Filozofija i društvo, godište XXIX, broj 4
izdaje / published by
Institut za filozofiju i društvenu teoriju
Kraljice Natalije 45, Beograd, telefon: +381112646242
Email: institut@instifdt.bg.ac.rs
www.instifdt.bg.ac.rs

redakcijafid@instifdt.bg.ac.rs

Prilozi objavljeni u *Filozofiji i društvu* indeksirani su u Web of Science (ESCI), Scopus, ERIH PLUS, Philosopher's Index, EBSCO, PhilPapers, ResearchGate, Genamics JournalSeek, Google Scholar, J-Gate, ProQuest, ReadCube, Europeana Collections, Journal Index, Baidu Scholar

Dizajn: Milica Milojević

Lektura: Edvard Đorđević

Grafička obrada: Sanja Tasić

Štampa: Donat Graf, Beograd

Tiraž: 300. Časopis izlazi četiri puta godišnje.

Cena 350 dinara; godišnja pretplata 1200 dinara.

Objavljivanje časopisa finansijski pomaže Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije.

Univerzitet u Beogradu
Institut za filozofiju i društvenu teoriju

FILOZOFIJA I DRUŠTVO
PHILOSOPHY AND SOCIETY

broj 4 2018.
godište XXIX

Beograd 2018.
YU ISSN 0353-5738 UDK 1+316+32
3

REVIEW ESSAY
KRITIČKI OSVRT

REVIEWS
PRIKAZI

I

EDITED BY MILJANA MILOJEVIĆ

EXTENDED MIND AND EXTENDED COGNITION

PROŠIRENI DUH I PROŠIRENA KOGNICIJA

UDK: 159.95
https://doi.org/10.2298/FID1804473V
Original Scientific Article
Received: 05.08.2018. Accepted: 26.10.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Gottfried Vosgerau

VEHICLES, CONTENTS AND SUPERVENIENCE

ABSTRACT

In this paper, I provide an argument for the assumption that contents supervene on vehicles, which is based on the explanatory role of representations in the cognitive sciences. I then show that the supervenience thesis together with the explanatory role imply that the individuation criteria for contents and vehicles are tightly bound together, such that content internalism (externalism) is in effect equivalent to vehicle internalism (externalism). In the remainder of the paper, I argue that some of the different positions in the debate stem from different research questions, namely the question about the acquisition conditions and the question about the entertaining conditions for mental representation. Finally, I argue that the thesis of externalism is much more interesting if understood as a claim about how mental representation works in our world as opposed to how they work in all metaphysically possible worlds. In particular, I argue that this "nomological" understanding of the thesis is able to explain how and why the experimental methods used in contemporary cognitive sciences are able to provide insight into behavior generation.

KEYWORDS

externalism, semantic externalism, vehicle extrenalism, supervenience, empirical methods

There are two aspects of each (mental) representation, namely the content and the vehicle, which should not be confused (Hurley 1998). Since there is a clear conceptual difference between the two, theses about them clearly state different things. For example, the thesis of content externalism (Putnam 1975, Burge 1979) and the thesis of vehicle externalism (Clark 2008) are about two very different state of affairs; so different indeed that some claim the two theses to be independent of each other (Rupert 2004). The aim of this paper is to show that, although there is a clear conceptual difference, the two concepts are nevertheless strongly bound together; so strong that claims about the one entail certain views on the other. Of course, all this heavily depends on our understanding of what contents, vehicles and representations are. My understanding of these terms is based on the praxis that we find in behavioral sciences and that led to the so-called "cognitive turn"; thus, I will only focus on understandings of these terms that are compatible with behavioral science praxis.

My argument will proceed in three steps. The first step shortly discusses the notion of representation with respect to its explanatory role as we find it in behavioral sciences. I will argue that representations are assumed (or assigned to cognitive

Gottfried Vosgerau: Professor, Institut für Philosophie, Heinrich-Heine-Universität Düsseldorf; vosgerau@hhu.de.

systems) as theoretical entities that explain flexible behavior, which serves as the first premise of my argument (section 1). The second step establishes the second premise, namely that contents must supervene on vehicles if they are to fulfill this explanatory role. Since there are multiple possible ways of understanding the notions of content, vehicle and representation, I will try to clarify my understanding of them by defending my argument against possible objections that might arise on the basis of different understandings not compatible with behavioral science praxis (section 2). In the third step, I show that (given the premises) content externalism logically entails vehicle externalism, and that vehicle externalism entails content externalism (not strictly logically but in effect; section 3). I conclude that every argument for content externalism is in effect an argument for vehicle externalism and the other way round.

