, above all, on social redistribution of property, including primarily incomes. In this case, the declared aim is to: increase the level of abilities of these persons, who are in some way handicapped. In practice, it goes to transferring a part of goods that have been worked out or possessed by the society. However, in general, the mentioned goods do not constitute the property of the whole society, but they become it as a result of their appropriation. This process is conducted in the manner that some part of the possessed property is taken away from some categories of citizens defined by law, which will later supply completion of projects directed to support the *ability to do* of the other categories of persons. However, we should notice that the process of this type is based on using force in social life [26, pp. 89–115]. Therefore, it naturally seems to have an impact on the negative liberty of society members. Does it happen so in the real world?

We will make an attempt to visualize this problem by referring to the example, in which we will compare two different manners of transferring goods. We will use an example of donation here, which according to us provides important features of the whole problem. A voluntary donation will be the first one, whereas the second one will be a donation made under compulsion. On the basis of the analysis of these cases, we will make an attempt to define possible changes in the scope of liberty, which can take place in relation systems created in a donation process. If it turns out that there are differences in the scopes of liberty, the next step will be to try to compare a significance of both liberty types. In the further part we will try to interpret both of these situations in the perspective of their importance for the possibilities of fulfilling by persons their humanity. It will be, above all, an attempt to define, which influence of redistribution of properties is made on the ability of persons to initiate acts of moral nature.

#### **4. Thought Experiment of *Donation***

A donation is a relation made between three types of entities: a donor, an agent and a donee. Before we start to make a more detailed analysis, we will refer to the status of entities before a donation. It is an initial situation, in which there are three entities that are not related to each other with any relation. We assume that each of them shall be entitled to the same negative liberty and some differentiated value (level) of positive liberty. A future donor disposes of “freedom to” with a higher level than a future donee. Disparities, which take place between them, in the level of this disposition are not significant, as the object of our interest will be the differences between an initial situation and a final one of individual entities.

A free donation is making a relation between entities and is based on the non-compulsory transfer of means by a donor, collected by an agent (for example a charity organization) [25], [20, pp. 393–394], which then disposes of it by supporting a donee. In the scheme of this relation, a donor is somebody who initiates action in this case. His act will be repeated by an agent. The situation is totally different in a case of compulsory donation. Three elements also take part in it, however, their mutual relations are different than the previous ones. The party which initiates an action is an agent in this case, which forces a donor to give away a part of possessed resources. Then it is passed to a recipient.

Let us try to interpret the aforementioned relations now in the light of a question about consequences, which are brought for the liberty of its members. We take a simplified assumption in

this case and also in the whole work that the transferred goods are freely exchanged for capabilities and the last ones reflect the scope of positive liberty. The first point of reference will be the initial situation before a donation. When looking for the whole idea from the perspective of liberty, the final result shows extensive changes in relation to the initial point.

In conditions of voluntary donation, a donor has resigned from a part of his abilities (and despite this fact and maybe thanks to it, he fulfilled his positive liberty) and strengthened this type of liberty in a donee. A donee has had his positive freedom expanded. Moreover, an agent has received, in relation to a status, which has been before a donation, an additional aspect of positive liberty that is based on the ability to transfer means from a donor to a donee. Let us also notice at the same time that none of the entities has sustained a loss in the field of negative liberty, which he or she is entitled to. However, if we take into account compulsory donation, the situation will look otherwise.

Positive freedom of a donee has been increased in this situation in relation to the original one, whereas its level has been decreased in the case of a donor. Therefore, it seems that there are no differences in relation to the characteristics of voluntary donation in this respect. However, a change has taken place here and it is of vital importance. While in the case of a voluntary donation a donor fulfils his or her positive liberty, they use it in some sense, by decreasing their possibilities by such action (we use money for the same purpose by spending them, which does not mean that we lose it, as we simply change it and materialise its value), so the same possibilities are taken away from a donor in the case of a compulsory donation. He does not show his abilities in this case. He does not do this, which expresses his will, but acts in a way that is in compliance with the will of the other people. In the case of an agent, a change has taken place in relation to a state before a donation and towards a voluntary donation. It has revealed an additional scope of positive liberty, which is based on the possibility to transfer means from a donor to a donee, but also the ability to force a definite action on a donor, which has not seen before. This special ability to force actions on other entities is called power in other words<sup>5</sup>. It means that an agent has power over a donor.

As regards the scope of negative liberty, we need to acknowledge that its scope in a donor has been limited both in relation to the original situation and also in comparison to a voluntary donation. The other entities, which participate in a relation, will not record any changes in this respect.

When we look at relations between entities as a certain closed system, then we can attempt to evaluate an increase and decrease of both types of liberty in the perspective of the whole. Thus, the whole system before a donation has been in a state, which has later changes together with donations. A voluntary donation has resulted in passing an aspect of positive freedom to a donee. It has co-existed with its fulfillment in a donor at the same time. As we have already indicated, such materialization is not identical to a *loss*, but rather an *exchange*. The scope of this type of liberty in an agent has increased in the same way. As a result, our interpretation goes in the direction of an ascertainment that a voluntary donation has caused an increase of a positive liberty in the whole system, with maintaining the same status of negative liberty.

The consequences of a compulsory donation seem to be significantly different. A transfer of positive liberty from a donor to a donee has taken place within the framework of the system, in which it has been taken from a donor, but not fulfilled at the same time. Moreover, an agent has gained one more ability – power – in comparison to a voluntary donation. In conclusion, negative liberty had decreased for the whole relation system with reference to the state from before a donation but positive liberty increased. However, the level of the last one has increased as a result of the transfer between a donor and a donee, as a part, which has been taken away from a donor, has been given to a donee,<sup>6</sup> therefore, its value has remained the same for the whole system. The status of positive freedom has increased in the system by making a relation of an agent who has power over a donor. Therefore, an occurrence of a phenomenon of power is an additional aspect of positive freedom in this case. However, the scale of negative liberty has decreased for the whole system in compulsory donation in relation to voluntary donation, whereas positive liberty has

remained at the same level (adding the ability to do in the form of power of an agent has been balanced by a loss of ability to do of a donor as a result of its fulfilment).

## 5. Consequences of Equal Liberties

In the case of equalizing differences in “freedom to” that is implementing the procedure of equalizing *abilities to do* in society in practice, it should be expected that the maximum positive liberty of some entities, coming from the fact of possessing capabilities by them, will be considerably reduced. A transfer of resources could not otherwise take place, which increases a status of positive liberty of other entities that have less. Such actions will definitely have an impact on the negative liberty of an entity which has the greatest capabilities. Then it will have to lead to the annihilation of the postulate of equal “freedom from”. Therefore, it seems that equality in negative liberty and aiming to equalize positive liberty cannot be reconciled. William Dustin claims, that clashes between representatives of these views „represent not conflicts of style, but vastly different moral orders” [10, p. 33]. One of these postulates must give place to the other one.

In the light of the aforementioned information, we still can attempt to answer a question related to the significance of both postulates. Which one of them is more important and/or consequences of which one of these solutions we are more inclined to accept. The case of compulsory donation, which has been described above, has shown how a transfer of liberty looks like within the framework of the whole relation system. The increase of positive liberty has finally taken place there, as well as the decrease of negative liberty compared to the status from before a donation, the decrease of negative liberty and the lack of a change in the scope of positive liberty with reference to voluntary donation. Therefore, it is worth to consider, if the increase of “freedom to” compensates, for the whole system, the decrease of the level of “freedom from”. The answer to this theoretical question would help us understand the significance of both types of freedom in real social life.

## 6. Imaginary Situation of Competition

If we would like to consider this problem, then we can refer to an imaginary situation again. Let us rethink, which equality in liberty we should value more and which one we can sacrifice, if it is not possible to combine both of them. In order to study this problem, we will concentrate on the situation, in which we assume the same context of capabilities for two entities (*A* and *B*) as well as different level of them *ceteris paribus*.

1. In the first case, both entities are equally independent of an external unjustified compulsion. The entity *A* possesses considerably greater capabilities than the entity *B*. Therefore, *A* disposes of bigger possibilities, so it has greater positive liberty than *B* has.

2. In the second example *A* and *B* dispose of equal scopes of positive liberty (they have the same capabilities), however, *B* possesses a considerable advantage in the scope of negative liberty. It means that *A* is not only under greater risk of external unjustified compulsion (unjustified protection from a compulsion), but is also subject to an external limiting compulsion to a greater extent than *B* is.

3. The last example presents a situation, in which *A* has greater positive liberty (disposes of greater capabilities) than *B* and is at the same time hindered towards the other one in the scope of negative liberty. *A* has less negative freedom than *B*.

Now we will consider three aforementioned cases for two circumstances. Firstly, the relation of competition between entities – zero-sum interaction; secondly, for the relation characterized by the lack of competition between entities – non-zero-sum interaction. We will understand the zero-sum interaction in a special manner here, namely that two entities in a mutual relation cannot win at the same time. If one of them wins, then the other one needs to lose. In contrast to this situation, the non-zero-sum interaction (in this specific case, we take into account a positive-sum), which will be defined here as the one, in which each party (at least subjectively) acknowledges that the party has



gained more on the transaction than lost from it.<sup>7</sup> Let us make an attempt a situation by placing it in a field of economy. Two competing sellers or two competing buyers will take part in the first relation. The second relation describes the occurrence, which will happen between a seller and a buyer.

Competition – interaction with zero-sum.<sup>8</sup> The situation of an auction will be taken as an example.

Re 1. The buyer *A* has greater abilities (more money) than buyer *B*. The external conditions for both of them are the same, then we should expect that *A* will win.

Re 2. Both buyers have the same abilities (the same amount of money), but *A* can be excluded from the auction, for example by intimidation or implementing a ban for participating in the auction for some categories of buyers. Therefore, *A* is subject to discrimination. *B* will win.

Re 3. The buyer *A* has greater possibilities than *B*, but can also be excluded from participating in the auction (*A* is subject to discrimination). Therefore *B* will win again.

The case no.1 is most similar to sport situations. Despite the lack of equality in positive liberty, the equality in negative liberty (the same rules for all participants) results in the situation that a victory seems to be honest. The cases 2 and 3 confirm the situation, which has been earlier indicated by an example of compulsory donation. Despite the fact of possessing capabilities in theory, fulfillment of positive liberty does not take place due to interference in the independence of an entity (decreasing the scope of negative liberty). Therefore, depreciation of negative liberty can also result in reducing the positive liberty of these entities, for which negative liberty has not been infringed. Negative liberty in entity *A* has been reduced in both of these cases, which has also become a barrier for fulfilling its positive liberty.

No competition – interaction with non-zero-sum. We will consider a relation of a seller – a buyer in this case.<sup>9</sup>

Re 1. The seller *A* that possesses more capabilities, proposes his product *x* to the buyer *B*. The last one is neither forced to its purchase nor is it made impossible for him.

Re 2. Two possibilities are differentiated in this case:

- a) The seller is *A* that has the same resources as *B*, yet is subject to external compulsion to sell *x*, for example at a reduced price;
- b) The seller is *B* that has the same resources as *A*, yet can be forced to buy *x* at a higher price or to buy it in general, for example as an indispensable element of equipment of the product *y* offered by *B*.

Re 3. The same case as the previous one:

- a) The seller *A*, despite the fact of possessing greater capabilities than the buyer *B*, is forced to sell *x* to *B* at a lower price;
- b) The buyer *A*, despite the fact of possessing greater resources than the buyer *B*, is forced to buy *x* at a higher price.

The relations reviewed here also show that until positive liberty of some does not have an impact on the negative liberty of the other ones. The relations seem to be honest, despite the fact of a difference in capabilities. The next cases in the example of interactions with non-zero-sum value show something else. The first case (of equality in negative freedom) remains an interaction of positive sum. Both parties of the relation have an open way to fulfill positive liberty, which has been granted to them, while maintaining negative liberty at the same time. The other cases fall into the category of interactions with zero-sum, in which one party must lose, so that the other one can win.

However, the provided examples raise a significant doubt. Does inequality of capabilities (positive liberty) really lead to fair competition? The example of sport, in which abilities of competitors almost always are unequal seems to indicate the affirmative answer to this question should. Our analysis is made difficult by the fact that real equality of capabilities in conditions of functioning of the society seems not only to be incomparable, but also unachievable. As we have already highlighted, equality of capabilities can only be contextual. However, persons are found in this meaning in various contexts at the same time.

Now back to the considered postulate of equal positive liberty. It seems that it should be acknowledged one important aspect. Even if we set a goal to lead to equalize everyone in an arbitrarily selected context, then we should still expect that we will be able to achieve it at the price of actual discrimination of at least some persons. When we write about postulates of equality in positive liberty, which exclude each other, we should also add some more information here. Thus, the postulate of equal positive liberty can have an impact on equal negative liberty. In contrast, the postulate of negative liberty does not infringe positive liberty. The requirement of common equalizing of the state of lack of compulsion does not change and therefore, does not infringe capabilities of anyone.

## 7. Conclusion

What is the significance of the aforementioned deliberations related to donations and the importance of both types of liberty to justify the postulate of popularising equality in positive liberty? Such postulate in politics of many contemporary countries is fulfilled in the form of social redistribution of resources. It seems to be analogical to a situation of compulsory donation, which has been presented here. However, if such a situation takes place, its consequences for freedom of both types should be taken into account. The topic of the greatest interest for us is not a political but ethical aspect of the problem. The problem of moral aspects of redistribution can be important for that as human beings are treated.

It is commonly acknowledged and probably justifiable that out of all known creatures, a human is somebody special.<sup>10</sup> We name the others, who are similar to us as *someone*, and not *something*. Why is that? Various answers were provided to this question. We will refer in this case to some remark of Aristotle that a property that differentiates a human is his ability to discern good from evil and justice from injustice [2]. Not the ability to discern itself and to think in moral categories seems to be the only important one. It is also the fact that people have the ability to act in a moral manner. Persons are active entities, who discern just from unjust acting and can knowingly behave in a just or unjust manner. Therefore, the most significant factor seems to be the ability to make a choice in the light of possible options.<sup>11</sup>

Immanuel Kant pays attention to this property in *Groundwork of the Metaphysics of Morals* by making freedom as the condition of morality in general [16]. The freedom that Kant talks about, is understood as the possibility of acting in this or another way. When looking from our perspective, we should interpret it as positive freedom. This freedom is nothing else but an ability to act. The perfect condition for the fulfillment of the fullness of humanity would be the state of the greatest possible freedom of making decisions for all persons. It would be such a state, in which people will not necessarily make the greatest consequence of actions, but it would be the state, in which acting would be the most related to the notion of the moral duty. It means that the greatest possible number of acts would depend on free decisions of individual persons.

It is natural that the situation, in which there would be the maximum scale of acting of persons that discern good from evil and justice from injustice does not necessarily mean that these persons would prefer good and just actions as well as that they would select these acts which refer to them. According to Nozick, it doesn't mean yet that the state should violate human freedom: "The state may not use its coercive apparatus for the purpose of getting some citizens to aid others" [23, p. ix]. What would in such a case constitute a barrier preventing from bad and unjust actions of maximally free persons? The answer in this case of justice is certainly securing the maximum negative liberty and the greatest possible level of equality in it. Then we will have two different states in this case. Firstly, with maximum freedom and secondly with equal freedom. We should highlight their difference here. The maximum negative liberty in real social conditions is such a situation of a given entity, in which this entity is hampered with limitation to the lowest extent. However, equal negative freedom is characterized by the same scope of limitations for everybody. It neither limits nor has an impact on a change of a positive freedom status.

We should pay attention to a paradox of positive liberty. Both, the greatest positive liberty and also equal positive liberty are obtained as a result of actions, of which, depreciation of negative liberty can be expected, at least in some entities. The postulate of equalizing of positive liberty can also lead to a situation which can be evaluated as morally destructive. Considering a compulsory donation, it firstly can lead to limiting the ability to act in a moral way by these persons, whose property has been taken away. Secondly, we should place a donee in a situation that is morally ambiguous, as he profits from taking away of someone's property. The situation of a donee seems to be similar to a situation of a person, who has received a present coming from a theft. Finally and thirdly, an agent becomes a person who deprives persons of their moral subjectivity as he treats them as objects, means that are used to achieve his own aims.

In this text we have basically concentrated on some consequences of the postulate of equal freedom. This proposal refers to property redistribution in practice. Therefore, we cannot avoid a question related to an admissible scope of this redistribution. As it seems, the conclusions resulting from the presented examples should incline us to approve the thesis that a redistribution, which aims at equalizing scopes of positive liberty has its not extendible limit. It is a violation of equal negative liberty of entities that are full society members.<sup>12</sup> Thus, each it's exceeding results in reducing the independence of persons. However, it turns out that it can lead to reducing their real abilities. Therefore, fulfillment of the postulate of equal abilities can have an impact not only on freedom from compulsion, but also in fulfillment of positive liberty in some persons. What is more, some forms "freedom for" can generate occurrence and increase of power relations.

Apart from losing currently possessed possibilities and liberties by some persons, another consequence can also occur in the form of limiting their properties as a moral creature. As they have been deprived of a part of their current negative and positive liberty, which they have been entitled to earlier and as a result they have also been limited in their moral duties that are contingent upon these liberties. Their moral autonomy for this reason depreciate. However, this can result in the reification of these persons. As a consequence, it can happen that they will be reduced from the level of *somebody* to *something*. The history tells us that depriving selected groups their autonomy, has born the bitterest fruits. Therefore, it seems that the postulates the implementation of equality in positive liberty coerced by state institutions, should be treated with restraint and with special prudence.

## References

1. Alexander, J. M. *Capabilities and Social Justice: The Political Philosophy of Amartya Sen and Martha Nussbaum*, Aldershot: Ashgate Publishing, 2008.
2. Aristotle. *The Politics*, Indianapolis: Hackett Publishing, 1998.
3. Ayau, M. F. Property Rights and the General Theory of Exchange: Frequently Neglected Implications of the Division of Labour and the Law of Comparative Advantage, *Economic Affairs*, 26. No. 1, 2006, pp. 48–53.
4. Berlin, I. Two Concepts of Liberty, In *Four Essays on Liberty*, Oxford: Oxford University Press, 1969.
5. Boethius. *Theological Tractates*, London: Heinemann, 1973.
6. Brenkert, G. G. *Political Freedom*, London: Routledge, 2006.
7. Ciulla, J. B. Worthy work and Bowie's Kantian theory of meaningful work. In D. G. Arnold & J. D. Harris (eds.), *Kantian Business Ethics: Critical Perspectives*, Northampton: Edward Elgar Publishing, 2012.
8. Darwall, S. L. (ed.). *Equal Freedom: Selected Tanner Lectures on Human Values*, Ann Arbor: University of Michigan Press, 1995.
9. Denier, Y. *Efficiency, Justice and Care: Philosophical Reflections on Scarcity in Health Care*, Dordrecht: Springer Science & Business Media, 2007.
10. Dustin, W. *Toward an Ethic of Citizenship: Creating a Culture of Democracy for the 21st Century*, New York: toExel Press, 2000.

11. Epstein, R. A. The Rise of Transfer Payments. In G. C. Bryner & N. B. Reynolds (eds.), *Constitutionalism and Rights*, Provo: Brigham Young University Press, 1987.
12. Gołaszewska, M. Roman Ingarden's Moral Philosophy. In A. Tymieniecka (ed.), *Ingardeniana*, Dordrecht: Reidel, 1976.
13. Gould, C. C. *Rethinking Democracy: Freedom and Social Co-operation in Politics, Economy, and Society*, Cambridge: Cambridge University Press, 1990.
14. Green, T. H. Liberal Legislation and Freedom of Contract. In R. L. Nettleship (ed.), *Works of Thomas Hill Green*, vol. 3, London: Longmans Green, 1991.
15. Hayek, F. A. von. *The Constitution of Liberty*, London: Routledge, 1960.
16. Kant, I. *Practical Philosophy*, M. J. Gregor (ed. and trans.), Cambridge: Cambridge University Press, 1996.
17. Kant, I. *The Metaphysics of Morals*, Cambridge: Cambridge University Press, 1991.
18. Locke J. *Two Treatises on Government*, P. Abrams (ed.), Cambridge: Cambridge University Press, 1967.
19. MacCallum, G. C. Jr. Negative and Positive Freedom, *Philosophical Review* 76, 1967, 312–344.
20. Macey, J. R. On the Failure of Libertarianism to Capture the Popular Imagination. In E. Frankel Paul, F. D. Miller, & J. Paul (eds.), *Problems of Market Liberalism: Volume 15, Social Philosophy and Policy*, Cambridge: Cambridge University Press, 1998.
21. Narveson, J. *Respecting Persons in Theory and Practice*, Boston: Rowman and Littlefield, 2002.
22. Nolan, P. *Capitalism and Freedom: The Contradictory Character of Globalization*, London: Anthem Press, 2007.
23. Nozick, R. *Anarchy, State, and Utopia*, New York: Basic Books, 1974.
24. Sen, A. Well-being, Agency and Freedom. *Journal of Philosophy* 82. No 4, 1985, 169–221.
25. Tanner, M. *The End of Welfare: Fighting Poverty in the Civil Society*, Washington: Cato Institute, 1996.
26. Vaubel, R. Socialism. Do the Ends Justify the Means. In S. Pejovich (ed.), *Socialism: Institutional, Philosophical and Economic Issues*, Dordrecht: Springer Science & Business Media, 2012.
27. Weinstein, D. *Equal Freedom and Utility: Herbert Spencer's Liberal Utilitarianism*, Cambridge: Cambridge University Press, 2006.
28. Wright, R. *Nonzero: History, Evolution & Human Cooperation*, London: Abacus, 2001.

## Notes

- 
1. For example, MacCallum assumes that the two kinds of the freedom are indivisible elements of one liberty, and the dichotomy of them is false [19, pp. 312–334].
  2. It is sufficient to imagine two men in the prime of their life: a wealthy and a poor one, which have found themselves in the storm of the bolshevik revolution in Russia to realize relatively the notion of *greater capabilities*. As in this case and this context, greater material resources of the first person, make his situation worse at the same time, in which he is under a threat of persecution from the others, who would like to make a new order.
  3. About connections between freedom, resources and abilities [6, pp. 236–242].
  4. Wider scope of equal liberty from different points of view see [8].
  5. Richard Epstein compare government holding this kind of power to Robin Hood “taking from the rich and giving to the poor. No one disputes that individuals are entitled to make voluntary charitable contributions. But it is a different matter when some people try to fund their gifts with cash taken from their neighbor's pockets” [11, p. 34].
  6. The natural evaluation of *the ability to do* depends on a broader context. A significant impact on it will undoubtedly have the level of usability of transferred capabilities, which will be greater if the capabilities of a donee are smaller, whereas it will be lower if capabilities of a donor are greater.
  7. About positive sum in economy see: [6].
  8. In this issue see: [28], [22, pp. 55–56].
  9. Some theorists hold different views on this matter. For example Joanne B. Ciulla considers relations between a *seller* and *buyer* (in her case *employer* and *employee*) in terms of zero-sum game [7, p. 124].

- 
10. Long time the basic notion in this respect referred to the Boethius definition: „The person is an individual substance of rational nature” [5, pp. 84–85].
11. Maria Gołaszewska adduces on the matter to Roman Ingraden’s opinion. A man losses his freedom “when he becomes like a mechanism surrendering to command of the moment, when he himself is not the source of his decisions” [12, pp. 73–104].
12. For instance inviolability of negative freedom can be limited by common consent for the actions that rely on redistribution of goods, see: [21].

## Selfishness and Cooperation: Challenge for Social Life

*Konrad Szocik*

University of Information Technology  
and Management in Rzeszow,  
Poland

*e-mail:* kszocik@wsiz.rzeszow.pl

*Stig Lindberg*

Kyoto University  
Graduate School of Letters  
Christian Studies,  
Japan

*e-mail:* stigberg1@gmail.com

### *Abstract:*

Cooperation is a great challenge for natural selection. Some scholars assume that cooperation could not evolve within the framework of natural selection. It is undeniable that natural selection, at least at the individual level, favors selfishness and defectors. Nonetheless, this selfish tendency does not necessarily imply that cooperation could not evolve by means of natural selection. In this paper, we specifically acknowledge certain basic challenges for the evolution of the human ability to cooperate at the level of large groups. In this paper, we discuss topics like the human ability for “supercooperation,” the importance of repetition and reputation, and Multilevel Selection Theory as the basic mechanisms of evolution of cooperation.

*Keywords:* Cooperation, kin selection, natural selection, direct reciprocity, indirect reciprocity, Multilevel Selection Theory.

## 1. Introduction

Martin Nowak calls humans “supercooperators” [15]. David Sloan Wilson points out that a distinctive feature of humans is that they are not only “groups of organisms,” but builders of

“groups as organisms.” It is especially apparent today when global communication technology such as the Internet enables fast and spontaneous connection between unrelated humans. Humans can unite for the purpose of various social actions, from revolts and protests against authority or some morally wrong phenomena (like a protest against war, crime, environment pollution, etc.) to crowdfunding campaigns and charitable actions. Unlike non-human primates that usually fight for mates and resources, humans are able to moderate and/or suppress eliminative and selfish in-group selection forces, and they are affected by in-group selection forces against inter-group selection. Despite the fact that humans are able to act altruistically, altruism is costly and an unnatural phenomenon. It is so costly and unnatural that altruism in the form of thoughts and intentions does not exist even within various religious frameworks – precisely where one would most expect to find it. Of course, religious altruism does exist, but only as acts, not as thought or intention. That is to say, one must engage in self-sacrifice and contravene one’s natural inclination in order to conduct a religiously motivated act of altruism. This level of analysis is the most important one from the evolutionary point of view because altruism at the level of action may be affected by various motives and causes [23, pp. 49, 84, 142]. One of the proximate causes of altruistic action may be a pure intention, like in the case of psychological altruism. Another cause of altruism may be selfish motivation, like concern for reputation or future welfare (like belief in an afterlife). Some religious practitioners are motivated to altruistic acts by what they perceive to be the will of God. The above, then, represents a rough cross-section of the taxonomy of human altruism.

As alluded to above, there is a good reason to question whether altruistic thought exists within religious frameworks. Certain mechanisms such as the concept of an afterlife and the concept of a supernatural agent could reasonably affect one’s moral life. The concept of an afterlife can affect eusociality. One useful evolutionary mechanism which could motivate altruistic acts is empathy [20]. However, the mechanism of empathy is context-dependent and occurs in degrees, and is therefore “environmentally variable.” [18] Paul Bloom went one step further and points out that empathy is a mechanism that is harmful to social life and should be replaced by compassion [3]. The case of empathy suggests that we cannot rest on any fixed biological mechanism that could automatically provide altruistic and cooperative behaviors. Human life can be understood as a constant tension between short-term selfish approaches and long-term cooperative efforts [15, p. 280]. In this paper, we discuss some challenges for the human ability to cooperate at the level of large groups with unrelated individuals.

## **2. Kin Selection and the Challenge for Cooperation**

Many authors underline the puzzling nature of human cooperation as something strange and in some sense abnormal from an evolutionary point of view [7, p. 422]. Cooperation can be defined as a work toward a common goal in which one agent sacrifices and another benefits [6, p. 4]. A slightly different definition of cooperation is proposed by Samuel Bowles and Herbert Gintis. They define cooperation as “mutually beneficial activity.” [4] They point out that cooperation often functions as a kind of mutually beneficial interaction. Nevertheless, cooperation includes also altruistic behaviors, and this kind of cooperation is more familiar with the type of behavior mentioned above which implies cost for one side of an interaction [4, p. 2]. The reason that cooperation is a puzzling phenomenon is because Natural Selection is generally thought to favor immediate benefits. An agent who provides benefits for others will be outcompeted by the selfish

recipient. Consequently, Natural Selection should not affect the evolution of cooperation and altruism [7, p. 422]. In a one-shot game, both agents have a strong motivation to cheat and/or to defect; consequently, defection works as a default behavioral mode. However, when the probability of repeated encounters between the same agents is sufficiently high, the number of cooperative behaviors increases meaningfully [2, pp. 139–141]. Altruism understood biologically means that one agent minimizes his own fitness in favor of the fitness of another unrelated agent.

It is possible to view altruism and cooperation as cognate behaviors. A specific phenomenon is a eusociality that works as an extreme kind of altruism [10, p. 341]. It is worth noting that the crucial distinction revolves around proximate and ultimate mechanisms and causes. At the proximate level, we are looking for altruism and cooperation that occur in thoughts and feelings. Such an approach refers to conscious states and right intentions. The ultimate approach talks instead about action and effect. In this approach, we can find selfish motivation for altruistic acts, and this motivation is not important from the standpoint of ultimacy [23, pp. 60–65]. From a biological point of view, we talk about altruism in the context of effects of a given act, not in the context of its motivation [4, p. 201]. For this reason, religious components can work effectively as motivational tools to altruistic behaviors even if believers execute them out of fear of God or by hope for eternal reward. In both cases, he is affected by own selfish interests but it does not exclude to talk about altruism in terms of effects.

In the natural world, altruism does not exist [10]. It seems that we cannot find any example in which Natural Selection designed one species for providing benefits for other species. We can find many examples of mutualism, but mutual relations are a kind of reciprocity, and some of them are deeply rooted in parasitism. The case of yucca that are pollinated by yucca moths, is one example of mutualism which has some parasitic origins [10, p. 440]. In the context of a “selfish” natural world, some theorists assume that evolution of cooperation in humans could not evolve naturally within the framework of Natural Selection. The alleged selfish nature of the natural world requires some special instruments that are assumed to be unique to humans in order to explain acts of altruism and cooperation. One of such instrument is the concept of supernatural punishment that is enhanced by religious beliefs, especially by the concept of God/gods who can observe, detect, and punish morally wrong human behaviors [12, pp. 174–177].

Cooperation is a crucial factor for humanity [23]. Human evolution may be explained as the transition from small groups of hunter-gatherers to the current large societies. This process required increasing cooperation within the group. Its importance is so high that some scholars assume that the incestual taboo is not genetically engendered but is rather effected by social rules to provide benefits caused by inter-familial cooperation. Nevertheless, cooperation is not a crucial ability only for human species. In a living world, we can find many examples within species and between species of cooperation. Like Douglas J. Futuyma points out, almost all species have adaptations that make possible cooperation between species [10, pp. 425, 430].

If cooperation is a crucial factor, it is puzzling that it would be so hard to establish and to maintain via natural processes. The evolution of cooperation is hindered by at least two factors. One of them is Natural Selection’s rule that favors selfishness over cooperation. At the group level, it causes the challenge of free-riders. The second challenge is associated with another evolutionary rule that states that the most common kind of relations is genetic [11]. Individuals are more prone to collaborate and to conduct altruistic acts towards their kin not only in small traditional societies, but also in modern, large-scale societies. One strategy for going beyond kinship limits is by practice of



fictive kinships. This process can be maintained by kin-like linguistic forms. This strategy is used especially within the framework of religious language. Nevertheless, artificial social networks never replace the priority and strength of kinships. Like multiple authors have pointed out, members of large families take into account first their family members before looking for unrelated cooperators [8, pp. 20, 25].

It is possible to explain the evolution of cooperation and altruism by individual selection or group selection. The dominant approach states that cooperative behaviors including altruistic ones are affected by mechanisms of individual selection. The basic mechanisms include manipulation, individual advantage, reciprocation, or kin selection [10, p. 339]. The central idea assumed here is that a cooperative and/or altruistic individual cannot intentionally reduce his fitness by selflessly enhancing the fitness of other individuals. If he does so, he will wait for reciprocity from an agent in a pair of interactions, he is manipulated and/or cheated by another individual, or he helps other agents who are genetically related. In the last case, an individual cares for his own genes that are located in his offspring (kin selection) or other related individuals like siblings or cousins (inclusive fitness theory) that are not his own offspring. In this evolutionary landscape, it is difficult to expect that natural selection will affect cooperation and altruism that requires genetic self-sacrifice. Some scholars try to explain this phenomenon that is incompatible with the idea of natural selection by reference to the concept of group selection. However, the leading approach assumes that group selection is too weak evolutionary force [21].

### **3. Multilevel Selection Theory and Evolution of Altruism**

On a biological level, it is possible to discuss the conflict between individual and group benefits. Individual behaviors designed to maximize individual fitness are different from behaviors that are designed to provide group fitness. Wilson points out that group dynamics favor selfishness. Consequently, altruism cannot evolve within a group because Natural Selection favors individual benefits, not group benefits. Altruism can evolve only on the level of inter-group selection when individuals are able to provide benefits for other individuals or the whole group, and such behavior is advantageous for an individual. According to Wilson, we can talk about altruism when a group-level functional organization is at work. Wilson points out that all kinds of explanations of altruism – except for inter-group selection – reduce altruism into selfishness. It refers to Kin Selection and Inclusive Fitness Theory, Selfish Gene Theory, and Evolutionary Game Theory, including direct and indirect reciprocity. Evolution of altruism is affected only by group selection because internal group dynamics, including those mentioned individual selection's mechanisms, favors selfishness [23, pp. 20, 22–23, 29, 32–33, 38–39]. Consequently, it can be assumed that altruism is inexplicable from the point of view of Selfish Gene Theory and from such levels of selection like genetic and organismic levels [22].

Evolution of altruism that is definitely against intra group selection account can work as an argument for group selection theory. Differential group reproduction affects the evolution of altruism by group selection level [9, p. 1523]. Wynne Edwards suggests that animal populations can evolve adaptively at the group level to avoid the risk of exploitation of resources [25]. In this case, we could talk about group selection that leads to group adaptations. However, the idea of group selection is still discussed, and its main explanatory competitor is the concept of Kin Selection and Inclusive Fitness Theory [13].

One possible explanation of human altruism is the idea of inter-group conflict and wars that enhance intra-group altruism by increasing hostility toward outsiders. The focal point of this approach is an assumption that parochial altruism is a behavioral and social phenomenon that started altruism by using default hostility and distrust towards members of other groups. Intra-group selection favors evolution of tolerant nonaltruists. Only inter-group selection and inter-group competition affects evolution of parochial altruists who are hostile towards members of other groups and who are prone to self-sacrifice for the benefits of their group. Evolution of altruism can be connected with the warlike genetic predisposition of humans [5, pp. 636–637, 640].

#### **4. Eusociality**

It is worth mentioning the mechanism of eusociality, which is a domain of some social insects. Wilson points out that eusocial insects are the most dominant species on Earth. They have one of the longest life spans among other species, and they dominate ecologically because they produce the greatest amount of biomass. Biomass of insects comprises the majority of the total biomass. Among biomass of insects, biomass produced by eusocial insects like ants and termites greatly surpasses that of other species of insects. Eusocial species do not go extinct and they affect the evolution of other species. Eusocial life, like Wilson points out, is more beneficial in terms of fitness and provides such benefits like better inheritance of resources or more efficient homeostasis [24, pp. 4, 26–27, 36, 40]. In the case of eusociality, the unit of selection is a colony, not individual organisms. However, there is a discussion between two approaches, Kin Selection and Inclusive Fitness Theory on the one side, and Multilevel Selection Theory, on the other side [17], [1]. This kind of group selection makes a colony a functional superorganism that possesses some extra properties that are unavailable at the individual level. One of them is an ability for mass communication. Mass communication is used to make decisions at the level of the whole group. Another property is adaptive demography that means regulation of the rate of birth and death in the colony that works to achieve an optimal level of division of labor. Wilson points out that eusocial organizations provide success and dominance for eusocial insects [24, pp. 72, 97].

As mentioned above, the evolution of eusociality can be explained in terms of Kin Selection. On the one hand, the eusocial population seems to be caused by genetic relatedness among members. A eusocial queen with small numbers of offspring-workers has much higher rates of reproduction than solitary insects [17]. On the other hand, genetic relatedness in eusocial populations can be a consequence, not a cause, of eusociality. In this case, eusociality could evolve among unrelated individuals who united themselves for the purpose of common defense of their nest and resources. Eusociality among unrelated individuals is explained by Multilevel Selection Theory, including Group Selection, not by Kin Selection or Inclusive Fitness Theory.

#### **5. Human Cooperation is Affected by Repetition, Reputation, and Social Networks**

It is no doubt that cultural evolution, especially cultural group selection was a crucial factor for the evolution of cooperation. It is important to distinguish here cultural and genetic evolution. Some cultural factors like religion may use biological and/or psychological components that make them by-products from a genetic point of view. However, the same factor can work as adaptations at the

level of cultural evolution because they provide benefits for a given group that can be measured in terms of survival and reproduction [23, p. 79].

Human cooperation can be treated as a unique and special phenomenon in a living world. One of the reasons is specific human biological and cognitive equipment, including such cognitive devices like language or theory of mind. According to Dominic D. P. Johnson, these devices impede human defection and selfishness because such selfish and socially detrimental acts become easy to detect and punish [12, p. 169]. The crucial difference between humans and non-human primates – which is due to a lack of aforementioned cognitive capabilities among non-human ancestors – is the impossibility of punishment by absent individuals. Morally wrong human behaviors can be punished because linguistic skills and mind-reading ability enable detection of prohibited acts and their perpetrators even if punishing agents were absent from the event. In this context, Johnson suggests that evolution of the theory of mind and complex language among humans made selfishness a very costly act that can be easily detected and heavily punished by the entire group [12, p. 172].

There is no doubt that language played a crucial role in human evolution. However, a causal correlation that worked in the field of evolution of language and of cooperation is not clear. Nowak points out that natural selection favored people who could communicate by an extended set of linguistic signs [15, p. 177].

Bowles and Gintis put language and cultural transmission of social norms among the tools that have enabled the evolution of cooperation. They underline a crucial role that was played by uniquely human cognitive and linguistic abilities. Human evolutionary group dynamics could work according to them by following social structures: intergroup competition including warfare, within-group cooperative practices, and institutions that were used to promote and to enhance preferred values and rules [4, pp. 196–197]. According to some theorists, the phenomenon called the Tragedy of the Commons requires institutional (third party) support for providing collaboration for the purpose of the common good. In a default natural environment, people tend to cheat and defect not only for selfish motives but also out of concern over being cheated by others [15, p. 208]. In other words, they would rather “beat others to the punch.”

Nowak points out that the care for reputation works as an important motivational factor [15, pp. 215–216]. Reputation is a crucial concept for indirect reciprocity [16]. One of the mechanisms that is used to strengthen and to control cooperation is the practice of punishment. Nowak rejects the concept of altruistic punishment because he finds selfish motivations beyond the practice of punishing others. He talks about costly punishment instead of altruistic punishment. According to Nowak, punishing others is not a basic mechanism but it is associated with direct and indirect reciprocity. The leading role played by direct and indirect reciprocity is the consequence of the human evolutionary history of living in small groups. Living in small communities was affected by the high probability of encounter of the same agent and by concern for reputation. According to Nowak, it is impossible to remove the criteria of repetition and reputation from everyday human life. However, Nowak points out that punishment is not a useful mechanism for the evolution of cooperation because people who gain the best results in economic games do not punish. Consequently, people who prefer to punish do not win. The practice of rewards is much more effective than the practice of punishment for establishing cooperation and the common good. It is worth bearing in mind that behavioral dynamics in terms of reward, punishment, and cooperation is context-dependent. In one survey, participants from Greece and Russia were much more prone to

retaliate instead of accepting punishment for their deceptions than participants from other countries, especially from the West [15, pp. 225–227, 229–233].

Humans are organized into cooperative units. Like in the case of intersexual mating, we find the phenomenon of assortative mating (when someone is looking for a mate with whom he shares some similarities like origin, ethnicity, religion, hobbies, etc.) in the rest of everyday life activities humans organize into assortative sets. Such sets join humans according to their interests, preferences, etc. Set membership creates networks that can provide effective cooperation among unrelated individuals. Network reciprocity works in the most efficient way if there is an optimum point between low and high rates of mobility. Such an intermediate level provides chances for building stable relations with opportunities to avoid exploitation by defectors.

Nowak enumerates five mechanisms for the evolution of cooperation. The basic one is a repetition that is associated with encounters with the same agents and with the so-called tit for tat rule. It is a domain of direct reciprocity. Another crucial mechanism is indirect reciprocity that is based on the concept of reputation. Nowak mentions also spatial selection (network reciprocity). Multilevel Selection (group selection) means a selection between groups. Nowak points out that Multilevel Selection works more effectively at the level of many small groups than at the level of a few large groups. Finally, Nowak talks about the kin selection that affects cooperation by genetic relatedness.

Nowak is a proponent of the concept of cooperation as a principle of Natural Selection. He points out that such features like hope for reciprocity, forgiveness, and generosity in everyday human interactions often are default behavioral modes that are used to strengthen and to maintain cooperation. Nowak points out that mankind is linked together by global technology that provides an opportunity of global communication for the first time in human history. It is important for possible cooperation for the purpose of neutralizing negative effects of climate change [15, pp. 258, 262–263, 270–272, 278].