In section 4, I introduce the distinction between acquisition conditions and entertaining conditions as two different explananda for a theory of representation: It is one thing to explain how it is possible to acquire a certain mental representation and another thing to explain what it means to entertain a representation. This distinction is scrutinized in terms of supervenience relations. Then I sketch the debate between Alva Noë and Ned Block, showing that Noë argues mainly for acquisition conditions, while Block offers good arguments regarding entertaining conditions. Although both authors contend to say something about the metaphysics of mental representations (i.e. about the metaphysically necessary conditions for mental representation), I conclude by arguing that externalism is better understood as a claim about the nomological conditions that constrain mental representation. In this way, its impact on behavioral sciences can be much better accounted for.

1 Representation and Explanation

Let me start with the notion of representation.[1] Mental representations are introduced in order to explain behavior. This is at least the idea on which cognitive (behavioral) science is built: There are kinds of behavior, namely flexible behavior, which cannot be explained by stimulus-response patterns. Flexible behavior is understood as behavior that, given one and the same type of stimulus, can still differ. (This general description is meant to include cases in which the behaving system does not have a relevant input at all.) Since this implies that there is no simple one-to-one relation between stimulus and behavior, flexible behavior is not explainable by simple stimulus-response patterns. Rather, or so the reasoning goes, we have to assume that some kind of inner state of the behaving system also has an influence on what kind of behavior is selected given a specific stimulus. These inner states are, moreover, assumed to stand for something else and are hence called "mental representations". In the words of J. Haugeland: "[...] if the relevant features are not always present (detectable), then they can, at least in some cases, be represented; that is, something else can stand in for them, with the power to guide behavior in their stead. That which stands in for something else in this way is a *representation*"

1 I am not going to sketch an account of mental representation here, but rather introduce some basic considerations on why mental representations are introduced, i.e. on what the explanatory role of them is. For an account of mental representation see (Vosgerau 2009).

(Haugeland 1991: 62; original italics). In this sense, we have to assume some kinds of mental representation to explain flexible behavior.[2]

A typical example of (a simple) flexible behavior that is explained by mental representation is the so-called "homing behavior" of the desert ant: The ant is able to go straight back to its nest after an unsystematic search for food (cf. Gallistel 1993). If the ant is displaced before it starts its way back, it will not return to the nest but take a route parallel to the "correct" route that leads it to the point where the nest would have been, had it been displaced in the same way as the ant. This already shows that the ant does not have the relevant features available, i.e. it has no access to relevant features of the nest that could "directly" guide its behavior. Thus, the ability of the ant to return to its nest (i.e. the homing behavior) can only be explained if we assume that the ant has some kind of representation of the location of its nest (Vosgerau, Schlicht and Newen 2008). It is a representation of the location of the nest *because* it explains the ant's ability to find its nest (under normal or favorable circumstances; i.e., errors may occur under certain circumstances). If the ant behaved differently, e.g., if it went to trees instead of its nest, we would have to assume a different representation, namely a tree-representation.[3] A very similar argument was brought forward by R. Cummins against teleological theories of mental representation; his example is: if we observed bee dances that indicate the location of piles of rocks, than we would have to assume that they represent the location of piles of rocks, "even though we would be mystified about the evolutionary significance of the whole business" (Cummins 1989, 86). Therefore, the content of representations is, in the first place, determined by the object of the behavior to be explained. (There might be further criteria for content individuation, e.g., Frege's principle of simultaneous believability [Frege 1892]. However, such further criteria can never tell us why it is a nest-representation rather than a tree-representation.) In other words: If the ant would not interact[4] with its nest, we would have no reason to assume that it has a nest-representation at all.

The point to be made here is not just epistemological but rather conceptual: The concept of representation can only be applied to things that stand for something (the represented entity). "Standing for something" means (at least in the case of mental representations) that the representation can be used by a behaving system to engage in behavior directed towards the represented entity without having direct sense-contact with it.[5] This means that the representation enables the behaving system to act as if it had direct sense-contact with the represented entity.

---

2 This idea can also be found in the pioneering work that led to the assumption of mental representations, e.g. in Tolman and Honzik (1930), Tolman (1948); (cf. Vosgerau 2010).

3 The formulations just given might be read to point to a dispositional understanding of mental representations. Because of space limitations, I cannot discuss this idea—which I reject—here. For a detailed discussion see Vosgerau (2010).