Cooperation can be genuinely treated as a natural mechanism as Nowak asserts. Some moral emotions can work as tools that affect prosocial behaviors. One recent study by Nowak in the field of Evolution of Cooperation suggests that reward and forgiveness are much more effective than retaliation and punishment in pro-cooperative policy. The discussed study shows that guilt is more useful for establishing cooperation than retaliation and anger [19].

## **6. Conclusion**

Evolution of Cooperation is a highly discussed topic today. This discussion attracts attention from biologists and psychologists as well as from philosophers and theologians. Some of them try to find similarities between human and non-human animals and show that the human ability to cooperate is not a unique mechanism. Others point out that humans are unique biologically and socially [14]. They show that biological and cognitive human uniqueness affects cultural evolution, which in turn shapes human social relations.

It is no doubt that only humans build global social networks. Only humans can interact spontaneously with other unrelated humans. This unique capacity is the combined product of cognitive devices like language and mind with highly selective pressure for collaborative patterns. Ancestors with higher pro-cooperative tendencies probably had greater chances of survival and reproduction than less cooperative individuals. Human cooperation is not affected only by the need

for help and care since other non-human animal species also require care and help. Human cognitive devices could enable the evolution of special prosocial tools and strategies. It seems that they are based on criteria of repetition and reputation, and on all good and bad effects that are affected by them. Of course, some individuals can try to calculate means of avoiding negative effects of cheating or defection. Personal reputation is also a strong deterrent from cheating or defecting. Defectors and cheaters should understand that the others in their community will no longer collaborate once outed. Like in the case of repetition, also in the case of reputation someone can assume that he does not have to care for his own reputation. However, mentioned special cognitive human abilities like mind and language make reputation a really important factor in human communities.

In this context, we could say that human cooperation is the product of biological and cognitive abilities that make social relations exceptionally important, much more than in the case of any other animal species. Cultural tools like religion/religious components, legal systems, or some traditionally inherited cultural patterns, are secondary agents. It can be assumed that they can regulate or enhance the power of mechanisms of repetition and reputation, but their role is not as important as some proponents of cultural evolution assume.

## References

1. Abbot, P. et al. Inclusive fitness theory and eusociality, Brief communications arising from M. A. Nowak, C. E. Tarnita & E. O. Wilson *Nature* 466, 1057–1062 (2010), *Nature*, 24 March 2011, Vol 471.
2. Almenberg J. & Dreber, A. Economics and evolution. Complementary perspectives on cooperation, In M. A. Nowak & S. Coakley, *Evolution, Games, and God. The Principle of Cooperation*, Cambridge, London: Harvard University Press, 2013.
3. Bloom, P. *Against Empathy. The Case for Rational Compassion*, New York: Ecco, 2016.
4. Bowles S. & Gintis, H. *A cooperative species. Human reciprocity and its evolution*, Princeton and Oxford: Princeton University Press, 2011.
5. Choi, J.-K. & Bowles, S. The Coevolution of Parochial Altruism and War, *Science*, Vol. 318, 26 October 2007.
6. Coakley S. & Nowak, M. A. Introduction. Why cooperation makes a difference, In M. A. Nowak & S. Coakley, *Evolution, Games, and God. The Principle of Cooperation*, Cambridge, London: Harvard University Press, 2013.
7. Cownden, D., Eriksson, K., Strimling, P. A popular misapplication of evolutionary modeling to the study of human cooperation, *Evolution and Human Behavior* 38 (2017).
8. David-Barrett, T. & Dunbar, R. I. M. Fertility, kinship and the evolution of mass ideologies, *Journal of Theoretical Biology* 417 (2017).
9. Eldakar, O. T. & Wilson, D. S. Eight Criticisms Not to Make About Group Selection, *Evolution*, 2011 June, 65(6).
10. Futuyma, D. J. *Evolution*, Massachusetts Sunderland: Sinauer, 2006.
11. Hamilton, W. The genetical evolution of social behavior (I/II), *Journal of Theoretical Biology*, Vol. 7, Issue 1, July 1964, 1–52.
12. Johnson, D. D. P. The uniqueness of human cooperation, cognition, cooperation, and religion, In M. A. Nowak & S. Coakley, *Evolution, Games, and God. The Principle of Cooperation*, Cambridge, London: Harvard University Press, 2013.
13. Kohn, M. *The needs of the many*. *Nature*, Vol. 456, 20 November 2008, 296–299.
14. Laland, K. *Darwin's Unfinished Symphony. How Culture Made the Human Mind*, Princeton: Princeton University Press, 2017.

15. Nowak, M. (with R. Highfield). *SuperCooperators: Evolution, Altruism and human behavior or Why We Need Each Other to Succeed*, Edinburgh: Canongate, 2011.
16. Nowak, M. A. Five Rules for the Evolution of Cooperation, *Science* Vol. 314, 8 December 2006.
17. Nowak, M., Tarnita, C. E., Wilson, E. O. The evolution of eusociality, *Nature*, 466, 26 August 2010.
18. Peterson, G. R. Is my feeling your pain bad for others? Empathy as virtue versus empathy as fixed trait, *Zygon*, vol. 52, no. 1, March 2017, 232–257.
19. Skatova, A., Spence, A., Leygue, C. & Ferguson, E. Guilty repair sustains cooperation, angry retaliation destroys it, *Scientific Reports* 7, Article number: 46709 (2017).
20. Waal, de F. Putting the altruism back into altruism: the evolution of empathy, *Annu Rev Psychol.* 2008, 59:279–300.
21. Williams, G. *Adaptation and natural selection*, Princeton: Princeton University Press, 1966.
22. Wilson, D. S. & E. O. Wilson. Rethinking the theoretical foundation of sociobiology, *The Quarterly Review of Biology*, December 2007, Vol. 82, No. 4, 327–348.
23. Wilson, D. S. *Does Altruism exist?* New Haven and London: Yale University Press, Templeton Press, 2015.
24. Wilson, E. O. 1990. *Success and Dominance in Ecosystems: the Case of the Social Insects*, Oldendorf/Luhe: Ecology Institute.
25. Wynne-Edwards, V. C. *Evolution Through Group Selection*, Boston: Blackwell Scientific, 1986.

**Selected Ethical Issues in Artificial Intelligence,  
Autonomous System Development  
and Large Data Set Processing**

*Paweł Zgrzebnicki*

SWPS University of Social  
Sciences and Humanities,  
Warsaw, Poland

*e-mail:* pzgrzebnicki@st.swps.edu.pl

*Abstract:*

Due to information technology development and its industrial adaptation, the dilemmas so far specific to philosophical speculations and science-fiction literature and films have become the real problems of the contemporary world. On one hand, these issues are related to an unprecedented scale on which computational algorithms are currently used as well as a level of complexity of mutual connections; on the other hand, these are linked to their autonomous behavior. States, industry, and users themselves demand formulation of understandable ethical categories and determination of transparency standards and legal norms for these algorithms' functioning in the near future.

*Keywords:* artificial intelligence, ethics, industry, autonomous systems, algorithms.

## **1. Scale**

The world has been expanding its network of connections used to exchange information and this expansion not only consists of building new data transmission paths. The Internet architecture has been already steady and settled; however, its scale and, most of all, computing power are increasing. The scale of infrastructure and the scale of impact that IT technologies have on the lives of the people of the world today are now higher than ever before. It is precisely this scale that brought humanity closer to the point at which the existing legal solutions and philosophical ideas, in particular, ethical and human relation to non-humanity, turn out to be unreliable. First and foremost, the scale enabled the emergence of well-functioning artificial intelligence based on *deep learning* algorithms because the latter requires the availability of *big data*. Finally, it is also due to the scale of autonomous devices becoming so accessible that they require new legislation, and the social consequences of the technological revolution become so significant that they raise massive concerns not only about privacy but also about keeping jobs and even a complete change of political and economic order in the world.

What is this scale? According to one of the reports [48], the cloud data storage market suppliers will earn 120 billion dollars in 2020, and the sector's annual growth rate has been estimated to be 38.4%. It has been noticed that IaaS leaders experience even faster growth

compared to the rest of the market segment – 67.8% annually. According to other data [12] until 2021, the cloud computing market will double in relation to 2016, thus in the next five years. The perspectives of computing power development and its scale may be indirectly visible in sale predictions concerning backup power systems. According to the report [13], this market is growing faster than linearly and this trend is going to last at least until 2020. The scale of the undertakings is also visible in a physical area covered by the data centers – from 40.000 to over 100.000 square meters [47] while the unquestionable leader in this scope is Chinese Range International Information Group using the infrastructure of the area of almost 630.000 square meters [31]. End users who store their data on these companies' servers often do not pay for it at all or at least that is the impression. Making your data accessible to the companies as Facebook, Google, Microsoft or Apple is partly free and the companies make a profit thanks to the possibility of data processing. On a scale of billions of users sending as many queries to databases [19], large corporations have enormous amounts of information, among others, personal information as sensitive data concerning political views, sexual preferences or health records. Photos, films, metadata associated with voice communication, electronic mail patterns, emotional reaction patterns revealed in comments to articles in social networks as well as website opening sequence patterns or geolocation tracking are only some of the numerous pieces of information provided by every person connected to the common network as a “payment” to private companies and governments which are not necessarily their national governments. When algorithms are fed with big data and learn new skills as well as acquire the ability to recognize new patterns (*machine learning*), the problem occurs described by Daniel Tunkelang as “privatizing our past” [52].

So society is not only a beneficiary of “universally accessible and useful” [26] services but also, and perhaps above all, a living resource, fuel, by which all modern computer machinery can operate at all. For this reason, infrastructure impetus translates directly into legal, ethical and social issues. An example here can be a scandal caused by revealing a sociological experiment carried out by Facebook social network. The website, which currently has 1.86 billion active users [51], five years ago, that is in 2012, was used to analyze its clients' behavior on a sample of 689.003 people. The experiment of which the participants were not informed, consisted in showing, in most part, negative contents to the selected group of users while the other group was shown positive content mostly. Here, the fact that Facebook does not display all the information from pages liked by the users but only those which are recognized to be best adjusted algorithmically is worth mentioning. The algorithm itself is obviously a secret. After a week of manipulation, they investigated the users' reaction being the result of the supposedly induced mental state. As it turned out after the test was completed, the group, which was exposed to negative information, showed the much worse frame of mind compared to the one which had access to positive message [28]. The experiment stirred up emotions and ethical controversies. The political milieu even asked the question whether the CIA can trigger a revolution in some country by manipulating the mood of the public [5]? The questions concerning the idea proved to be logically valid taking the controversies related to United States presidential election in 2016 [1] and future 2017 election in Germany [41] into consideration. In the latter case, German authorities have their sights set on, among others, the aforementioned world's largest social network [50].

Large-scale ethical issues related to the processing of data from huge resources can be divided into two groups. One discusses the use of directly acquired data, a way of their processing and the fact that personal stories and traces left by each one of us are used as natural resources extracted by both private industry and governments in order to obtain particular benefits and the users are not necessarily the beneficiaries. The other group of moral issues is related to a dilemma regarding the way the obtained data are used towards the receivers who are often also the data providers. The aforementioned manipulation is just one of the examples but the class of objections can be expanded, also towards data processing and establishing decision thresholds. For instance, there is a system operating in Austria registering patient information, their medical appointments, picked up prescriptions and diagnoses and other detailed information including geolocation [39], and it is an example of a complex network that can be subject to a multidimensional analysis [23].



The system's designers expected that the citizens equipped with smart cards and 12.000 physicians would provide detailed data on an ongoing basis which would be further processed [15]. However, the system has raised serious ethical objections. Among others, there are questions of patient's right to decide what is done with their body, transparency and international medical data transfer [34]. An analysis of a great amount of information enables one to draw useful conclusions. Among others, through comparing a certain illness history and previous treatment profile for this and other diseases with similar cases, one may, with great probability, determine how the illness will develop in a particular human being in the future [49]. The ability to predict the further development of a disease and its transformation into a number of new disease entities may result in offering suitable preventive treatment to a patient as it is much cheaper to an insurer than later treatment of many possible future diseases. But what if the algorithm finds that the expected cost is too high? Will the patient be qualified for treatment or even informed of a threat when the predicted expense is too high to the system? What if the case lies below the threshold used to make a preventive treatment decision? Is there a place for empathy in this kind of algorithmic system at all or is it only governed by statistics and economic calculations? This is the example where automatic data processing resulting in data set classification implies the questions of the moral character of division and control over the result. Not only private companies but also states and international organizations operating for the “public good” are currently under pressure of determining the context of their actions, providing proper transparency and public participation in implementation and control of the tasks executed by algorithms [30].

## **2. Artificial Intelligence and Autonomous Devices**

The ethics of artificial intelligence has only recently been discussed, mainly within the frameworks of philosophical discussions based on visions created in science fiction. The previous questions have been focused mainly on fear of humanoid robots taking over the world or even destroying the humanity in its biological form. The visions presented in the series of the Terminator movies starring Arnold Schwarzenegger have become in the popular culture the symbol of social fear of superior intelligence form. Also, the “Matrix” trilogy, a masterpiece of filmmaking from a dozen years ago, frightens the viewers with the humanity apocalypse brought about by ruthless machines using a biological component as their power source.

Already in 1942, similar analyses concerning a possible result of human and machine interaction led Isaac Asimov to formulate the three laws of robotics which he presented in one of his stories [2]. According to these laws (1) A robot may not injure a human being or, through inaction, allow a human being to come to harm, (2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law, and (3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws. In his text from 1985 [3], the author added one more law to precede the others, which sounds exactly like the First Law except a “human being” phrase was replaced with a “humanity” word.

If one discusses a real class of problems related to artificial intelligence application, Asimov's laws seem to be too general and impracticable, for instance, with regard to the perspective of using an autonomous weapon. What is more, the notion of “harm” itself is ambiguous and dependent on both situational and cultural contexts. Does an automaton treating a patient and inflicting temporary pain on him in the dentist's office but eventually relieving the ailment do harm to a human being or not? Is an autonomous car in the situation of an unavoidable accident supposed to save passengers at the expense of passers-by or the other way round [20]? In many cultures, the good of the society is ranked higher than the good of an individual. However, the degree to which individual's rights are protected, even within the specific culture group, depends on the context. For instance, according to Article 3 of the Universal Declaration of Human Rights, everyone has the right to life and liberty [53], and that does not prevent states from imprisoning people and sentencing them to death in the particular situations. This happens in the name of “public welfare” that is the situation where the expected benefit of the group is ranked higher than

individual well-being. When introducing his zeroth law, Asimov must have been aware of this dualism; however, his rules are still too general to be something more than possible guidelines for a practical code of ethics for machines.

The discourse on the law, ethics, economic and social consequences of development in the area of artificial intelligence and autonomous devices has already gone beyond the fictitious space and has become a real problem in the face of technology developing more with every year [10] [18]. This has happened because of accessibility of great computing power which has enabled one to make complex calculations in a short time (or even in a real time), accessibility of large data sets which are indispensable for training and testing algorithms as well as developing mathematical methods used to execute the tasks of artificial intelligence. Within the ethical issues worth discussing here there are three areas requiring a detailed analysis: (1) determining acceptable goals that can be pursued by an autonomous algorithm or a device using this algorithm, (2) understanding the decisions made by artificial intelligence objects and (3) making supervision over their actions possible.

The first area (re. 1) is related to two classes of questions: (1a) which human actions are supported or taken over by artificial intelligence and (1b) what is the overall objective it is supposed to pursue? Additionally, in relation to these two classes, the primary question is what is the general purpose of these algorithms? The silent assumption that autonomous devices or programs are to “help with something” or “replace somebody in doing something” is often *a priori* used in discussions from the anthropocentric position; however, one may ask the questions whether animal or climate protection is equally important or more important than well-being of one, selected *homo sapiens* species, or a specific individual? After all, it is possible that a being much more intelligent than a human will come to the conclusion that he is the one standing in its way to achieve a seemingly ethical goal? Therefore, should one consider the question of human relations with the environment from the perspective of human ethics or a broader one that includes ecology in general? There can be a practical question asked concerning “doing harm” or “doing good” not only in relation to an individual and society, but also the environment they are functioning in – both biological and technological one – with which they are inseparably connected. If one should take care of plants essential to our lives, maybe they should also take care of devices; especially, if they demand so, governed by their own form of self-awareness? Therefore, the next question of this discussion should concern the aim and foundation of ethics in general. Depending on the answer, one can discuss further dilemmas, including those that have been already mentioned.

However, even if one assumes that artificial intelligence is to put the spotlight on a human being and only him, there should be a permissible area of its functioning determined (re. 1a). Is fully autonomous weapon a right application, and if so, what is the scope and method of its design [21], [32], [46]? Will totally autonomous medical robots and diagnostic systems be able to do every task? Or there still should be a human supervision, for instance, to tell the psychologically painful diagnosis as gently as possible? It seems that these issues still need a discussion focused on human emotional well-being; on the other hand, there is an overabundance of analyses perceiving next technological improvements as solutions to every problem. Thus the cold showers of *post factum* reflections engineers are forced to take every now and then when a system without “parental” supervision starts to transform into an ethical monster dangerously fast [6]. Even if one allows a machine some leeway, they should define its goal (re. 1b). It seems that to a living organism, its gene pool survival is of primary importance, and a device also should know its task. As in the case of a search algorithm, one should determine whether the aim is an increase in human interaction with the system for marketing purposes, pure economic profit or something else? In either case, the technology will execute its tasks the best it can; however, when formulating the instructions one should be extremely careful in expressing their wishes. It has been already known that even a simple form of artificial intelligence is capable of eliminating everything that stands in its way ruthlessly [29] or stealing somebody’s resources [42].

The second, already mentioned area of analyses (re. 2), consists in discussing the possibility of understanding a decision made by an intelligent system. The algorithms based on machine

learning function like black boxes [11], [38] and currently their proper functioning is verified on the basis of their effectiveness and not by means of tracing a decision process. Although there are some attempts to solve the puzzle of understanding [17], the problem seems to be difficult due to the very essence of such an algorithm. A network takes action on the basis of many interdependent elements, entangled units, and as in the case of a human brain, where it is impossible to indicate which neuron is responsible for a certain reaction (especially when it is complex), the same is with the discussed algorithm – it is impossible to provide a simple answer where particular components would be separated from input data, learning history and mutual relations [8, pp. 532–545]. Like the boomerang, in a somewhat transformed form, the old question of who takes the blame, the criminal or the society where he was brought up, is bouncing back. Although one can give a simple answer given that there is free will, assuming the null hypothesis, according to which there is no free will, leads to the same issue of responsibility and its blur and, most of all, to the ontological question of existence and its physical limits. In the case of people, nature has offered a way out of this difficult situation, namely, psychological rationalization. In some cases, a human being is able to give a reason for his behavior and explain the course of events that have led him there. Leaving the truthfulness of the provided human explanation and his actual awareness of real reasons aside, this action is sufficient enough for the majority of social interactions and self-understanding. It is possible that autonomous systems will also have to be equipped with a system able to give a credible justification for their actions. Here, one should really ponder over how to verify whether these statements are true or false since deception might be an as good strategy as any other action.

This discussion indirectly brings on the third group of questions (re. 3), that is, giving a man possibility of supervising machines. Are less intelligent individuals truly able to supervise the one that makes decisions which are unclear to them? Is it possible to control the system which, in fractions of seconds, conducts analyses totally impossible to a human being? Eventually, it is probable that every, even the most sophisticated control system, including a self-destruction system could be evaded by a sufficiently intelligent, self-adapting and self-modifying device. This is a great problem that perhaps can only be solved in a quite drastic way - when people start to integrate with machines and create a hybrid being joining human moral values with engineer excellence of intelligent machines together. It is impossible to predict today whether morality would be needed at all in the future. If one perceives ethics as a tool used by a social group to survive it can lead them to have concerns about its redundancy in the world where welfare (whatever its definition) is provided with the use of other, much more efficient strategies.

### **3. Legal and Ethical Regulations**

There is a growing number of people, companies, governments, and organizations recognizing the scale and complexity of the problems of artificial intelligence, autonomous systems as well as processing and accessing big data. There is a dispute over whether the new technology has in itself anything of the essence that can be supervised legally [8] or whether problems and threats should be considered only as emergent and contextual [4]. There is already a historical context of the problem [7], and there are suggestions on how to regulate the market for new services, but above all, there are problems of philosophical nature without which it is impossible to create effective and satisfactory solutions. People ask questions of what “a being” and “ethics” are and who has the “natural” right to what exactly as well as who has established this right. The questions which have been so far addressed by religion in the anthropocentric world and which have been discussed in a closely knit group of philosophers are now attracting the attention of lawyers, company boards, engineers, certifying authorities and members of the public. The European Union calls for establishing a new law and sees a possibility of appropriate regulation in introducing the notion of an “electronic person” [43]; Japanese scientists are implementing their version of a code of ethics that is to be helpful in their work with artificial intelligence [25], while IEEE – the world’s largest association of electrical and electronics engineers and representatives of branches related to computer technology and telecommunication – is working on standards concerning ethics during a

designing process [35], transparency of autonomous systems [36] and privacy issues in data processing [37]. In December 2016, the same association drew up the 136-page report outlining the assumptions and a range of dilemmas one must face within this area [16]. Two months earlier, the office of the President of the United States referred to the problem of the future of artificial intelligence, among others, stating that “AI can be a major driver of economic growth and social progress, if industry, civil society, government, and the public work together to support development of the technology, with thoughtful attention to its potential and to managing its risks” [40, p. 39].

In the context of such documents, a meta-analysis of the problem and question of who establishes ethical values, which are subsequently confined within the legal framework or subjected to standardization, seem to be justified. The world in which a law established in one country is only binding in this particular country is disappearing, especially in the context of international products and services. The contemporary algorithms are worked out by international teams and introduced to the market by companies with departments in many countries. Target receivers of the products are usually people from all around the world. Whose ethics should be binding in this case? Is there a universal ethics? Isn't it a post-colonial effort of those who have the capital to dictate the only right standard to the public? Even if determining standard design solutions is based on an open discussion taking place with the use of online conference software, what is the representation of people from countries with no access to the Internet, those who are undereducated, do not know English or are subject to restrictions on their own governments making the connection impossible? Who are the engineers to decide on ethical standards? Are they becoming contemporary priests, revealing moral truths to the world because they are best-educated caste controlling the functioning of the technological world? Aren't the representatives of exact sciences trying to categorize the area that is actually beyond their competence?

Quite suddenly, it turned out that ethical values are essential to practical application and one can no longer rely on good manners, silent assumptions, social pressure or the system of punishments and rewards which have developed to regulate cooperation and coexistence of human groups. The companies want to know their scope of liability, and people want to know what they can expect from much more intelligent “beings” with which they are willing to cooperate. Ethics has become a valuable asset, a rare semi-finished product the world wants to use to create a predictable future for itself. Attempts to standardize moral values interwoven with an industrial process may also mean the near end of ethical relativism, popular in times of neoliberalism.

#### **4. Social Outcomes**

The Federal Republic of Germany has announced it is getting ready for a turning point that will be the fourth industrial revolution [33]. Digitalization of economy and basing it on intelligent machines are going to bring changes that are to be so serious that Australia also expects rapid changes and cooperates in this scope with the European partner actively [24]. The perspective of machines taking over human competences on a large scale raises concerns regarding the future of the labor market. For instance, JPMorgan bank has implemented new software which does the same things that took lawyers 360,000 in total in just a few seconds [45], while Bridgewater Associates, a hedge fund managing 160 billion dollar portfolio, is going to entrust their investments to fully automated system [44]. The concerns related to machine replacement for a human being are not new [9]. As they were expressed already in the Victorian era in England, and in the 50s in the United States, one may think there is really nothing to fight over. However, the current change is taking place on an unprecedented scale, in the world that has never been so interrelated before. Historically speaking, the quality of the network consisting in the fact that a disturbance in just one part of the world may have an almost immediate influence on the whole area is totally new. In this sense, one is dealing with a qualitatively new situation. Soon, thanks to the technology, intelligent devices connected to one network and working “for the society's purposes” can totally replace people in some branches by redeveloping, improving and adapting themselves to changing

conditions. There are opinions that the democratic system and Euro-Atlantic social model may not last out this revolution [22]. Some see the basic income as a solution as people would receive it regardless of whether they work or not [14], others are rejecting this idea [27], but whatever the solution might be, economic redistribution of money will not bridge the emotional social gap people with no work or aim and thinking they are useless will be struggling with. A mob of the unemployed with any basic income will surely generate a gigantic social and emotional dysfunction in the cultures where work constitutes a virtue and sense, where the ethos of work determines values making people proud of their own abilities and feel valued by a family and society. This is a great problem since one cannot reconstruct the system of values carried along by many generations, religious systems, myths and canons in a dozen years. It seems impossible, and in this context, the worries whether societies and states survive such a revolution are extremely real. Even from the economic point of view, the potential problem cannot be limited to unemployment but it is related to a collapse of the whole branches of economy. The skyscraping office building where millions of corporate employees are doing their painstaking work right now may become empty due to artificial intelligence. The arduous work of IT specialists, accountants, analysts, logistic managers, drivers, traffic organization specialists and any work that can be dressed up in an algorithm and learnt can be also taken over by more or less autonomous devices that do not need a vast physical space and can be digitally transferred to a totally different country. The real estate market disaster that took place in the United States in 2008 may turn out to be only a prelude to another collapse in this branch, this time in the commercial space segment.

There is also an equally important issue of the countries less technologically developed than the United States, Western Europe, India, Japan or China. How their citizens will find themselves in a new economic reality on which they have hardly any influence? Will the colonial hierarchy, destroyed by the post-war ideas, be rebuilt in a new posthumanist reality? So far it seems that even the citizens of the countries which would call the shots here are not entirely ready for it.

## 5. Summary

In the world being on the verge of a new industrial revolution and possibility of broad interaction with artificial intelligence, the issues related to ethics become pressing problems to the governments, companies, and societies. Everyone emphasizes the need for new legal regulations based on the transparent ethical declarations; however, codification itself is not the most difficult matter here, unlike the common agreement on what the humanity expects from further development and how it perceives itself in juxtaposition with intelligent machines which objectives may evolve contrary to human expectations.

## References

1. Allcott, H. & Gentzkow, M. *Social Media and Fake News in the 2016 Election*, National Bureau of Economic Research, 2017.
2. Asimov, I. *Świat robotów*, Warszawa: Varia APD, 1993.
3. Asimov, I. *Roboty i Imperium*, Poznań: Dom Wydawniczy 'Rebis', 2014.
4. Balkin, J. M., The Path of Robotics Law, *The Circuit*, Paper 72
5. Booth, R. Facebook reveals news feed experiment to control emotions, *The Guardian*, 2014. Available at: <https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds/>, accessed: 2016-02-28
6. Bright, P. Microsoft terminates its Tay AI chatbot after she turns into a Nazi, *Ars Technica*, 2016, Available at: <https://arstechnica.com/information-technology/2016/03/microsoft-terminates-its-tay-ai-chatbot-after-she-turns-into-a-nazi/>, accessed: 2017-03-01.
7. Calo, R. Robots in American Law, *Legal Studies Research Paper*, vol. 2016-04, 2016.
8. Calo, R. Robotics and the Lessons of Cyberlaw, *Cal. L. Rev.*, vol. 103, pp. 513–563.

9. Campa, R. *Technological Unemployment. A Brief History of an Idea*, International Sociological Association eSymposium, 2017.
10. Campa, R. *Humans and Automata. A Social Study of Robotics*, Frankfurt am Main: Peter Lang, 2015.
11. Castelvechi, D. Can we open the black box of AI? *Nature*, vol. 538, no. 7623, 2016, pp. 20–23.
12. *Data Center Outlook – Strong Demand, Smart Growth, North America 2016*. Jones Lang LaSalle, IP, Inc., 2016.
13. *Data Center UPS Market Analysis By Product (Small Data Centers, Medium Data Centers, Large Data Centers) And Segment Forecasts To 2020*. Available at: <http://www.grandviewresearch.com/industry-analysis/data-center-ups-market/>, accessed: 2016-02-28.
14. Diamandis, P. If Robots and AI Steal Our Jobs, a Universal Basic Income Could Help, *Singularity Hub*, 2016. Available at: <https://singularityhub.com/2016/12/13/if-robots-steal-our-jobs-a-universal-basic-income-could-help/>, accessed: 2017-03-01.
15. Dorda, W., Duftschmid, G., Gerhold, L., Gall, W. & Gambal, J. Introducing the electronic health record in Austria, *Studies in health technology and informatics*, vol. 116, 2005, pp. 119–124.
16. *Ethically Aligned Design*. IEEE, 2016.
17. Falk, C. *Auditing Deep Neural Networks and Other Black-box Models*, 2016. Available at: <http://thesis.haverford.edu/dspace/handle/10066/18664/>, accessed: 2017-03-01.
18. Frankel, S. 5 predictions for artificial intelligence for the coming year, *ReadWrite*, 2016. Available at: <http://readwrite.com/2016/12/15/5-predictions-for-artificial-intelligence-in-2017-dl1/>, accessed: 2017-03-01.
19. Google Annual Search Statistics. *Statistic Brain*, 2016. Available at: <http://www.statisticbrain.com/google-searches/>, accessed: 2016-02-28.
20. Greene, J. D. Our driverless dilemma, *Science*, vol. 352, no. 6293, 2016, pp. 1514–1515.
21. Guetlein, M. A. *Lethal autonomous weapons – ethical and doctrinal implications*, Joint Military Operations Department, Naval War College, 2005.
22. Helbing, D. et al. Will Democracy Survive Big Data and Artificial Intelligence? *Scientific American*, 2017. Available at: <https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/>, accessed: 2017-03-01.
23. Herbek, S. et al. The Electronic Health Record in Austria: a strong network between health care and patients, *European Surgery*, vol. 44, 2012, pp. 155–163.
24. *Industry 4.0. Australian Government, Department of Industry, Innovation and Science*, Available at: <https://industry.gov.au/industry/Industry-4-0/Pages/default.aspx>, accessed: 2017-03-01.
25. Japanese scholars issue ethical guidelines on AI, *NHK World* (2017), Available at: [https://www3.nhk.or.jp/nhkworld/en/news/20170228\\_36/](https://www3.nhk.or.jp/nhkworld/en/news/20170228_36/), accessed: 2017-03-01.
26. Jurevicius, O. Google mission statement 2013, *Strategic Management Insight*, 2013. Available at: <https://www.strategicmanagementinsight.com/mission-statements/google-mission-statement.html>, accessed: 2016-02-28.
27. Kharpal, A. Mark Cuban: Basic income ‘the worst possible response’ to job losses from robots, *CNBC*, 2017. Available at: <http://www.cnbc.com/2017/02/22/mark-cuban-basic-income-worst-response-to-job-losses-from-robots-ai.html>, accessed: 2017-03-01.
28. Kramera, A. D., Guilloryb, J. E. & Hancockb, J. T. Experimental evidence of massive-scale emotional contagion through social networks, *PNAS*, vol. 111, no. 24, 2014, pp. 8788–8790.
29. Leibo, J. Z., Zambaldi, V., Lanctot, M., Marecki, J. & Graepel, T. Multi-agent Reinforcement Learning in Sequential Social Dilemmas, *ArXiv Prepr. ArXiv170203037*, 2017. Available at: <https://arxiv.org/abs/1702.03037/>, accessed: 2017-03-01.
30. Lepri, B., Staiano, J., Sangokoya, D., Letouzé, E. & Oliver, N. The Tyranny of Data? The Bright and Dark Sides of Data-Driven Decision-Making for Social Good, *ArXiv Prepr. ArXiv161200323*, 2016. Available at: <https://arxiv.org/abs/1612.00323>, accessed: 2017-03-01.

31. Lima, J. Top 10 biggest data centers from around the world, *Computer Business Review*, 2015. Available at: <http://www.cbronline.com/news/data-centre/top-10-biggest-data-centres-from-around-the-world-4545356/>, accessed: 2016-02-28.
32. Lin, P., Bekey, G. & Abney, K. *Autonomous military robotics: Risk, ethics, and design*. California Polytechnic State University, 2008.
33. MacDougall, W., *Marketing & Communications & Germany Trade & Invest. Industrie 4.0*, Germany Trade and Invest, 2014.
34. Mueck, T. *Privacy Issues in the Austrian EHR Project "ELGA"*, Available at: [https://www.nii.ac.jp/issi/pdf/1/2Thomas\\_Mueck.pdf](https://www.nii.ac.jp/issi/pdf/1/2Thomas_Mueck.pdf), accessed: 2017-03-01.
35. P7000 – Model Process for Addressing Ethical Concerns During System Design, IEEE, 2016. Available at: <https://standards.ieee.org/develop/project/7000.html>, accessed: 2017-03-01.
36. P7001 – Transparency of Autonomous Systems, IEEE, 2016. Available at: <https://standards.ieee.org/develop/project/7001.html>, accessed: 2017-03-01.
37. P7002 – Data Privacy Process, IEEE, 2016. Available at: <https://standards.ieee.org/develop/project/7002.html>, accessed: 2017-03-01.
38. Pagels, M. *The Black Box of Deep Learning*, SC5, 2016. Available at: <https://sc5.io/posts/black-box-deep-learning/>, accessed: 2017-03-01.
39. Philippi, T. *The Electronic Health Record (ELGA) in Austria*, ELGA GmbH, 2015.
40. Preparing for the future of artificial intelligence. Executive Office of the President National Science and Technology Council Committee on Technology, 2016.
41. Reuters. Germany investigating unprecedented spread of fake news online, *The Guardian*, 2017. Available at: <https://www.theguardian.com/world/2017/jan/09/germany-investigating-spread-fake-news-online-russia-election/>, accessed: 2016-02-28.
42. Reynolds, M. AI learns to write its own code by stealing from other programs, *New Scientist*, 2017. Available at: <https://www.newscientist.com/article/mg23331144-500-ai-learns-to-write-its-own-code-by-stealing-from-other-programs/>, accessed: 2017-03-01.
43. Robots – Legal Affairs Committee calls for EU-wide, *European Parliament News*, 2017. Available at: <http://www.europarl.europa.eu/news/en/news-room/20170110IPR57613/robots-legal-affairs-committee-calls-for-eu-wide-rules/>, accessed: 2017-03-01.
44. Solon, O. World's largest hedge fund to replace managers with artificial intelligence, *The Guardian*, 2016. Available at: <https://www.theguardian.com/technology/2016/dec/22/bridgewater-associates-ai-artificial-intelligence-management/>, accessed: 2017-03-01.
45. Son, H. JPMorgan Software Does in Seconds What Took Lawyers 360,000 Hours, *Bloomberg.com*, 2017. Available at: <https://www.bloomberg.com/news/articles/2017-02-28/jpmorgan-marshals-an-army-of-developers-to-automate-high-finance/>, accessed: 2017-03-01.
46. Sparrow, R. Building a Better WarBot: Ethical Issues in the Design of Unmanned Systems for Military Applications, *Science and Engineering Ethics*, vol. 15, no. 2, 2009, pp. 169–187.
47. Special Report: The World's Largest Data Centers. *Data Center Knowledge*, n.d.. Available at: <http://www.datacenterknowledge.com/special-report-the-worlds-largest-data-centers/>, accessed: 2016-02-28.
48. Sverdlik, Y. Top Cloud Providers Made \$11B on IaaS in 2015, but It's Only the Beginning. *Data Center Knowledge*, 2016, Available at: <http://www.datacenterknowledge.com/archives/2016/06/07/top-cloud-providers-made-11b-on-iaas-in-2015-but-its-only-the-beginning/>, accessed: 2016-02-28.
49. Thurner, S. *Understanding co-morbidity through medical data on national scales*, International Conference on Synthetic Populations, Luca, 2017.
50. Toor, A. Germany grapples with fake news ahead of elections, *The Verge*, 2017. Available at: <http://www.theverge.com/2017/1/19/14314680/germany-fake-news-facebook-russia-election-merkel/>, accessed: 2017-03-01.
51. Top 20 Facebook Statistics – Updated January 2017, Zephoria Digital Marketing, 2017. Available at: <https://zephoria.com/top-15-valuable-facebook-statistics/>, accessed: 2016-02-28.

52. Tunkelang, D. Privatizing Our Past, *Hacker Noon*, 2016. Available at: <https://hackernoon.com/privatizing-our-past-148d7f5c2084>, accessed: 2016-02-28.
53. Universal Declaration of Human Rights. United Nations General Assembly, 1948.



## **Artifacts and The Problem of Ethical Extensionism – Selected Issues**

*Magdalena Holy-Luczaj*

University of Information Technology  
and Management in Rzeszow,  
Poland

*e-mail:* mholy@wsiz.rzeszow.pl

### *Abstract:*

This paper addresses the problem of extending ethical obligations toward usable things. The first part reconstructs current debates on the metaphysical and ethical status of artifacts. Next, drawing upon Tadeusz Kotarbinski's reism, I describe artifacts as concretes, focusing on the possibility of their damage and destruction. The core part of the article analyzes ethical implications of the following issues: 1) using artifacts, 2) their production, 3) purchase and sale of artifacts, and 4) the post-use period.

*Keywords:* ethical extensionism, artifacts, postnatural extensionism, environmental philosophy, philosophy of technology.

## **1. Introduction**

The debate about expanding ethical obligations of human beings toward nonhumans has been going on for several decades. Arguments for including different animal species, plants, other living organisms, and finally inanimate entities in the boundaries of ethics have successively appeared in numerous works dealing with this issue. Now – as I attempt to show – it is worth establishing possible grounds for extending human moral obligations toward artifacts as beings which status can be improved or deteriorated through human activity.

Usable things, on which this paper concentrates, belong to a subset of technical artifacts. It consists of single objects created by humans, which have a determined material identity and serve established functions. Hence, from this classification are excluded electronic/digital beings (such as a computer software), since they are not associated with particular, physical carriers.

The above specification is necessary to focus on problems, which are situated at the intersection of environmental philosophy and philosophy of technology, such as production and use of usable things, or managing them after utility period. I shall present them in the context of regarding usable things as moral patients.

## **2. Emancipation of Artifacts**

Philosophy for centuries, actually from its very beginnings, disregarded technical artifacts [46, pp. 1–2], [2, pp. 99–100], [34, pp. 9–10]. It was not until the last few decades when interest in them

increased significantly. It can be observed both in the tradition of analytic and continental philosophy, as well as in the currents trying to cross this divide.

## 2.1 Analytic Philosophy

Studies on artifacts diverged in analytic philosophy as a part of its reflection on technology [46, p. 4] in order to do justice to artifacts as fully-fledged beings [15, p. 3], [2, p. 100]. Ontological emancipation of artifacts is usually justified using one of two strategies. On the one hand, scholars attempt to describe artifacts as generally identical with natural beings, presenting differences between them as secondary for their identity [3, p. 60]. On the other hand, artifacts are portrayed as completely different from natural beings, which, however, does not mean that they are of inferior status [16, pp. 123–124], [47, p. 11]. Many philosophers argue here with Aristotle [16, p. 124], [3, p. 50], [47, pp.7–8], [46, p. 29], who did not consider artifacts (beings which are produced) to be genuine substances, even though they are particular individuals (there is a significant difference in this regard between his *Categories* and *Metaphysics* [22, p. 13], [2, pp. 104–105]). According to him, artifacts are flawed ontologically, because, in contrast to natural beings, they do not have the principle of origin in themselves, but this principle is located in man as their creator.

In analytic philosophy, there are basically two main determinants of the identity of artifacts. The first one is their non-naturalness [15, p. 123]. Yet, currently, many philosophers highlight the problematic division of beings to the natural and artificial in the context of the development of science and technology and their interconnections [3, p. 64], [6]. Moreover, it is argued that if human beings are natural entities, their products are natural as well [see 3, p. 65]. The second one is that artifacts are functional objects [15, p. 123]. This characteristic is often even identified with the essence of the artifact. As Houkes and Vermaas point out a large number of artifacts is named in functional terms, such as ‘screwdriver’ [15, p. 124]. However, recently there is more and more visible tendency toward examining artifacts not by their functions, but by the way they are produced and used [14, p. 52]. Therefore, the current definition of artifacts is well framed in the expression “produced-to-use”. It means that technical artifact is the material object with a certain function and a use plan, which was designed and constructed by human agents. It is of particular importance that this definition highlights that the agent, who constructs the artifact is not a single person who produces relatively simple artifacts, starting with a (perhaps personal) desire and ending with a finished product (“the artisan”), but this is a team of designers, engineers and producers [16, pp. 403–404].