4 I use "interaction" in a wide sense, which includes avoidance behavior. E.g., when a beaver flees from an eagle, it interacts with the eagle in this sense; it displays, as I will say, an eagle-directed behavior, and the eagle is the object of the fleeing-behavior.

5 This is true at least for perceptual representations. There might be further (conceptual) representations that are built by combining representations. Such conceptual representations can be far more abstract in that they might represent things that cannot be the target of behavior. However, it is very plausible to think that such conceptual representations are based on perceptual representations (cf. Vosgerau 2009).

Crucially, the basic idea is still that representations fulfill a certain explanatory role: *Because* representations are substitutes for the represented entities, the bearer of the representation is able to show this and that behavior. In a full account, however, this basic principle has to be spelled out in more detail, for example by explaining the explanatory roles in terms of causal roles. Nevertheless, on a more abstract level of description, the content of the representation, which specifies the represented object, will explain the behavior. In the case of the ant, the content of the representation can be described as location of the nest, and it enters the explanation of the ant's homing behavior at the place where the location of the nest (or the detection of the relevant features of the nest, resp.) would occur if the ant had sense-information about the nest.

2 Supervenience as the Relation Between Contents and Vehicles

Let me now turn to the second step in my argument, namely the defense of the claim that contents supervene on vehicles. The distinction between contents and vehicles is widely accepted, and it is widely agreed that we should not confuse the two (Hurley 1998). Accordingly, a strict difference between content externalism (going back to Putnam 1975, Burge 1979) and vehicle externalism is drawn, since they amount to radically different claims (Rupert 2004). While I agree that these are, in principle, two different issues, the question arises how they are connected. Is it, for example, possible to hold the one but to deny the other? The answer to this question depends upon one's view on content and vehicle individuation: If contents are individuated independently from vehicles and vice versa, it is possible to separate the two issues clearly. But if the individuation of vehicles is dependent on the individuation of contents, theses about the "location" of contents have implications for the location of vehicles. In this paper, I argue that there is a systematic relationship between content individuation and vehicle individuation, namely the supervenience of contents on vehicles.

Contents are introduced in order to explain behavior (see above). Such explanations are usually understood as causal explanations ("the ant goes back to its nest because it has a representation with this and that content"). If we now ask what the vehicles of these contents are, it seems quite natural to say that contents supervene on their vehicles. My argument for this claim takes the form of a *reductio*: Firstly, vehicles are things that play causal roles in the behaving system that displays the behavior. Now assume that the content of those vehicles does not supervene on them; then it would be possible that one and the same vehicle has different contents (this follows from the definition of supervenience). This would mean that there could be two different content-based explanations for a certain behavior (which involves one and the same vehicle) for which there is only one causal story on the vehicle level. From this we can infer that the contents of mental representation would not be apt to figure in causal explanations of behavior, i.e. they would have no causal roles. This, however, contradicts our idea of providing causal explanations of behavior based on content. From this contradiction we can infer that the assumption was wrong, and therefore, that contents supervene on vehicles.

Of course, contents are but one of the causes of behavior among other causes. This could lead to the following reasoning: Contents are based on (or involve) more

causal roles then just the causal role of the vehicle, i.e. content is individuated externalisticaly and vehicles are individuated internalisticaly. Take, e.g., the mental state of seeing a tree. It could be argued that the tree plays a causal role relevant for the individuation of the content of this mental state, yet the vehicle is the according brain state that is internal. How then would the vehicle be individuated? It could be individuated as the internal part of the supervenience base of the content. However, this criterion for vehicle individuation is rather ad hoc and does not seem to have an independent support (indeed, it would dogmatically presuppose vehicle internalism). Moreover, it would give us two different causal explanations: one on the content-level and a different one on the vehicle-level.[6] And what then would make this internal state the vehicle of this content (rather than another content)? There is no principle in sight which could answer this question. So, if we want to construe the content of seeing-a-tree externalistically (i.e. construing "seeing" as a success-verb), we should take its vehicle to involve the tree as well; if we want to talk about the brain state as the vehicle, we should construe the according content as having-the-impression-of-seeing-a-tree (we could be mistaken after all), which is individuated internalistically as well.

It might be argued that my reductio is way to simple, given the sophisticated arguments by Burge (1979) to the contrary effect. But, Burge's arguments for content externalism do not contradict my claim, since Burge merely establishes the thesis that the content of mental states does not supervene on brain states (or, more generally, inner states of the representing system). However, he does not claim that brain states are the vehicles of mental content. Therefore, his claim is compatible with the claim that contents supervene on vehicles. It is only incompatible with the claim that brain states are vehicles: "CE [content externalism] has it that the supervenience base for states with mental content includes external physical features" (Sprevak and Kallestrup 2014: 82).

Furthermore, Burge's argument (the twin earth argument) is a variation of Putnam's (1975) argument for externalism of linguistic meaning. Burge adds the premise that the content of mental states is individuated by (or even identical to) the content of the linguistic expression we use to describe the mental state. This notion of mental content is, however, quite obviously not compatible with behavioral science praxis: The explanation of the homing behavior of the ant is not dependent on how we describe its mental representation. So, if we want to go along with cognitive science and ascribe a mental content to the ant that has a role to play in the explanation of the ant, we better not use a notion of content that is language-dependent such as Burge's notion (cf. Newen and Vosgerau 2007).

Another line of reasoning might be that not only content but also behavior could be individuated broadly or narrowly. And if so, content could play a role in explaining broadly (i.e. externalistically) individuated behavior, while the vehicle only plays a role in explaining the narrowly (i.e. internalistically) individuated behavior. In this way it might be that vehicles are internal and contents external, implying that contents do not supervene on vehicles. Again, the notion of broadly

6 While it is plausible that content explanations might be on a different level of abstractness (or fine-graindness), we still expect a systematic correspondence between the two levels if both are considered to deliver causal explanations (Soom 2011).

individuated behavior is at odds with the praxis of behavioral sciences: It is crucial that the experimentally induced cases, in which the ant does not reach its nest, are equally subsumed under the same type of behavior, namely the homing behavior. If behavior would be individuated broadly, these cases would have to count as cases of a different type of behavior (because it does not involve the nest), in which case they could not give us any evidence about the case we are interested in, namely the ability of the ant to find its nest. Thus, the notion of behavior that is used in the behavioral sciences has to be the narrowly individuated one.

Sprevak and Kallestrup (2014) claim that at least content externalism has to be confined to a claim about weak supervenience instead of strong supervenience (Sprevak and Kallestrup 2014: 82), building on arguments from Stalnaker (1989) and Jackson and Pettit (1993) concerning narrow content. And if so, then it simply is not true that one and the same vehicle cannot have different contents: This is only excluded within the same possible world but not in two different possible worlds. However, the argument given for the individuation of narrow content cannot be transferred to content externalism, at least not to the Burge-style content externalism they, and I, have in mind: The Twin-Earth story involves two different possible worlds; thus, it would not tell us anything about the supervenience base for contents then, especially not that "the supervenience base for states with mental content includes external physical features" (Sprevak and Kallestrup 2014: 82). In particular, it would cease to be an argument for the thesis that mental content does not supervene on brain states. Thus, I will proceed with the assumption that strong supervenience is the better option for content externalists.

Last not least, let me shortly comment on the claim that vehicle externalism is the claim that "the mechanisms of human cognition extend outside the brain and head" (Sprevak and Kallestrup 2014: 83). Interestingly, later in the paper the relevant formulations of the different versions of vehicle externalism include reference to, e.g., "mental processes/states" (Sprevak and Kallestrup 2014: 85). Although I do not wish to argue that this understanding is somehow wrong or inadequate, I would like to point to a few problems that lead me to adopt a more specific notion of vehicle than the one implied by the passages above. In particular, this will be the notion that is referred to earlier in their paper, namely the following: "The vehicle of content is the physical item that has, or expresses, that content – for example, a sentence, if we talk about linguistic content" (Sprevak and Kallestrup 2014, 78). Clearly, this is a state, not a process. The process of writing does not have or carry or express the content, rather the ink pattern on the paper does. And in parallel, it is some physical states (e.g. brain states) that are the vehicles of mental content, and not "some piece of cognitive architecture" (Sprevak and Kallestrup 2014: 78), let alone mental processes.[7] And while there is a version of vehicle externalism that talks about processes (playing Tetris and mental rotation, e.g.), my concern here is only with mental states, i.e. with the version of vehicle externalism that is about the notebook entry (not the writing down) of Otto, for example. And this is because my aim is to discuss the notion of mental representation that is used in the

7 Indeed, it seems pretty unclear what would make a mental process mental anyhow (cf. Fodor (2009) and Vosgerau (2010)), which is probably also reflected in the debate about a "mark of the mental" (Walter 2010).

behavioral sciences (see above), which is about the contentful states of animals. So, there are mental states that have two aspects: a content and a vehicle that has or carries the content. And for this understanding of mental states (representations), which I claim to be the relevant one for the praxis of behavioral sciences, I contend that my argument for supervenience holds.