One can obviously and reasonably recognize that the question of whether the artifact is produced by an artisan or whether it is the product of engineering process does not make any difference to the nature of the artifact (what counts is how and whether the artifact is actually used) and as such it is insignificant [16, p. 414]. Yet, it gains considerable importance in the context of the morality artifacts as opposed to the thesis of their neutrality. According to the latter, artifacts are not value-laden, but at most they constitute means or requisites used in activities that can be assessed in ethical terms. For this reason – some scholars hold – we cannot acknowledge that the work of designers and engineers may be subject to ethical evaluation [47, p. 16], [34, pp. 71–109]. Critics of this belief, however, indicate that some of the artifacts to a greater extent than other things can do harm to people and the environment, so they cannot be regarded as ethically neutral entities. In such a case production of these artifacts must be assessed in ethical categories. The main question is how engineers and designers can change the world for better or worse by creating specific artifacts [16, p. 417].

There is also another standpoint, according to which an artifact itself is a moral agent. In such a perspective artifacts are no longer merely passive instruments used for good or bad purposes, but they are seen as something that actively shapes human existence [34, pp. 123–137]. It is due to the fact that some decisions and activities are delegated to devices, an example of which can be speed control device in cars [46, p. 213]. However, there were numerous allegations against this concept, e.g. establishing technocracy, which could lead to totalitarianism. A second objection to

assigning morality to things concerns the fact that artifacts are not able to make deliberate decisions about their influences on human action. Things lack intentions, runs this objection, and thus cannot be held accountable for their behaviors [46, pp. 213–214]. Some scholars, however, believe that things carry morality because they shape the way people experience their world and organize their existence, regardless of whether this is done consciously and intentionally or not. Accordingly, the ethics of artifacts in analytic philosophy is understood as research on *moral agency* of artifacts. The possibility to recognize artifacts as *moral patients* is neglected or treated inexplicably as secondary (as David Gunkel shows with respect to robots [12, pp. 119–124]). The notion of *moral considerability* refers in this case to the fact that production of artifacts requires taking into consideration various factors [13, p. 227].

## 2.2 Continental Philosophy

Continental philosophy examines the issue of artifacts assuming that human beings and things are essentially different, which, however, by no means negates their bonds [see 46, p. 19]. The leading role is played by the interpretation and critique of the philosophy of Martin Heidegger. Some researchers analyze the early period of it, which belongs to the existential paradigm and concentrates on the human understanding of ways of being. Heidegger contended that the most intuitive is the one which can be grasped in the attitude which humans have toward tools, that is beings which are situated in the context of their functions and mutual belonging [7, p. 111]. On the other hand, numerous studies are devoted to the later philosophy of Heidegger, which is a project of “another thinking” that attempts to go beyond anthropocentrism. It assumes that we should think of tools as things, which are the spots for “enowning” – happening of being. According to Heidegger, things disclose themselves in their particularity. However, it is hard to notice this in the optics of modern technology, in which they are merely elements of “the resource” [see 5].

Ethical issues related to artifacts in continental philosophy are under the great influence of Heidegger’s thought as well, which is linked with concepts inherited from other important thinkers belonging to this tradition, such as the idea of otherness (the Other) brought by Emanuel Levinas [5], [21]. In result, the most significant categories of continental ethics of things (which can be seen as a consideration of the human relations with things) are respect for being and otherness, as well as care recognized as an existential phenomenon.

Continental tradition, however, is also developed in the more pragmatic direction. An example of this approach is postphenomenology of Don Ihde, widely regarded as one of the leading philosophers of technology. Although Ihde complains that Heidegger’s concepts are technophobic and entangled in romantic cliché which idealize premodern reality, he thinks that it is worth to draw upon Heidegger’s idea of worldliness as something toward which human beings orient their cognition, intentionality as the basic structure of our cognition, and technology as the filter, which modifies perception of a reality [see 20, pp. 103–15]. Ihde in his research concentrates on artifacts understood as objects, which mediate human knowledge and action, and understands this mediation as a form of activity of things, which can be subject to ethical evaluation [19].

## 2.3 Posthumanism

Another interesting trend is a pursuit to bring together continental philosophy with the perspective of science and technology studies. Scholars employ here the concept of actants, that is nonhuman actors, created by Bruno Latour [see 21]. According to Latour, we cannot understand reality if we downplay the role of nonhuman beings in lives of human beings because we cannot separate these two groups – they are intrinsically linked in the network of mutual relations. Furthermore, belonging to this network is the condition of gaining their proper identity and possibility to act. For this reason, Latour holds, we should not speak of human and nonhuman actors, but rather actants, which are natural beings (to include virtues, climate, species), technological infrastructure, various

objects from our everyday surroundings, law, social institutions, as well as ideas and values [6, pp. 158–159].

The desire to treat humans and nonhumans “symmetrically” culminates in the posthumanism. Keeping in mind that this is an umbrella term for a wide variety of projects developed in the last three decades [see 6, pp. 154–155], it can be concluded that one of the key postulates of posthumanism says that attitude reserved hitherto exclusively for humans (respect, attention to the preservation of welfare, avoiding harms) should be extended to all beings in the widely understood environment (which constitute natural animate and inanimate beings as well as artificial beings) and not only *homo sapiens* [8, p. 31].

Posthumanists highlight indivisibility of metaphysical universe, in which all actors/actants interact with each other, what means that also things affect humans. Posthuman theorists argue that taking away agency from nonhuman beings supported human hubris, which translated into an arrogant way of dealings with beings other than human [6, p. 155], [4, p. 12]. Accordingly, in posthumanism, the ground for taking artifacts into ethical consideration is primarily their moral agency.

## **2.4 Postnatural Environmentalism**

The problem of moral patiency of artifacts is important for the postnatural environmentalism. This approach criticizes traditional environmentalism for excluding built environment, humanized landscape and artifacts from the scope of its interest and being focused solely on natural beings [50, pp. 2, 88], [48], [49], [44]. In its pursuit to expand the reach of moral considerability beyond traditional limitation to humans (traditional) environmental philosophers include nature to it, but they do not worry about things which environ us on a daily basis: buildings and useful things [50, pp. 2, 88]. Such attitude toward artifacts has, as postnatural environmentalists show, negative effects. First, it brings a sense of alienation in human beings, because it assumes that their creations are different from natural ones and thus they do not deserve moral attention. Second, it results in the thoughtless and irresponsible way of dealings with artifacts. Postnatural environmentalists want to change it by enhancing the metaphysical and ethical status of artifacts.

It is of crucial importance to emphasize that postulate to include artifacts into ethics does not intend to devaluate natural entities nor to ground normative anthropocentrism (subordinating functioning of the entire ecosystem solely to human needs). On the contrary, Steven Vogel suggests that we should rather stop unreflective exploitation which tries to satisfy only human needs [50, pp. 137, 163]. At the same time, he underlines that his analysis has an introductory and preliminary character in ascribing moral patiency to artifacts and requires further research [see 50, p. 164]. I do believe it is worth taking them. In what follows, I map the most important areas which should be explored.

## **3. Moral Considerability, Patients and Agents**

Addressing the problem of ethical obligations toward usable things requires reconstructing conceptual framework and a detailed analysis of categories, which will be employed in such a project. The first one is moral considerability, which is a capacity to absorb moral consideration. That is to say, it is belonging to the class of individuals, which deserve moral consideration (assessing actions undertaken toward them in terms of ethics), even though such beings are not able to morally assess their own actions. This restriction is compatible with the distinction between moral patients and moral agents. Moral patients are beings to whom we consider that we owe moral obligations, and agents are those who are held to be morally responsible for their actions. What is of particular importance, all moral agents are moral patients, but not all moral patients are agents (as it is for example in the case of mentally impaired persons or little children) [10].

Many ecophilosophers refers to this division when they argue for shifting boundaries of ethics in order to include natural entities. Ethics of things can identify strategies, which they

employed in extending ethics towards various kinds of natural beings and critically assess whether they can be modified and adopted in attempt to take artifacts into moral consideration

Furthermore, it would be indispensable to examine the problem of egalitarianism of ethical extensionism, answering the question whether an extension of ethical consideration necessarily means that the interests of all beings should be considered on an equal footing. I will present various scenarios regarding graduality and egalitarianism of moral considerability concerning different entities and the possible principles that one can follow in case of conflict of life interests (distinguishing them from the so-called secondary and tertiary needs [see 9, p. 15]).

Another issue concerning ethical considerability, which should be explored, is the relationship between metaphysics and ethics. Theorists of moral consideration assume that assessment whether some being deserves it, depends on what is this being (that is to say, whether it has all necessary characteristics qualifying it to be considered ethically [see 13, pp. 28–39]). A close relationship between metaphysics and ethics was indicated directly both by environmental philosophy and posthumanism. Unfortunately, their representatives do not elucidate in more details the nature of this connection [see 27, p. 99], just like analytic philosophers quoted earlier. In my opinion, this connection can be explicated by referring to the concept of “idea” presented by Arthur Lovejoy, who recognized the idea as the disposition to think with the use of categories specific to the cultural context, which moves human imagination, behavior, and emotions. Following this path, it can be argued that metaphysics should not be treated as a logical justification of ethics, but metaphysical assumptions (metaphysical “ideas”) can be considered as a component of the motivation of human behavior.

The key question is as follows: what metaphysical image of usable things can induce us to include them in the scope of ethical reflection? Certainly, there are many possible solutions. My suggestion is to present useful things in accordance with well-rooted in Polish philosophical tradition Tadeusz Kotarbinski's ontology as “concretes.”

#### **4. Concretes**

Tadeusz Kotarbinski defined “concretes” in conformity with the common usage of the word “thing” as “objects located in time and in space and having certain physical characteristics” [30, p. 233]. Concretes, according to Kotarbinski, are humans beings, animals, plants and inanimate beings – both natural and artificial [31]. In the course of the development of his metaphysical theory, Kotarbinski decided that the notion of “concretes” is identical with terms “body” and “thing”, but it avoids difficulties related to those names – (“body” is usually used not in the sense of physics, but biology, where it is interchangeable with “organism” and is limited to livings beings, whereas “thing” is usually associated with inanimate entities, inorganic solids [see 32, p. 29]. In the context of the problem of the metaphysical status of usable things, it is important that Kotarbinski not only rejected dualism of animate and inanimate beings, but also resigned from strict separation of natural and non-natural beings – he listed a watch as an example of a concrete [31, p. 283].

Drawing upon Kotarbinski's theory it can be shown that usable things are as real as natural entities are and therefore the thesis that they do not actually exist is counterintuitive [2, p. 64], [24, p. 93]. Furthermore, concretism inclines us to see metaphysical universe rather nonhierarchically, what can support rejecting belief in ontological subordination of artifacts [2, p. 60]

In the characteristics of usable things, a premium should be put, however, on the fact that they – as particular individuals – can be irreparably damaged and destroyed. Even if the usable thing remains a material object after the destruction, it no longer has its integral identity, or more precisely, it has no longer a possibility to fulfill it. It is this immanent feature of usable things that makes it necessary to separate them from the broader group of technical artifacts, which includes, for example, computer software, which has significantly different material status (most researchers ignore this fact [2, p. 49], [47, p. 5], [see 34, p. 38]. The difference between physical beings and electronic beings consists in that the latter, unlike physical beings, are perfectly multiplicable. For digital-beings, there is no distinction between originals and duplicates. As a result, a single digital-

being literally can exist at multiple locations at the same time and can endure the test of time, as Joohan Kim underlines [28, p. 99]. Moreover, a digital-being is far less likely to be irreversibly destroyed.

It is worth highlighting that such a destruction of usable things can be caused by human activity. The possibility of irreversible destruction of a particular entity by human actions should move human ethical imagination and at the same time can define boundaries of moral considerability. Distinguishing between beings which can be hurt or benefit thanks to human activity and beings which can be damaged or repaired I find vague and hard to defend [1, pp. 31–32], [18], [11].

Likewise, lack of ability to self-maintain and self-repair should not disqualify artifacts from the moral domain. Many environmental philosophers argue that a difference in the degree of complexity (as, for example, between people and animals) should not automatically eliminate certain beings from the ethical domain. Yet, it appears that for artifacts these environmental thinkers do not adhere to this rule. In this case, they even fall into the trap of recognizing something as “mere.” This is a very dangerous pattern of thinking, because it gives us a kind of alibi. If something is seen as insignificant, we can deal arbitrarily with it. We can excuse ourselves from responsibility for such things – for “mere things” [18] – and this is one way environmental philosophy sees artifacts. For this reason, environmental philosophy does not do the right thing, when it grants us the right to be thoughtless about beings which constitute our everyday environment, even if they are less complex than other beings [see 50, p. 163].

Moreover, the identity of artifacts is not so poor as some environmental philosophers hold. Artifacts, from the moment they are produced, are autonomous in relation to human beings – they deteriorate, decay, refuse to work, and have unexpected applications. In short, human intention does not determine them completely [50, pp. 105, 113], [14, pp. 52–55], [3, p. 50], [46, p. 29].

It seems then that it not so absurd to challenge alleged obviousness of embargo on taking usable things into ethical consideration. Revisiting ontological status of usable things is of crucial importance for identifying the most serious dilemmas related to them, which we have to face. They can be divided into four groups, respectively linked with: using, producing, sale and purchase, after-use period and threats.

## 5. Using

The problem of moral considerability of usable things requires us to examine in the first place activity, which most commonly leads to the destruction of the thing: using. It is using which contributes to deteriorating of the thing (loss of certain qualities) and eventually its destruction (that is the “using up” of a thing). Subsequent attempts to regenerate the thing only delay this effect.

Furthermore, using, as theories of intrinsic value assume, introduce oppression and hierarchy – an entity, which uses another entity, makes it subordinate and disavows its ontological status. Due to this intrinsic value is often identified with non-instrumental value [see 9, p. 14].

Such an approach is another reason, why artifacts are excluded from the scope of ethics by environmental philosophy. If they have only usable character (what is even reflected in their name), there is no rescue for them – they are doomed to be exploited and disregarded. Natural beings, on the other hand, have different scenarios of their existence, mainly undisturbed development.

Such a belief, however, suggests a peculiarly atomistic picture of the world, which undermines the attempt of environmental philosophy to convince people to recognize the interdependence of the different parts of the natural the world. This contradiction is highlighted by environmental philosophers who are skeptical about the identification of the inherent or intrinsic value with the non-instrumental value [39], [36]. Using can be seen as a kind of cooperation, in which entities disclose various aspect of their identity. In order to reject the strict opposition of the inherent and instrumental values, it will be necessary to develop the argument that if some entity is not used, this does not mean that it becomes meaningless or it does not have any longer a right to exist, but it refers to the inability to manifest and fulfill its identity.



Revision of assessment of the phenomenon of use is crucial for our perception of artifacts – their usable character does not longer undermine the ethical dimension of dealing with them. Hence, it would be essential to ask if it is possible to achieve a balance between enabling things to realize their potential and protecting them from decay. Answering this question is one of the most fundamental tasks of the ethics of things. The starting point should be distinguishing between “proper use” and “exploitation” as suggested by Heidegger in *Anaximander’s saying* and *What is Called Thinking?*. We can expand on this concept by indicating when using the thing becomes improper and unacceptable from the perspective of its integral identity [see 14, pp. 52–54]. In this regard, it will also be important to distinguish between alternative and standard variants of the proper use [see 14, pp. 52–53].

## 6. Production

Except for already discussed issues, there is also one more which plays a significant role in excluding artifacts from the sphere of ethics by environmental philosophy – it is a fact of being produced by human beings. Just as Aristotle, so many environmental thinkers hold that the way some being occurs is decisive for its status and ontological autonomy [33, p. 49]. It can be argued whether more important is the way of creating or the final result of it, which in case of artifacts is concrete individuum [see 46, p. 29]. Yet, for our considerations, two problems would be the most important: being secondary to the material from which something has been produced and replaceability.

Many environmental philosophers highlight that artifacts, again unlike naturally-occurring beings, are secondary to the material, which they were made from [see 13, p. 50]. For example, a tree is not derivative of wood, while a wooden chair is. There are of course degrees of artefacticity in this regard. The material can be natural (e.g., wood), can be derived from natural material (e.g., plastic as made from oil), and can be constructed *de novo* (e.g., diamondoid material) [see 13, pp. 49–52]. The question of the material which is needed to create an artifact results in perceiving artifacts mainly through the prism of the use of natural resources or seeing them as difficult to recycle waste, which pollutes the environment. Such an approach is focused on the fact that artifact is produced at the expense on some natural being. Without a doubt, it cannot be denied that to produce artifacts we need to use some (more or less) natural materials. However, in nature, everything seems to be at the expense of other beings (wolves eating sheep, sheep eating grass, etc.). Do we see only this dependence when we think about those beings? Rather no. If we did, nature would be for us nothing more than a Manichaeian sphere of cruelty and darkness. Likewise, if we perceive artifacts only as something that requires using natural resources, we cannot have a positive attitude toward them, and this translates to specific manners of dealings with them (often undesirable from the point of view of ecology).

By the same token, positive attitude towards artifacts is disabled by the belief that artifacts, in contrast to natural beings, are replicable and replaceable because we can always produce another artifact. Such a belief is often expressed in the debate about ecological restoration. Some environmental thinkers are reluctant about it, because they believe it is a false impression that we can restore beings destroyed by humans. Critiques of ecological restoration emphasized that newly planted trees will never replace the concrete trees, which ceased to exist as a result of human activity [see 24, p. 72]. Not denying this claim, I believe that environmental philosophy exhibit unjust disregard for artifacts, because even though they are (mass) produced every artifact is always a particular individual, and as such is a unique entity, which cannot be replaced.

## 7. Purchase and Sale

Dismissing image of artifacts as replaceable has a significant meaning in the analysis of ethical dilemmas related to their purchase and sale. For example, in the case of sale of artifacts, I will, for example, try to answer whether the fact that it consists in selling new things justifies all sale

practices (how to assess a situation which leads to the purchase of several exemplars of the same artifact that do not have even a chance to be used). Analyzing ethical aspects of the purchase of new goods, I will consider if we have any liabilities to things we already own. This issue will be important to resolve the question what actually justifies purchase of a new thing, resignation from repairing the old one, etc. consider increasingly emerging problem of producing things which have factory-provided short term of performance (which broke just after the expiration of the guarantee), and their repair is more expensive than buying a new product [see 43].

Recognition of artifacts as moral patients will also encourage a critique of consumerism as not related to real needs purchase of new things, which are not perceived as particular individuals, but as signs of wealth, keeping up with fashion and so on. I will refer to the slogan introduced by deep ecology that consumption by improving the standard of life, deteriorates its quality [see 35, p. 222] and show that consumerism does not allow us to create bonds with things we possess. At the same time, my analysis will distance itself from minimalistic attitude, which calls for reducing the number of things we have, because they are seen only as a threat to human freedom. I will argue that such an approach should be replaced by the pursuit to create a positive attitude to things we own (having a sense of connection with them), which should more effectively support long-term and prudent using of artifacts, and in result counter the threat of consumerism.

## **8. Post-use Period**

Another important field of ethics of things should be issues related to the period when a usable thing ceases to be able to perform its default function. They can be discussed with reference to the ontology of trash [see 26], which focuses on the transformation of things in the so-called utility and waste. Ethics of things should concentrate on situations, which enable things to fulfill their potential after the standard period of use. These will be basically two phenomena. First one will be recycling, that is a process of converting waste materials into objects, which can be reused for the original purpose (glass bottles) or other goals (making notebooks from wastepaper). The second phenomenon will be a creative reuse, referred to also as repurposing. It consists in restoring things their functionality (restored furniture) or giving them a new form (e.g. making poufs or garden pots out of used tires). I will present these phenomena as something that create opportunities for things to complete the realization of their potential and identity and not only as something that enables people to limit the use of natural resources and reduce production of waste.

## **9. Threats**

Including artifacts in the scope of moral considerability requires us to examine negative issues related to the existence of certain artifacts. For example, it should be considered what we should do with objects whose standard function is killing (e.g., atomic bombs) or which pollute the air (e.g., old, inefficient cars). It seems we can adopt an analogical strategy as environmental philosophy, which argued that when some being (e.g., virus) is a direct threat to the existence of another being or when it generates only harms to others, they should be eliminated or at least neutralized. By the same token, usable things which are dangerous to people, nature, and other things should be destroyed or, as far as possible, recycled or upcycled.

## **10. Conclusion**

This presentation of the possibility of extending moral obligations toward usable things is only a sketch of the problem, which requires further analysis and development. The aim of it was to draw attention to possible, an alternative approach to artifacts. Many environmental thinkers perceive artifacts only as a threat to the environment and therefore postulate limiting their consumption. Meanwhile, as I have argued, it is worth trying to engage positively with things and recognize them as concrete individuals and as such worth ethical concern, what should also result in the reduction



of consumerism and the production of waste. Such an approach does not intend to undermine other perspectives, but to supplement them. I believe that we start to understand eco-sphere in its original meaning as a household (gr. *oikos*), only if we take into considerations artifacts as its elements.

## References

1. Attfield, R. *Beyond Anthropocentrism*, In A. O'Hear (ed.), *Philosophy and the Environment*, Cambridge 2011.
2. Baker, L.R. The Ontology of Artifacts, *Philosophical Explorations*, 7(2), 2004.
3. Baker, L.R. *The Metaphysics of Everyday Life*, Cambridge: Cambridge UP, 2009.
4. Bennett, J. *Vibrant Matter: A Political Ecology of Things*, Durham – London: Duke University Press, 2010.
5. Benso, S. *The Face of Things: A Different Side of Ethics*, New York: SUNY, 2000.
6. Bińczyk, E. Posthumanist Tendencies in Science and Technology Studies, *Political Dialogues. Journal of Biopolitics and Contemporary Political Theories*, 15, 2013.
7. Blok, V. Heidegger's Ontology of Work, *Heidegger Studies*, 31, 2015.
8. Braidotti, R. *The Posthuman*, Malden: Polity, 2013.
9. Devall, B. *Simple in Means, Rich in Ends. Practicing Deep Ecology*, Salt Lake City, 1988.
10. Goodpaster, K. On Being Morally Considerable, *Journal of Philosophy*, 75 (6), 1978,
11. Goodpaster, K. On Stopping at Everything: A Reply to W. M. Hunt, *Environmental Ethics*, 3(3), 1980.
12. Gunkel, D. A Vindication of the Rights of Machines, *Philosophy & Technology*, 27 (1), 2012.
13. Hale, B. Technology, the Environment and the Moral Considerability of Artefacts, In J.K.B. Olsen, E. Selinegr & S. Riis (eds.), *New Waves in Philosophy of Technology*, New York: Palgrave Macmillan, 2009.
14. Houkes, W. & Vermaas P.E. Actions versus Functions: A Plea For an Alternative Metaphysics of Artifacts, *Monist*, 87, 2004.
15. Houkes, W. & Vermaas P.E. Produced to Use: Combining Two Key Intuitions on the Nature of Artefacts, *Techne*, 13 (2), 2009.
16. Houkes, W. & Vermaas P.E. Contemporary Engineering and the Metaphysics of Artefacts: Beyond the Artisan Model, *The Monist*, 92 (3), 2009.
17. Hullinger, J. *Ikea Wants You To Stop Throwing Away Your Ikea Furniture*, [online:] <http://www.fastcompany.com/3055971/most-creative-people/ikea-wants-you-to-stop-throwing-away-your-ikea-furniture>, 2015.
18. Hunt, W. M. Are Mere Things Morally Considerable? *Environmental Ethics*, 2(3), 1980.
19. Ihde, D. *Expanding Hermeneutics*, Evanston: Fordham UP, 1998.
20. Ihde, D. *Postphenomenology*, Evanston: Fordham UP, 1993.
21. Introna, L. Ethics and Speaking of Things, *Theory, Culture & Society*, 3, 2009.
22. Katayama, E. *Aristotle on Artifacts: a Metaphysical Puzzle*, Albany: SUNY, 1999.
23. Katz, E. Understanding Moral Limits in the Duality of Artifacts and Nature: A Reply to Critics, *Ethics & the Environment*, 1, 2002.
24. Katz, E. Further Adventures in the Case Against Restoration, *Environmental Ethics*, 1, 2012.
25. Katz, E. Artefacts and Functions: A Note on the Value of Nature, *Environmental Values*, 2, 1993.
26. Kennedy, G. *An Ontology of Trash. The Disposable and Its Problematic Nature*, Albany: SUNY, 2012.
27. Kiepas A. *Człowiek wobec dylematów filozofii techniki*, Katowice: Wyd. UŚ, 2000.
28. Kim, J. Phenomenology of Digital-Being, *Human Studies*, no. 24, 2001.
29. King, R. 2002. Environmental Ethics and the Built Environment, *Environmental Ethics*, 22(2), 2002, pp. 115–131.
30. Kotarbinski, T. Fazy rozwojowe konkretyzmu, *Studia Filozoficzne*, 1, 1958, pp. 228–236.
31. Kotarbinski, T. Zasadnicze myśli pansomatyzmu, *Przegląd Filozoficzny*, 33, 1934, pp. 283–294.

32. Kozłowski, J. O reizmie somatycznym Tadeusza Kotarbinskiego, *Studia Humanistyczne Akademii Górniczo-Hutniczej. Prace Filozoficzne*, 27, 1985, pp. 28–39.
33. Lee, K. *The Natural and the Artefactual. The Implications of Deep Science and Deep Technology for Environmental Philosophy*, Lanham – Boulder – New York – Oxford: Lexington Books, 1999.
34. Lizut, R. A. *Technika a wartości. Spór o aksjologiczną neutralność artefaktów*, Lublin: Academicon, 2014.
35. Maskit, J. *Deep Ecology and Desire: On Naess and the Problem of Consumption*, In E. Katz, A. Light, D. Rothenberg (eds.), *Beneath the Surface. Critical Essays in the Philosophy of Deep Ecology*, Cambridge – London: MIT Press, 2000.
36. Morito, B. Intrinsic Value: A Modern Albatross for the Ecological Approach, *Environmental Values*, 12: 2003, pp. 317–336.
37. Naess, A. The Deep Ecology Movement, G. Sessions (ed.), In *Deep Ecology for the Twenty-First Century*, Boston – London, 1995.
38. Newman, L. & Dale, A. Celebrating the Mundane: Nature and the Built Environment, *Environmental Values*, 22: 2013, pp. 401–413.
39. Norton, B. Convergence, Noninstrumental Value and the Semantics of ‘Love’: Comment on McShane, *Environmental Values*, 17, 2008, pp. 5–14.
40. Ouderkirk, W. Katz’s Problematic Dualism and its “Seismic” Effects on His Theory, *Ethics & the Environment*, 7(1), 2002, pp. 124–137.
41. Piątek, Z. *Aspekty antropocentryzmu*, Kraków: Wyd. UJ, 1988.
42. Piątek, Z. *Etyka środowiskowa. Nowe spojrzenie na miejsce człowieka w przyrodzie*, Kraków: Księgarnia Akademicka, 1998.
43. Przybysz, A. Raport: czas użytkowania lodówek i telewizorów coraz krótszy. Wina producentów? *Gazeta Wyborcza*, 15.04.2015.
44. Thompson, I. Landscape and Environmental Ethics, In P. Howard, I. Thompson, & E. Waterton (eds.), *The Routledge Companion to Landscape Studies*, London and New York: Routledge, 2013.
45. Tymieniecka-Suchanek, J. (ed.). 2014 *Człowiek w relacji do zwierząt, roślin i maszyn w kulturze. Od humanizmu do posthumanizmu*, t. 1–2, Katowice: Wyd. UŚ, 2014.
46. Verbeek, P.-P. *What Things Do. Philosophical Reflections on Technology, Agency, and Design*, University Park: Penn State UP, 2005.
47. Vermaas P., Kroes P., & Franssen M. *A Philosophy of Technology: From Technical Artefacts to Sociotechnical Systems*, Morgan & Claypool Publishers, 2011.
48. Vogel, S. Environmental Philosophy after the End of Nature, *Environmental Ethics*, 24(1), 2002.
49. Vogel, S. The Nature of Artifacts, *Environmental Ethics*, 25(2), 2003.
50. Vogel, S. *Thinking like a Mall*, Cambridge: MIT Press, 2015.

**Salutary Meanings of Sublimation.  
Selected Soteriological Threads of Alienation  
Criticism of Religion<sup>1</sup>**

*Jowita Guja*

AGH University of Science  
and Technology in Cracow,  
Poland

*e-mail:* jowitaguja@poczta.fm

*Abstract:*

The article concerns a few soteriological threads of alienation criticism of religion whose feature is the creation of a new autonomous and transgressive subject. It focused on the presentation of this subject using Nietzsche's philosophy perceived within Freudian perspective.

*Keywords:* atheism, soteriology, alienation, sublimation, repression.

## 1. Introduction

Speaking about soteriology in the area of atheistic philosophy may seem disputable as a term “salvation” is typically reserved for religious traditions. Contrary to this trend, I am proposing a thesis that salvation of a man can have a completely secular meaning. Literally, it can be understood as liberation from a negative state or negative circumstances. That is the comprehension found within the Old Testament and in the Judaic tradition. In Christianity, in turn, salvation has several meanings, out of which restoration of lost inner oneness or completeness to a human being seems to be the most important. We deal with such a perception of salvation in, among others, the Letter to the Romans by St. Paul, within which a split of man into an old (outer) one and a new (inner) one is mentioned and discussed; the cleavage that is to be overcome in Christ. In atheistic philosophy, one can come across a certain analogy to both of these senses while simultaneously reversing the religious belief itself. On the one hand, there is a Promethean theme in it with its matter of liberation from a certain negative state: a specific sin of culture – in this case, the sin is equal to a projection of God and to negative consequences of this activity. On the other hand, some atheistic trends tell of a crack in a man that is to be filled. This is precisely the purpose of alienation criticism of religion. However, the cause of the break in a man is accounted for a mechanism of religious alienation. The secular salvation is, in this case, consists in the restoration of man's lost unity and their introduction into a path of authentic existence, which they are devoid of in religion.

Depending on an atheistic trend with which we are dealing, numerous soteriological threads may be mentioned. All consistent plots of thought are characterized by Prometheism associated with an idea of liberation by the truth. Besides, there may occur themes such as salvation through a

close to mystical experience of a miracle of existence or an aesthetic experience (which is characteristic for New Atheism, for example of Richard Dawkins's or Sam Harris's), salvation by action, transformation of social reality and liberation from suffering (noticeable, among other things, in Russian philosophy, for example in Vissarion Belinsky's thought or in revolutionary atheism) and finally salvation by creation of a new subject. The last thread itself is precisely the main topic of this article. It is one of the factors determining peculiarity of alienation criticism of religion, whose soteriology focuses mostly on trying to create a certain existential, authentic and transgressive ideal. At the same time, the authentic existence constitutes the opposite of the existence that has been alienated by religious projection.

## 2. Alienation Criticism of Religion – Description

Alienation criticism of religion represented among other things by Ludwig Feuerbach, Karl Marx, Friedrich Nietzsche, Sigmund Freud, Jean-Paul Sartre or Slavoj Žižek is an extremely strong standpoint in a discourse opposing to religion. All its arguments focus on indicating the effects of a projection of an idea of God within the constitution of human subjectivity. It is symptomatic that it does not enter into discussions with any proof of God's existence, there is no place here for this kind of polemics with supporters of a theological hypothesis. As opposed to New Atheism, represented among for example by Richard Dawkins or Daniel Dennett, it also rarely focuses on such aspects of religion as the generation of political conflicts. For alienation criticism of religion, the consequence of God projection for human subjectivity and how this projection influences it is much more important. Social phenomena related to religion are rather depicted as a result of the alienation of subjects that create culture.

The name "alienation trend" which I have adopted, indicates the fact that an idea linking all its representatives is a concept of subject alienation – a belief that through religion a man simply cuts off from a certain valuable part of their own being, objectifies it and treats it as a strange one. This problem seems to be paradigmatically articulated in Feuerbach's words: 'Religion is a relation of a man to their own essence. Its truth and healing moral force are found within this relation. However, religion is the relation to the essence treated not as its own but as the other own, not only different but even opposite' [2, pp. 334–335] Feuerbach is recognized as a father of modern atheism, and he is often given credit for introducing a scheme of religious alienation: a man projects his best qualities in the other world, cuts off them, and finally pays homage to them. This activity results in a false self-image: a cult of God's projection involves humiliation of a human being. In fact, a concept of alienation within various shades of meaning – from a totally secular one in a sense of alienation of a man from neighbors or social institutions, through metaphysical alienation of a Spirit that is performing its own self-objectivization to alienation in an entirely religious sense – which was described in a chapter of *Phenomenology of Spirit* concerning unhappy consciousness – gained significance in Hegel's thought. He writes about it using both: a notion *Entäusserung* to mark the life of the Spirit that externalizes itself objectifying itself in nature and culture, and a notion *Entfremdung* to mark unhappy consciousness which is usually identified with the Judeo-Christian consciousness. Feuerbach almost literally borrows the last term from him. It is indisputable it was him who popularized the notion within atheistic philosophy influencing consecutive critics of religion who then repeatedly followed a schema of alienation, modifying its meaning a bit each time. We find it in Marx's thought in an almost identical form and in a similar form in writings of Nietzsche who uses a term *entfremdete Geist* as well as he talks about *alteration* that takes place within religion. Freud and Žižek put the notion into words of psychiatry and psychoanalysis, Sartre speaks of an alienating, perverse, divine look. The ambiguity of the term "alienation" is also worth mentioning. It is used in numerous contexts within the topic discussed: beyond a classical Feuerbach meaning, one can notice among other things alienation of resentment subject (Nietzsche) that does not recognize actual reasons of their behavior and perception or

cynical subject (Žižek) alienated by declarations and rituals that are accompanied by a frivolous and reserved attitude. The common denominator of these concepts is understanding of alienation of a subject as a situation within which the subject does not recognize certain aspects of their own existence as actually their own or it perceives their false image, thereby remaining in an inauthentic attitude.

The alienation criticism of religion owes to Hegel a lot more than just defining a key argument against religion. At the same time Hegelian phenomenology is a starting point and a counterpoint of an original perspective of study of human subjectivity – the archaeological perspective. The essence of this perspective is very well outlined by Paul Ricoeur's essay *Freud and Philosophy: An Essay on Interpretation*. This French thinker analyzes a thought of one of the representatives of alienated criticism of religion, Sigmund Freud. He treats it as a kind of hermeneutics exploring a key idea of the psychoanalyst in question: the subject is not what in fact they are. In other words, the point of view of consciousness is usually a false point of view. To achieve the truth of the subject, we must adopt a perspective of reductive hermeneutics, demystifying, reaching deep into the self-unconscious subject, in their *arche*. To all appearances, it seems that this hermeneutics is exactly the inverse of Hegelian phenomenology, which performs a reverse movement – it is not interested in the *arche* of the subject, but in their *telos*. While Freudism goes deep into the subject looking for the sources of meaning within the unconsciousness, in Hegel's phenomenology we deal with a forward movement towards an assumptive goal that is a full spirit. According to Ricoeur, we do not just have to do just with a reverse but with a shift: Freudian metapsychology, like Hegel's *Phenomenology* but in an opposite direction, carries out a shift of the centers of meanings, the birth of sense. However, despite this shift, Freudian psychoanalysis can be read in Hegelian perspective and thus it can reveal mutual points of Freud and Hegel's thought. The first of these points is the *arche* of the subject, which contrary to appearances, is perceived similarly by both thinkers.

Both Freud and Hegel are convinced that conscious subjectivity (together with culture) is born in the desire movement, which on the one hand can be noticed in a Freudian drama of the subject struggling with their own desire, and on the other can be seen in a Hegelian struggle for recognition. In Ricoeur's opinion, this struggle, which is about forcing another to confess that "I am an autonomous consciousness" means that the desire is transferred to a sphere of spirit: concepts of domination and surrender, which belong to a Hegelian language, denote the fate of drives in a Freudian language [32]. This similarity is sharpened by Kojève's direct interpretation of Hegel's struggle for recognition as a struggle between two opposing desires. Secondly, despite the fact that Freud's psychoanalysis emphasizes a role of archeology of the subject, there is also *telos* present within it. The concept remains very close to Hegel's assumptions – an ultimate goal of psychoanalysis is to broaden the consciousness of the subject. An ideal end of psychoanalysis is equal to the creation of the subject that is not alienated, and within which, there is a stable self-conscious self at the center.

Achieving this point, as Ricoeur points out, is reminiscent of the Hegelian phenomenology, with its phase of the struggle for recognition when we look at a process of raising the subject's consciousness in analytical relation. The process of self-awareness goes through the consciousness of another, an analyst's consciousness, what contains an element similar to a Hegelian struggle of two self-knowledges. Nevertheless, one cannot forget about differences: in Hegel's thought, we do not just deal with the phenomenology of consciousness, but with the phenomenology of spirit. As Ricoeur writes, the genesis of meaning does not originate in consciousness, but the consciousness itself is rather inhabited by a certain sense. The Spirit is a dialectic of definite figures (a master and a slave, stoic withdrawal, indifference of skepticism, unhappy consciousness, menial consciousness, etc.). The consciousness is an internalization of movement, which is also explored in structures of objective institutions, monuments, works of art, culture. Each figure gets its meaning from the one that follows it. Self-overcoming of the spirit creates the truth of the earlier figures that do not know

about themselves without it. Hence the terminology – phenomenology of the spirit and not of the consciousness. Freud, in contrast to Hegel, is characterized by extreme skepticism about mechanisms of human consciousness. It is why Ricoeur calls him “a master of suspicion,” the term which he shares with Marx and Nietzsche.

The focus on an analysis of the unconscious *arche* of the subject and the *telos* which is about withdrawing alienation and disenchanting of false consciousness, is a characteristic feature of a whole alienation trend. Feuerbach, Marx, Nietzsche, Freud and their contemporary heirs, joining in a specific way Hegelian tradition, at the same time oppose to the well-grounded understanding of a human subject in philosophy as a self-evident, stable, self-conscious *cogito*. A conscious subject ceases to be trustworthy for them, hence the inclination to enter deeply into *arche* of a subject that characterizes each of them. The source of subjectivity is not located in the open (conscious), but in the hidden. Freud talks about unconscious complexes and drives, Nietzsche mentions will to power and resentment that is to wear various masks. Marx notices the importance of power relations and economic forces in the genesis of consciousness constructs, such as religion, art or philosophy. In all these cases, what is hidden becomes building material of that what is open, at the same time escaping from the domination of the subject. Resentment produces values, economic power creates ideology, anxiety and wishful thinking originate religion. Drives are expressed in sublimation, erroneous activities, neurosis and dreams. At the same time, alienated critics of religion clearly outline the *telos*, that is broadening of the field of consciousness and achieving the widest possible range of self-control. The attitude to religion presented by the discussed thinkers does involve such understanding of the subject.

### 3. Religion as a Dream of a Human Spirit

Understanding of religion in alienation criticism is already reflected in the language of its representatives. Let's have a look at some definitions of religion that can be found in their works. According to Feuerbach, religion is “a dream of a human spirit”, according to Marx it is “opium of the people”, “imaginary happiness.” Nietzsche speaks of “omma”, Freud of “obsessive compulsive disorder” and “a sleeping-pill.” Finally, for Žižek, religion is involved in “phantasms” underlying social reality. Comparisons to sleep, mental illness or narcotic illusion prove that religious beliefs cannot be reduced to a simple cognitive error and their genesis is more complicated. The concept of religion as a dream seems to be the most paradigmatic of all these definitions. It seems to achieve its full meaning when one reflects at it from a Freudian perspective where religion has a clearly similar structure to dreams. Although in Freudian psychoanalysis a definition of religion as “collective neurosis” has become much more popular, here we do emphasize an analogy of religion and a dream as it has a much broader scope and is used by almost all representatives of alienation criticism of religion.

Already at the beginning of *Interpretation of Dreams* Freud categorizes a dream within the same sequence of “abnormal mental formations” that includes diseases such as a hysterical phobia or insane and obsessive dreaming [4, p. XXIII]. A diagnosis of religion as an obsessive compulsive disorder by its structure is in line with the same group of the formation. The analogy between dreams and religious beliefs is later confirmed by Freud in *Totem and Taboo* [7]. *Interpretation of Dreams*, Freud's first larger work from 1900, makes a supported by numerous examples study into explaining etiology of dreams based on a complex number of causes, displaced complexes and camouflaged wishes. It suits perfectly the outlined concept of the subject who is not the master themselves.