3 The Relation between Vehicle and Content Externalism

I therefore take it that contents supervene on vehicles (if they figure in causal explanations). According to the standard formulation of supervenience (Kim 1984), properties of kind A supervene on properties of kind B iff: $\Box \forall x \, \forall F \in A \, [Fx \rightarrow \exists G \in B \, (Gx \wedge \Box \, \forall y (Gy \rightarrow Fy))]$[8]. In words: Necessarily, if something has A-property F, then there is a B-property G that it also has and, necessarily, anything else having G also has F. For our case, content types (the property of being an instance of this content type) are A-properties while vehicle types (the property of being a vehicle of this type) are B-properties. So, if contents supervene on vehicles, it is impossible that one and the same vehicle has two different contents. At the same time, it is possible that one and the same content has different vehicles ("multiple realization").

Equipped with the two premises (that representations explain flexible behavior and that contents supervene on vehicles), we are now able to move to the third step of the argument, which investigates the relation between content externalism (internalism) and vehicle externalism (internalism). To show that this relation is in effect an equivalence relation, I will discuss both directions of the biconditional in turn.

Assume that content externalism is true, i.e. that at least some contents essentially depend on factors in the environment of the behaving system. Then it is possible to change a content by changing some factor in the world while keeping the state of the behaving system fixed (exactly this is done in the Twin-Earth thought-experiments). Then, there are two different contents, *F1* and *F2*, for each of which there has to be some vehicle on which it supervenes. Since these two vehicles have to be distinct according to the supervenience definition (otherwise, the second implication would be false), and since the internal states of the behaving system are not distinct, the vehicles have to extend beyond the boundaries of the system into the world. Thus, if contents supervene on vehicles, content externalism implies vehicle externalism.[9] (By *modus tollens*, vehicle internalism implies content internalism.)

What about the other direction: does vehicle externalism imply content externalism? Hurley (1998: 3) argues that multiple realizability is the reason why not:

8 This is the standard definition for strong supervenience which is sufficiently adequate for the present discussion. Other versions of supervenience either are not apt for the present discussion or are designed to minimize certain difficulties which are not important for the present discussion. (See also above and the following footnote.)

9 There are attempts to formalize the notion of supervenience such that extrinsic properties like externalistically individuated content can be handled (e.g. Hoffmann and Newen 2007). However, the main point in these formalizations is to find a general constraint that excludes properties from the supervenience base that are irrelevant for the higher-order property; so, these accounts can be said to provide a formalization of the notion of a "minimal supervenience base". The presented argument is, however, independent of how exactly the minimal supervenience base is determined.

While there might be cases in which a content supervenes on internal states, this does not mean that all instances of this content do. In the definition, this fact is expressed in the existential quantifier: We do not know how many *G*s there are, and, for a given content *F*, some *G*s may involve external features while others do not. The reason for this possibility is, according to Hurley, that vehicles are token-explanatory while contents are type-explanatory. Thus, for different tokens of the same (content-) type, there can be different features playing an explanatory role.

However, this picture does not work at least for the purpose of explaining behavior. The reason is that the behavior which we want to explain is always a type of behavior. It might be (metaphysically) possible that the vehicle of the ant's nest-representation involves the nest sometimes (maybe even exactly the times at which the homing behavior is successful), and does not involve the nest at other times (namely in cases of misrepresentation). But how do we find out about this? We cannot, since the explanations we provide are type-explanations. This is just the way experiments work: Different factors are systematically controlled while the single tokens of behavior are counted as belonging to one type—the type that we call "the behavior" (e.g. the homing behavior of the ant). If we would not count cases of (experimentally induced) misrepresentation as belonging to the same type of behavior, then the experimental conditions could not tell us anything about the behavior we want to explain. And so in explaining "the" homing behavior, we have to assume that the nest does not play an essential part because it is not involved in all situations in which the behavior is displayed. If we are really interested in explaining one successful token of the homing behavior, we still resort to the general pattern which we explored for the type. Thus, there is no way of formulating a criterion for individuating the vehicle for this token without reference to the type: There is no reason to assume that in the successful case the internal features of the ant are radically different from the (experimentally induced) case of misrepresentation. Therefore, the explanation we give for tokens relies on the type-explanation we have given. In other words: If, like Hurley proposes, vehicles are token-explanatory, then we need individuation criteria independent from content (since content is type-explanatory); however, there are no such criteria,[10] and if there were, they would not help to explain behavior.