Distinctive features of a dream and a neurosis (including a religion, that is a collective neurosis) are as follows:

1. They hide their true motivation. They can cover it with another, quite consistent system. We always have to do with a kind of facade that is accessible to our consciousness, the one which at the

same time hides its true determinants. In a case of a dream this façade is created by the content of a dream story, whereas the hidden part consists of dream thoughts (a coherent complex of suppressed wishes that are expressed within the dream content but in a strongly camouflaged way). Work of a dream transforming hidden dream thoughts into a content of a dream that becomes accessible to consciousness is very complicated and uses a variety of mechanisms, such as condensation or dislocation. Similarly to a case of neurosis – its symptoms show coherence and consistency, but this consistency is apparent, because in reality the symptoms take their vitality out of the determinants that are hidden. This procedure also applies to a Freudian collective neurosis, that is religion.

2. They are not irrational – according to Freud, every dream, in reality, is a sensible psychic creature that should be placed in a locally identifiable mental state of vigilance. In other words, there is a close connection between a dream (neurosis/religion) and the reality that is known to a subject. Symptoms result from real events that take place in the subject's mental life and they express them. As long as we agree that religion in alienation criticism of religion has a similar structure as a dream in Freud's classic psychoanalysis, we must admit that it does not have much in common with a simple mistake of irrational thinking. Its ground is as real as possible. Ludwig Feuerbach does recognize this aspect of religion while writing: 'Religion is a dream of a human spirit. But also in a dream, we are not in nothingness or in heaven, but on earth, in a kingdom of reality, with a one and only difference that we see real objects not in the light of reality and necessity, but in a fantasy world of delusion and randomness' [2]. Certainly such a depiction of religion will stay in line with Karl Marx's view who looks for its etiology in the most tangible of factors – the "opium" does not come from nowhere, it is 'an expression of real poverty/misery and at the same time a protest against real poverty/misery' [19]. In spite of an illusory form that religion adopts its genesis is rationally explained by economic conditions. However, these determinants are not recognized by the subject.

3. The explanation of dreams and hence the criticism of religion which remains characteristic for alienation trend, in its basic design is a decryption that is intended to lead to broadening of consciousness and the subject's recovery of themselves.

4. A dream and all forms of mental life whose structure is similar to it contain both the truth and the untruth, or rather the truth in a form of illusion. Despite a strong camouflage using apparent irrationality, a dream contains a hidden truth about human nature in a pictorial form: 'Thus dreams would reveal the true nature of man, though not his whole nature, and they would constitute one means of rendering the hidden interior of the mind accessible to our knowledge' [3]. The truth about this interior very often does not have building character. Let's recall another passage from a Freud's work: 'In this, psycho-analysis is no more than confirming the habitual pronouncement of the pious: we are all miserable sinners' [5, p. 48].

Remembering the concept of religion as a dream, let's illustrate it with an example of Nietzsche's thought. In *The Daybreak* one can find two very important quotations. The first one is as follows: '*Awaking from a dream*. Noble and wise men once upon a time believed in the music of the spheres; there are still noble and wise men who believe in "the moral significance of existence," but there will come a day when this music of the spheres also will no longer be audible to them. They will awake and perceive that their ears have been dreaming' [20].

"For Nietzsche" the moral sense of existence" and religion are bilaterally related. On the one hand, the predominant moral system in a particular society is designated by religious outlook. On the other hand, the religion is founded on morality imposed on the dignified by the common majority. Both of them, as Nietzsche says, make the world "become a dream." In *Antichrist*, he diagnoses religion as follows: 'neither morality nor religion has any point of contact with actuality. It offers purely imaginary causes ("God", "soul", "ego", "spirit", "free will" - or even "unfree"), and purely imaginary effects ("sin", "salvation", "grace", "punishment", "forgiveness of sins"). Intercourse between imaginary beings ("God", "spirits", "souls"); an imaginary natural history (...); an imaginary psychology' [23].



Bearing in mind Freudian metaphor of a dream, there is no doubt that these imaginary elements (after all forming a very coherent system) are parallel to the façade that was a characteristic feature of a dream. Let's recall the second quotation from *The Daybreak: 'Dreaming and responsibility*. You would wish to be responsible for everything except your dreams! What miserable weakness, what lack of logical courage! Nothing contains more of your own work than your dreams. Nothing belongs to you so much! Substance, form, duration, actor, a spectator in these comedies you act as your complete selves' [20, p. 78].

This passage is really significant – even in dreams the subject is themselves, although it does not necessarily have to be conscious of a system of beliefs and morality that they share to and as well as all the frustrations that result from it. Let's refer to Freud again. The author of *Interpretation of Dreams* explains it in a very simply way – a dream is a fulfillment of a hidden wish. A dream is a place where we have to do with the real nature of a man, although because of a caesura a dream can take such a twisted form that it becomes brutal and traumatic. As everybody knows, according to Freud both dreams and all forms of similar structure, such as a phobia or a neurosis, including a collective one, that is religion, result from a complex mixture of a drive, a ban and a complex. Despite all that, the content of a dream does express some truth about ourselves.

Nietzsche has no doubt that any consciousness originates in the sphere of the subject that is guided by instinct and drives. According to, among others, Michel Onfray the whole Freudian concept of unconsciousness and impulsivity, including displacement and sublimation, originates in Schopenhauer's notion of "a will to live" and in Nietzsche's philosophy, especially in his theory of "will to power." The categories of will to power, as well as known from the Freudian psychoanalysis notions of displacement and sublimation which also find their origin in the philosophy of the author of *Thus Spoke Zarathustra* are clearly visible in Nietzsche's reflection on Christianity: 'The "Christian" – he who for two thousand years has passed as a Christian – is simply a psychological self-delusion. Closely examined, it appears that, despite all his "faith," he has been ruled only by his instincts – and what instincts! – In all ages – for example, in the case of Luther – "faith" has been no more than a cloak, a pretense, a curtain behind which the instincts have played their game – a shrewd blindness to the domination of certain of the instincts...' [23, p. 39]. Passing over a problem of engineering of religion that was discussed by Nietzsche among other things in his *Antichrist* and *On the Genealogy of Morality*, the fact that its genesis and its motive is a desire to reevaluate values in such a way that they would subordinate dignified people to the weak ones – let's pay attention to what religion is for the believer subject themselves. In *On the Genealogy of Morality* Nietzsche shows very interestingly what Christianity does to the real determinants underlying it. The basic determinants are situated in a *Trieben* sphere, that is the sphere of drives. As these determinants are to remain hidden, there are three possible types of psychological shifts: *Triebversicht* (displacement of a drive), *Triebaufschub* (adjournment of a drive) and *Triebverschiebung* (redirection of a drive). The drives certainly do not disappear, they do not become eradicated but they do change the way they are fulfilled. The drives in question which Nietzsche writes about are involved in the will to power that takes different forms, and which is fulfilled in a resentment and indirect way, at the same time remaining a form of aggression and cruelty towards of the strong by the weaker. Here the dream facade is a whole complex of religious beliefs and moral convictions that are related to them – the faith in God, the consecration of Christ as well as eternal life, humility, love of neighbor, rituals, dogma. Their rational, hidden and redirected determinants, however, are just resentment, lust for domination, hatred and revenge.

Even a cursory glance at a depiction of religion in alienation criticism draws attention to a difference between this trend and the one that seems predominant nowadays among the other atheist ones, that is New Atheism which is adopting the post-Darwinian perspective. For Dawkins, Dennett, or Harris, religion is a simple cognitive error that is relatively easy to uncover by indicating its irrational and inconsistent elements as well as natural genesis. A "sharp voice of reason" is completely enough to do it. The alienation trend in turn points out to a much more



complicated and to some extent “layered” structure of religion – behind a coherent façade (which does not mean the rational one) there are some motives, concealed secretly. They do possess their *ratio* but it is camouflaged well, disassociated by psychological censorship that is covering up that determinants of religion (alike those of other formations similar in structure to a dream) are neither to the credit of a believer subject nor to engineers of religion (see, for example Marx’s concept). However, it does not influence the fact that the façade is a symptom and its interpretation makes it possible to reach the determinants. Thus alienation criticism of religion tries to investigate why the Feuerbachian subject creates a notion of God that humiliates themselves, why the Nietzschean higher man is to be “seduced” by Christianity at the expense of resigning from life here and now, why the Sartrean subject denies themselves freedom in the name of the projection of the oppressive look that is chasing them, why the man according to Žižek resigns from being active and falls back on being alienated by the Big Other. The dream which is dreamt by an alienated subject has an enormous seductive power. That is precisely why in order to solve these puzzles, the alienation criticism of religion must reach into archeology of human consciousness, but assuming an overriding purpose, very close to Hegel’s goal – to make the subject cease to be a mystery to themselves.

#### **4. Unmasking of False Consciousness**

In accordance with the thesis I am proposing in this article, we can speak of a certain soteriological ideal within the alienation criticism of religion, the one which is realized by this trend by the unmasking of false consciousness. Next step involves the creation of a project of a new subject: full, real and capable of transgression. This project is not fully developed and its realization can be implemented in a couple of directions: both towards moral heroism and towards focusing the subject on themselves their own. Alienation criticism of religion does not say exactly what the “content” of a new subject is like (clear values, ways of acting, etc.). It rather focuses on a “formal” procedure of its construction. The context of which soteriological threads of this trend constitute is an affirmation of the truth proclaimed by it.

The conviction of a salvific role of the truth is one of the aspects that almost all atheistic philosophers have in common. One should bear in mind that it is not happiness, prosperity or peace that do save the subject. First and foremost, it is the truth that does it, even if that difficult to accept. Even if religion brings benefits to the subject – it acts as an escapist function, organizes social life, etc. – the subject pays far too much for that. A characteristic feature of all philosophical trends focusing on criticism of religion is Prometheanism. The trends aim to liberate a man from religious superstition and to direct him towards the truth. At the same time alienation trend emphasizes the existential truth, the truth of the subject as religious alienation is a fundamental self-imposture which should be disenchanting. Only the subject that exists authentically and is not contradictory in themselves, self-controllable and aware of one’s own capabilities, can be saved. Obviously, in a secular sense. An ideal of existential truth as the main postulate of alienation criticism of religion is already present in Feuerbach’s works, but it plays a vital role in Nietzsche’s thought to which subsequently Freud refers. The ideal is later developed by the successive representatives of the trend. In the Marxist tradition, in turn, the existential truth is associated with the proclamation of the unity of thought and action which in the alienation criticism of religion enable a real transformation of the one and only reality available – the reality here and now. At the same time it is worth noticing that according to the thinkers in question, the secular salvation is required not only by an individual subject but also all of their contemporary culture. A problem of the truth is also debatable here – the truth that culture is devoid of. Surprisingly, it is Feuerbach who already draws attention to the same thing that nowadays is being diagnosed by Lasch, Bell, Bloom, and Baudrillard: culture is being trivialized, subjects that constitute the culture are endangered by narcissism, routine, cynicism, automatism. There is probably one and only remedy for this situation: the transformation

of existence, recovery of autonomy and in the first place, a sense of efficiency. These threads are found in works by all representatives of the trend. They all associate trivialization and cynicism that are so characteristic for Western culture with supporting of a routine and an inauthentic form of human functioning. The first step which should be taken in order to change this situation is the unmasking of false consciousness. The second one would be a project of a new subject.

The unmasking of false consciousness consists primarily in unveiling a mechanism of alienation. Basically, all the representatives of the trend in question refer to the Feuerbachian scheme: we deal with self-imposture, projection, and lack of recognition of our own essential aspects of being as ours. The differences between them appear among other things once a question about a cause of alienation is asked. In *The Essence of Christianity* by Ludwig Feuerbach where one comes across a complete schema of religious alienation applied in a context of atheistic philosophy for the first time, they do not find a satisfactory answer to this question. In the later *Lectures on the Essence of Religion*, the author suggests a sense of dependence on nature as one of the reasons of alienation. However, it is not in line with the thesis presented in his previous (and undoubtedly more influential) work and causes a considerable dissonance between *The Essence of Christianity* and *Lectures*. Marx is the one who polemizes with this resolution. Having adopted the Feuerbachian schema of alienation from *The Essence of Christianity* he claims at the same time that it is not possible to understand the genesis of religious alienation without a reference to a political context. Religious alienation is one of manifestations of more broadly understood social alienation which in turn is caused by a political-social system that has been wrongly structuralized. In this case the origin of religion is reduced to two sources: on the one hand, in an imaginary and distorted way it reflects the prevailing and unjust social system; on the other hand, it supports it deliberately. It becomes “the opium,” which promises to replace the passive suffering and injustice of the subject with a reward after their death. The cause of religious alienation is therefore social injustice, manipulation, and suffering of the individual.

Nietzsche and Freud notice the origin of religious alienation in repression of drives. The theories of both thinkers were often compared with each other even during Freud’s life, although he strongly denied not only being inspired by Nietzsche’s thought, but also the knowledge of Nietzsche’s works. However, commentators draw attention to Freud’s strong dependence on Nietzsche’s thought which can be seen in his central depiction of drives as well as in the construction of a concept of unconsciousness, sublimation or repression of the drives. Even though, while comparing these thinkers it is impossible to forget about such differing components of shaping religious attitudes as Oedipus complex in Freud’s thought or rebellion of slaves in Nietzsche’s works, the psychology of the subject completing self-alienation in religion is outlined by both of them in a similar way. Nietzsche and Freud consider it always in a context of three factors: a relation of a man to themselves, the activity of their drives and a manner in which these drives are converted. For Nietzsche, the key word here is “resentment,” for Freud it is “a complex.” In both cases, an important role in the construction of religious beliefs is played by a motive of repressed and camouflaged revenge.

The categories that were later adopted by Freudian psychoanalysis began to appear in Nietzsche’s thought in the late 1870s in a work *Human All too Human* (1878), in which he mentions sublimation in a psychological sense for the first time. In *Beyond Good and Evil* (1886) Nietzsche emphasizes a sexual drive that is displaced or transformed in religion and like Freud he compares religion itself to neurosis. However, the most interesting works are those from the late 80’s relating precisely to religion: *On the Genealogy of Morality* and *Antichrist*, within which a notion of “resentment” becomes the key word. Nietzsche does not provide a precise definition of resentment, pausing at such depictions as “spiritual self-intoxication”, “telling oneself lies,” “targeting at the outside,” but following Max Scheler we can accept understanding of resentment as a constant mental attitude that arises when emotional reflexes and passions are systematically repressed [34]. These reflexes and passions involve: a reflex of revenge, hatred, malice, jealousy,

etc. The repression is closely connected with the disintegration of the subject, their self-alienation and appearance of pathological symptoms. An intersubjective effect of this phenomena, however, is the emergence of a new, re-evaluated system of values and religion. A similar mechanism is described by Freud in *Totem and Taboo*. Skipping the aspect of totemic feast and focusing on the psychology of an individual subject – the source of religion is the transformation of a drive and covering the tracks of a primal motivation. For both Nietzsche and Freud, this genesis is connected with such psychological mechanisms as displacement and compensation in the world of illusion. Nevertheless, Freud elaborates on this theory associating it with a father's complex whereas Nietzsche presents religion as a tool of revenge that is performed by the miserable on the dignified.

In Sartre's opinion, religious alienation is a form of escape from responsibility. French philosopher equips the subject with freedom of conferring a sense which turns out to be a burden. Trying to reduce unavoidably related fear, the subject is likely to escape into routine actions and colloquial or daily morality. All these ways involve negating one's own transcendence, perceiving oneself as a thing of established nature and sense. Religion is just one form of such an escape. The aspect of absolute freedom that the subject is equipped with is shown in the project outside and attributed to an illusory being. This very being is at the same time a response to the longing of the subject for a certain ontological whole: the synthesis of being in itself and being for itself.

Apart from differences, a general diagnosis – alienation of a subject resulting in an inauthentic existence – is the same for all the representatives of the trend. They also ask a question whether unveiling the subject involved in religion in front of the consciousness will result in liberation. The answer is rather negative. Feuerbach represented a very optimistic attitude and in his opinion "disenchanted" of the false consciousness would be an effect of appropriate education. At the beginning of his work Marx used to agree with such reasoning, however, later he proposed a more sensible thesis: a change of the most important aspects of consciousness is impossible without a change of social conditions. Marx's gloomy continuator (and a mocker at the same time) is Žižek who claims that a dominant attitude in our society is a cynical one. Even if the subject does not believe "spiritually," the subject participates in a religious paradigm, supports it with their own practice, and still remains alienated due to a fundamental dissonance between a thought and an action. Despite these sad conclusions, alienation critics of religion do not resign from their faith in a possibility of secular salvation of the contemporary subject. They believe that authenticity, working activity, true and active secular faith instead of cynical inertia is possible and should be pursued. A trial to construct a new subject will be shown at the example of a thought by Friedrich Nietzsche, one of the most influential representatives of alienation criticism of religion.

## 5. Salutary Meaning of Sublimation

The term that often runs through this article to designate a man facing alienation is an epithet "authentic." Although it is quite a problematic word because of its historical background, in fact, there is not a better one. Without going into historical details of this notion, one should remark that it has been related for a long time with a discovery of an inner self which is much more intimate and real in relation to what is presented by the subject outside. One can find it used in such a sense in works by Hölderlin or by Rousseau who is said to have propagated the notion. "Discovering one's self" has become a slogan and is used unusually often in our culture in different contexts, both serious and pop cultural ones. However, significantly it has nothing in common with a concept of a new man who is proclaimed as a soteriological ideal of alienation criticism of religion. First of all, alienation trend suggests an anti-substantial understanding of the notion of authenticity - a new man is the one who dissociates from expressions that would captivate him, creates himself without relying on any "ready" and "existing" self. One of the sources of such a depiction of a new man is Nietzsche's anthropology.

Friedrich Nietzsche's thought is a real breakthrough in a tradition of understanding the term authenticity. His aphorisms are so ambiguous that he can be recognized as both: a critic and an eulogist of this idea. However, he sounds critical only when he negates the concept of inner self that can be discovered and known. "True, inner self" – he writes in *Untimely Meditation* – it is a fantasy or "mythology." Nevertheless, we can try to "look inside ourselves." What will we find there? Nothing except a bundle of antagonistic drives. Originally the subject is scattered and multiple. Yes, it is capable of achieving their inner coherence – as Nietzsche writes in *On the Genealogy of Morals* – in a resentment and twisted manner that additionally conceals assiduously their own motivation. At the same time, he suggests that there may be another, more recommendable way of dealing with the bundle of drives that constitute us. Nietzsche speaks of a true way of being. It is also based on a certain project but the project which is proper and desired. It is evident that a serious shift in understanding of a notion "authenticity" is a characteristic feature of Nietzsche's thought if it is analyzed in a perspective of the history of ideas. The term no longer designates fidelity to the uncovered inner core of personality but is realized in a project whose implementation is determined by a certain procedure. In *Joyful Wisdom*, Nietzsche proclaims that basically a man is assigned to enrich their style, add some autonomy and creativity to their character as these features are distinguishing for humanity. However, let's not forget that according to Nietzsche a man is a reservoir of contradictory drives. The statement makes him explicate a basic condition that constitutes a new man: the continuous overcoming of oneself.

Presented by Nietzsche existential ideal is to be complemented by a modern man. Similarly to the rest of the representatives of alienation criticism of religion, he also draws attention to fundamental complaints of his time, a fact that a man has fallen into conformism, automatism, routine, banality and negation of anything that possess value. The only alternative to this phenomena is the realization of an existential project based on a crucial role of transgression – the transgression that is necessary if we do not want to become slaves to our disruptive and contradictory drives. Nietzsche uses one of his key words to describe it: *Überwindung* – self-overcoming.

Nevertheless, the concept of *Überwindung* in Nietzsche's works is ambiguous and generates some problems with its understanding. Let's ask a basic question: what is actually being overcome and how? The notion of self-overcoming gains its meaning in the later Nietzsche's works, from the 80s. At that time, he does not write about numerous autonomous drives any longer and brings them all together as a will to power. The case is quite controversial, as evidenced by a discussion between Walter Kaufmann and Giles Deleuze. These commentators argue quite sharply about the relationship between the forces that are being overcome. According to Kaufmann Nietzsche in his early works inclines to a statement that contradictory principles are strength and reason. On the other hand, within a developed stage of his philosophy, which is the main point of reference for us, Nietzsche turns to an attitude classified by Kaufmann as monism. The factor that is to overcome and is being overcome at the same time is solely the will to power manifesting itself in numerous ways. At the same time Nietzsche rejects the sovereignty of reason that is also reduced to the will of power. Despite all that, a man is not its play. Although it constitutes the metaphysical principle of reality, it does not exclude human activity and autonomy. That is why Nietzsche, in spite of all, values reason and intellect more than, for example, the realization of a sexual drive that can cause captivating of the subject. Moreover, Kaufmann claims that only being reasonable can guarantee a man power of self – determination and self-control.

The preceding description comes down to the fact that Kaufmann ultimately classifies Nietzsche as a "dialectical monist." In his depiction self-overcoming of the will to power leads to a synthesis of its aspects. In conclusion, Kaufmann notices a clear parallel between the Nietzschean *Überwindung* and the Hegelian *Aufhebung*. In turn, Deleuze sees Nietzsche in a different way. He is convinced that Nietzsche's philosophy is clearly anti-dialectical in its character. It is not so much dialectical as pluralistic. Deleuze argues that in a situation when we have to do with human

transgression, the ratio of the overcoming power to the one that is being overcome is never understood as a negative element. In a relation between two forces, the one which imposes obedience negates neither the other one nor anything that constitutes itself. It just affirms the difference and enjoys it.

Once we try to apply psychological categories in place of the Hegelian terms, the difference between Kaufmann and Deleuze might appear less serious. It is clear that transgression which is crucial for a creative human being consists in overcoming some manifestations of the will to power by means of the other ones. However, there is a question whether the transgression is a continuous struggle that is just fated to failure? Does it come to an end? Does human creativity possess a better or worse form? When can we say a man “possess themselves” and when that they are “divested of themselves”?

A distinction between repression and sublimation that occurs in Nietzsche’s late philosophy presents this problem in a slightly different light. As it was mentioned before, religious behaviors are primarily related to a mechanism of repression – a man denies all that is fundamental for them (although displaced elements do come back), at the same time alienating themselves. Similarly to Freud, Nietzsche was never a supporter of an immediate, debauched realization of all the drives. Despite some liking for “blond beasts,” the ideal affirmed by Nietzsche is a self-controlled man that is capable of overcoming themselves. Nevertheless, it is not so easy while a man is defined as a reservoir of contradictory drives.

The problem of the distinction between repression and sublimation sets us on an appropriate trail of understanding a soteriological ideal of existence. Whereas repression generates “unreal” existence, sublimation is the one and only proper way of dealing with one’s own impulsive nature, which, as it was mentioned before, always expresses the will to power. As Kaufmann notes, the process of sublimation in Nietzsche’s philosophy has been either disregarded or omitted for a long time [17]. Even in such a “psychologizing” work as Jasper’s study of Nietzsche, sublimation is only mentioned as one of the numerous ways to “cope” with a drive and the author neither emphasizes it nor attributes any special role to it.

In fact, the term sublimation understood according to in Nietzsche’s late thought belongs to key words and is closely related to an ideal of self-overcoming as one of the soteriological threads. Not only does Nietzsche anticipate Freud in this case, but also is a figure who for the first time in the history of ideas formulates a psychological meaning of the word sublimation.

In *Human, All too Human* he defines sublimation in a way that is close to Freudian interpretation and the approach of later psychoanalysis. Sublimation, as he writes, consists of a process within which “what is rational” comes into being from “the irrational,” “logicalness” from “illogicalness” and “impartial” from “greed.” However, the problem of distinguishing between repression which results in an inauthentic existence and sublimation that plays a key role in forming a creative and unalienated existence is quite difficult. As an example, one can have a closer look at Christian love of neighbor. As Nietzsche points out, it is an effect of resentment and in fact, it is nothing else but a veiled vengeance. Briefly put, it results from repression, not sublimation. Nevertheless, it does not change the fact that the “social effect” itself is positive. Why cannot we just say that simply a sublimation of negative drive has taken place? After all, “impartiality” has emerged from “greed.” Certainly, one can remark immediately that in a broader perspective of the Nietzschean theory of religion and the revaluation of values in order to subordinate “the strong” to “the weak” it is actually apparent. Still, the question of Christian love of a neighbor shows how subtle and at the same time significant distinction we are to deal with when it comes to repression and sublimation. This problem becomes even more evident when Nietzsche’s reflections on Freud’s theory are confronted. Ken Gemes in his article *Freud und Nietzsche on Sublimation* shows relations and differences between them and proves that contrary to all appearances, the opposition of repression and sublimation is clearer and much more consistent in Nietzsche’s thought that preceded Freudian psychoanalysis [12].

To begin with, let's recall Freud's distinction between repression and sublimation, bearing in mind that like Nietzsche he took the view that sublimation of a drive determines a healthy subject. Suppression of drive results in "restitution of the displaced" in the form of pathological symptoms which are some sort of "substitute" of a direct target on which the displaced drive is directed. Sublimation prevents from uncontrolled "restitution of the displaced" and at the same time its effects are also "substitutes." However, the thing is whether there is any reliable criterion that clearly distinguishes these substitutes from the symptoms connected with the restitution of the displaced. According to Gemes, Freud's solution is not entirely satisfactory because in order to distinguish between these two types of psychic formations he had to refer to arguments beyond a discourse of psychoanalysis. It is evident within a comparison of two cases that he analyzed: Leonardo da Vinci's and Daniel Paul Schreber's.

*Leonardo da Vinci, a Memory of his Childhood* (1910) is one of the most important Freudian studies devoted to a problem of sublimation. Just at the beginning, Freud distinguishes three ways of dealing with a sexual drive in childhood that ultimately determine what type of person the subject will become in their adulthood. The first manner encompasses displacement of a drive into unconsciousness along with the reduction of a desire to achieve knowledge (lack of interest in sex, also within a sphere of *ratio*, that is questions addressed to parents, etc.). The second one also includes displacement of a drive into unconsciousness but this time the drive returns within a sphere of knowledge. Finally, the third way is connected with sublimation and characterizes people who transform a drive into a craving for knowledge which strengthens a strong research drive. For Freud an example of this third type of man is Leonardo da Vinci, a man who sublimated his own sexual drives into artistic and research activities at the same time – he presumes – remaining an ascetic in his sexual life. Interestingly, Da Vinci himself is either portrayed by Freud once as a genius of sublimation or as an "obsessive neurotic" fixed at an infantile phase of his own development. Freud devotes a lot of attention to Da Vinci's relationship to his mother and father that he finds and presents as extremely problematic. Nevertheless, Leonardo's first depiction prevails: he avoids repression and the effects are well-known in a form of fantastic works of art and intriguing research sketches.

An analysis of a case of repression, presented in a certain sense as an antithesis of Leonard's attitude, can be found within reading another of Freudian sketches: *Psycho-Analytic Notes upon an Autobiographical Account of a Case of Paranoia (Dementia Paranoides)* (1911). Like in a study of Leonardo its author analyzes a case of a person he did not examine personally: this time it is the former president of the Saxon Senate, Daniel Paul Schreber [35]. As analysis material, Freud used Schreber's autobiography. Its author survived a three-times episode of mental illness, two of which were described in details in his book (the third one is known only from hospital documentation). In a case of the second episode that was treated in a closed psychiatric hospital, among other things we have an access to a brief description of an illness prepared by the court to whom Schreber applied for a release from the institution. In this document, his character is recognized to be aspiring to save the whole world and to restore a state of lost bliss. The mission is likely to succeed only if the character in question transforms himself into a woman. The condition was finally qualified as a manic formation.

Freud analyzes this case in much the same way as Leonard's case but classifying its cause as a classic form of repression of a drive. Devoting a lot of attention to an analysis of Schreber's relationship with his father and his physician Dr. Flechsig who used to treat the first of the sickness episodes Freud concludes that the cause of Schreber's problems was – like in Leonardo's case – unaccepted homosexual impulses. As far as Leonardo is concerned, Freud comes to this conclusion mainly by analyzing his works in which he sees idealized figures of men or women of an androgynous appearance. In Schreber's case among other things, he focuses on dreams within which the man fantasizes about becoming a woman fulfilling her matrimonial duties as well as on the content of manic fantasy in which similar images appear but this time gaining a religious

context. In both cases, we are to deal with a homosexual impulse. The difference lies in a way it is transformed: Leonardo sublimates his drive in a creative activity, Schreber falls into a religious mania.

A key question that should be asked at this point is about an exact difference between these two cases? As Freud proves we are never able to avoid any form of displacement of a drive but we can transform it in a different way in the long run. Actually, why are we to assume that Schreber with his religious obsession is less healthy than Leonardo obsessed with making sketches of flying machines and other never-completed projects? Is the difference between sublimation and repression really the one of an antagonistic character?

Ken Gemes claims that a fundamental problem in this distinction is that following Freud, we would ultimately have to end up explaining the results of sublimation emphasizing their social value. Affirming Leonardo's character, Freud speaks of "high cultural accomplishments," "redirecting of a sexual drive towards higher goals" such as science or art. This means introducing non-psychoanalytic and valuing terms into the definition of sublimation. In other words, the same behavior can be once classified as pathological and another time as sublimated one [9]. That is also the reason why the characters of Leonardo and Schreber were evaluated by Freud in a completely different way. According to Freud obsessive behavior and thinking focused on God do not have any social value whereas the works created by Leonardo do have.

At this point, it is worth going back to the very sources – that is to Friedrich Nietzsche who is a bit clearer about the issue. In his view, the distinction between repression and sublimation is not a measure of "health" or "sickness" but a difference between a weak and resentful subject and a strong one that is capable of auto-creation. Both of them face an internal conflict between drives that cause them to experience internal "uproar" and "struggle." The measure of the power of the subject will be an ability to overcome this sensation. A sublimating subject will primarily be the one that transgresses themselves (overcomes themselves) and the one that "exists in a real way" in a sense of their capacity of full self-control.

According to Gemes a notion of "splitting off" which is used in the first place in reference to repression is at the same time the one that helps to understand the Nietzschean distinction between repression and sublimation. On the one hand, the term signifies a separation of a right impulse from consciousness. On the other hand, it means a total separation of an impulse and its proper realization. The "splitting off" can be also treated as a measure of a subject alienation – the subject separates themselves from their own motives which, as psychoanalysis points out, do not disappear but just slip out from their domination. Nietzsche – like Freud in his study of Leonardo – is convinced that the problem of "splitting off" affects both the weak and the strong. However, while the weak one surrenders to resentment, the strong one is capable of struggling with themselves and is able to redirect their drives towards a higher goal. It is also significant that in an act of sublimation redirected drives are put to appropriate use entirely, which means that traces of their possible displacement will not return in an uncontrolled form but totally surrender to a domination of the subject. While in a case of repression a man becomes a "slave" of their own drives, in sublimation from the beginning to the end they do serve them. It is also important that – in Nietzsche's view – in sublimation process, there is no falsification of a primal drive or its complete transformation into the opposite one. The drive is subordinated to the subject and transformed. Although it cannot be realized in a direct and uncontrolled form, it still retains its authentic character. It seems obvious once we refer to an opposite case of Christian love of the neighbor. According to Nietzsche it "grows out of a trunk of hatred" and what is also worth mentioning it never recognizes its true motives. A Freudian opposition between the cases of Leonardo and Schreber could be recalled again and one should make an attempt to interpret the contrast in question in a Nietzschean manner, certainly bearing in mind the hypothesis put forward by Freud. In Leonardo's case the homosexual drive "has got refined" as a result of subordination to scientific research and artistic activity but at the same time, it has not lost its chance of adequate expression.

Hence his concentration on a study of a human body and imitating male beauty in pictures of men and androgynous women. Da Vinci is entirely in control of himself subordinating all his impulses to what he perceives as superior: a drive to a scientific and artistic creativity at the same time incorporating his own sexual drive in the activity. Therefore, if we analyze Leonardo's case from a Nietzsche's point of view a factor that evaluates his attitude is not social value of his works as it is suggested by Freud but an ability to control himself while simultaneously integrating the drives and subjugating them. On the other hand, as far as Schreber is concerned, we have to do with psychological disintegration and the opposite of a full, unified "I" that could be noticed in Leonardo's case. Displaced homosexual impulses explode here at different stages of religious mania which is entirely beyond his control and which turns out to be a completely inadequate form in relation to its primal motives. Besides, it is worth noticing that apart from the episodes of mania, Schreber presents himself as a rational, skeptical of religion and responsible politician as well as a staid spouse. In the Nietzschean psychology of drives, as Gemes points out, sublimation means that stronger drives integrate the weaker ones which enables these weaker to express. On the other hand, repression noticed in Schreber's case consists in a fact that stronger drives try to suppress the weaker to such an extent that their expression is sometimes restrained completely. Still, it more often reveals in a strongly distorted form, usually aimed at the opposite of the original. This is exactly what in Nietzsche's opinion happens in case of Christianity: hate, jealousy or a desire for domination present themselves in a form of fraternal love.

Although Nietzsche's theory is not as thoroughly prepared as Freud's concept it does hold an advantage over it. As opposed to Freud, distinguishing between repression and sublimation it remains strictly within the internal life of the subject without resorting to any external criteria. The social value of deeds does not decide whether the subject is "healthy" or "sick." It is determined by their own relation to themselves and a degree self-control. The difference between a subject who victimize their drives and the one that sublimates them is ultimately reduced to a degree of internal, mental integration and self-control. A figure of self-control has a great significance for Nietzsche and is crucial in his creation of existential ideal. Its actual meaning is revealed only when we understand thoroughly the difference between repression and sublimation. This difference reveals an outstanding feature in the psychology of alienated man as well as the psychology of man from alienation. At the same time it refers to a notion of transgression whose comprehension is also determined by a proper understanding of sublimation. An analysis of a concept of sublimation in Freud and Nietzsche's works reveals two distinct components of a soteriological existential ideal: on the one hand there is "I" that has been liberated from alienation, more unified, restored to themselves and complete; on the other hand there is a self-controlled and able to transgress one. The notion of transgression itself is not characterized by an uncompromising struggle or pluralistic coexistence of Deleuze's pluralism. If we have a closer look at self-transgression taking place within the subject from a perspective of the Nietzschean concept of sublimation, we realize the necessity to cope with integration that is simultaneously related to precise self-control.

Finally, let's refer to a concept of an excess of life and passion that is related to it. A proper form of existence suggested by Nietzsche associates with vitality, self-control, and enigmatic figure of "passion" understood not just as a blind drive but as a figure organizing the rest of drives. According to Kaufmann, this figure is equal to the firmly affirmed by Nietzsche pure reason. In my opinion, it is not so obvious. Gemes writes about "a dominant drive." What is this "dominant drive"? Nietzsche does not answer this question explicitly, restricting himself to pointing out the only measure that defines it – absolute commitment to the impulse chosen as prevailing – the mentioned before – "passion."



## 6. Conclusion

The model of an authentic subject presented by Nietzsche presents explicitly the main elements of a soteriological existential ideal created by alienation criticism of religion. Its starting point is an alienated subject that is internally “divided,” unaware of the truth about themselves and consequently unable to control themselves. An answer to this problem, shared totally or partly by the other representatives of the trend, seems to be the subject found within Nietzsche’s concept. It is autonomous, autocratic and creating oneself in the truth. On the other hand, an analysis of Nietzsche’s theory points out some inevitable problems associated with understanding the subject in such a way. We have to do with the one that has liberated themselves from an alienating projection of God. However, who is the subject now? The same one as “before” alienation or maybe someone completely different? Alienation critics of religion know that the return to a hypothetical pre-alienation state is impossible. Instead, they emphasize an element of creation. After getting rid of dependence on an idea of God, a man is capable of being exactly the one that they have created themselves. Nietzsche’s philosophy suggests that defining a path to existential salvation of the human being understood in such a way comes down to a certain procedure: you can become whatever you want to be like. However, the manner of self-construction is not arbitrary. In this case, Nietzsche speaks of consistency, active organization of our contradictions and creative applying a style to our character in harmony with principles that we find the most important. Nevertheless, the problem of choosing the principles according to which we are to accomplish our auto-creation does reveal basic antinomy of the soteriological ideal proposed by alienation criticism of religion: the problem is to what extent the choice in question is identical with our will or our desire? Where is a boundary between existential freedom and total arbitrariness? These are the questions that the alienation criticism of religion is still trying to find answers for, opting for a practical aspect. Which of our choices can be described as authentic is ultimately determined by a degree of our involvement in it and by our awareness of responsibility. The representatives of the alienation trend more and more often write about a necessity of real, secular faith whose verification would be an explicit proclamation and in the first place acting in accordance with it. Conclusively, our authenticity can only be proved by our practice.

## References

1. Assoun, P. *Freud and Nietzsche*, London: Continuum, 2000.
2. Feuerbach, L. *Das Wesen des Christentums*, In L. Feuerbach, *Gesammelte Werke*, W. Schuffenhauer (ed.), Berlin: Akademie Verlag, 2006.
4. Freud, S. *The Interpretation of Dream*, Transl. by J. Strachey, New York: Basic Books, 2010.
6. Freud, S. *Totem and Taboo. Some Points of Agreement between the Mental Lives of Savages and Neurotics*, Transl. by J. Strachey, London: Routledge, 2001.
8. Freud, S. *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. 24 vols, Transl. and ed. By J. Strachey, London: Hogarth Press, 1961.
9. Gardner, S. *Irrationality and the Philosophy of the Psychoanalysis*, Cambridge: Harvard University Press, 1971.
10. Gay, V. *Freud on Sublimation – Reconsiderations*, Albany: State University of New York Press, 1992.
11. Gemes, K. Freud und Nietzsche on Sublimation, *The Journal of Nietzsche Studies*, 38, 2009, pp. 38-59.
13. Gemes, K. Nietzsche on the Sovereign Individual, Free Will, and Autonomy, *Aristotelian Society* (Suppl.), 80 (2006), pp. 321–39.
14. Golomb, J. *Nietzsche’s Enticing Psychology of Power*, Jerusalem: Magnes Press, Hebrew University, 1989.

15. Groddeck, G. *The Book of the It*, New York: Vintage Press, 2006.
16. Hacker, F. Sublimation Revisited, *International Journal of Psycho-analysis* 53 (1972), pp. 219–23
17. Hegel, G.W.F. *Phänomenologie des Geistes*, Berlin: Duncker und Humblot, 1841.
18. Heidegger, M. *Nietzsche*, vol. 1: *The Will to Power as Art*, Transl. by D. F. Krell, New York: Harper and Row, 1979.
19. Jaspers, K. *Nietzsche, An Introduction to the Understanding of His Philosophical Activity*, Cambridge: Johns Hopkins University Press, 1997.
20. Kaufmann, W. *Nietzsche: Philosopher, Psychologist, Antichrist*, Princeton and Oxford: Princeton University Press, 2013.
21. Lehrer, R. *Nietzsche's Presence in Freud's Life and Thought*, Albany: State University of New York Press, 1995.
22. Marx K. *Critique of Hegel's 'Philosophy of Right,'* Transl. by A. Jolin & J. O'Malley, Cambridge: Cambridge University Press, 1982.
23. Nietzsche, F. *Daybreak. Thoughts on the Prejudice of Morality*, Transl. by J. Hollingale, Cambridge, Cambridge University Press, 2006.
25. Nietzsche, F. *Human, All Too Human*, Transl. by G. Handwerk, Stanford: Stanford University Press, 1997.
26. Nietzsche, F. *On the Genealogy of Morals*, Transl. by D. Smith, Oxford: Oxford University Press, 1996.
27. Nietzsche, F. *The Antichrist*, Transl. by A. Knopf, Białobrzegi: Masterlab, 2014.
28. Nietzsche, F. *The Gay Science*, Transl. by J. Nauckhoff & A. Del Caro A., Cambridge: Cambridge University Press, 2001.
29. Nietzsche, F. *Thus Spoke Zarathustra*, Transl. by A. Del Caro A., Cambridge: Cambridge University Press, 2006.
30. Reginster, B. Nietzsche on Ressentiment and Valuation, *Philosophy and Phenomenological Research* 57 (1997), pp. 281–305.
31. Ricoeur, P. *Freud and Philosophy: An Essay on Interpretation*, Transl. by D. Savage, New Haven: Yale University, 1970.
32. Schacht, R. *Nietzsche*, London: Routledge and Kegan Paul, 1983.
33. Scheler, M. *Ressentiment*, Transl. by W. W. Holdheim, New York: Schocken, 1972.
34. Schreber, D.P. *Memoirs of My Nervous Illness*. New York Review of Books: New York, 2000.
35. Segal, H. *Dream, Phantasy, and Art*, London: Routledge, 1991.
36. Žižek, S. *The Plague of Fantasies*, London: Verso, 1997.
37. Žižek, S. *The Sublime Object of Ideology*, London: Verso, 1989.

## Notes

1. The article was written as part of “Sonata” grant no 2013/09/D/HS1/00348 by Narodowe Centrum Nauki.