This point can be demonstrated more formally with the help of the notion "minimal supervenience base". The definition of strong supervenience is formulated such that it is always possible to include in *G* every arbitrary feature for any token of *F* (the "irrelevant feature problem"; cf. Hoffmann and Newen 2007). However, what we are interested in is the minimal supervenience base, i.e. the smallest set of *G*s (which does not contain irrelevant features) which renders the definition true (for some specific *F*). Cases of misrepresentation, in which the represented object is not present (or does not have the represented features), show that the minimal supervenience base for these cases does not include features external to the behaving system. This provides very good reason to think that the minimal supervenience base in cases of successful representation is internal, too. Indeed, this is the rationale of experimental research (as discussed above). If there is no case of

10 Vehicles could be a natural kind with essential properties for which the reference is fixed by ostention. This, however, seems to be a very mysterious claim.

misrepresentation, then the "minimal supervenience base" is likely to include external features—however, these cases are not cases of mental representation since the displayed behavior is not flexible. For this reason, cases of misrepresentation constitute very good cases for internalism (pace Hurley).

One might argue that further external features are involved in the ant's representation since it has been shown that the ant represents the direction of the way to the nest relative to the location of the sun. However, this only leads to a more specific characterization of the content of the ant's representation: location of the nest in terms of the direction relative to the sun (and distance). Although this implies that the ant has to have information about the location of the sun in order to successfully use the representation, it does not follow that the sun is part of the minimal supervenience base of this content. However, such a "relativization" of the content is not possible for the nest itself, and thus the nest is not only excluded from the minimal supervenience base, but it also does not play a role in the explanation of the behavior (unlike the location of the sun).

To conclude, vehicle externalism does not imply content externalism logically. However, conjecturing that for some tokens the vehicles are (partly) external (although the content is fully internal) leads to a non-testable hypothesis that contradicts experimental praxis and scientific explanation. Such a thesis would claim that the vehicles of misrepresentations are systematically different from the vehicles of successful representations with the same content. The reason for its non-testability is that tokens of vehicles cannot be individuated independently of content, so that the same criteria for vehicle individuation apply to both successful and misrepresentations. I therefore conclude that every sensible version of vehicle externalism goes hand in hand with content externalism and vice versa. Therefore, the sharp distinction between content and vehicle externalism is very helpful at the conceptual level, but it is not very important in arguing for one or the other.

4 Acquisition and Entertaining of Representations

Indeed, philosophers arguing for externalism (for some mental representation, e.g. perceptions) usually claim that the external features play an *essential* role for representation, i.e. that in every case of representation, the external feature is involved. Combining this with the thesis that content supervenes on vehicles, this kind of argument is both an argument for vehicle as well as for content externalism. As an example I will discuss Noë (2005), who argues that the vehicles of percepts extend beyond the boundaries of the perceiving system. He claims that the perceived object plays an essential part in constituting the percept, because percepts are based on "knowledge" about the systematic contingencies between one's own movements and the changes in the sense input coming from the perceived object.[11]

11 Noë (2005) sometimes seems to make the more moderate claim that only parts of the body are an essential part of representations because movements are (and thus representations are) in the body as opposed to only in the brain. However, the systematic contingencies do in fact involve the environment, such that I will concentrate on this claim. Moreover, he suggests that every occurrence of a percept has to involve an active component—a claim which I will not discuss here (for a critical discussion see, e.g. Block 2005, Jacob 2006).

Block (2005) criticizes this position by claiming that Noë does not talk about the minimal supervenience base. He holds, contrary to Noë, that in the end, vehicles and contents have to be internalistically individuated, even if the represented objects play an essential part in the *normal* way of acquiring the representations in question.

In order to evaluate both positions and the involved arguments with respect to the explanatory role they assign to mental representations, I will first introduce the distinction between the acquisition conditions and the entertaining conditions for mental representations: It is one thing to explore the conditions that have to be fulfilled to acquire a certain representation, and another thing to explore what conditions have to be fulfilled to entertain a given content (provided that the system has acquired this content). With this distinction at hand, I will show that the debate between Noë and Block arises only because they are talking about different things; there is no real controversy. Before doing so, I will now introduce this distinction in detail.

4.1 Historical Externalism

Consider, again, the homing behavior of the ant: The ant has a nest-representation (rather than a tree-representation) because it displays nest-directed behavior. If there were no nests, there would be no nest-directed behavior, and thus there would be no nest-representations. In other words, the ant could never perform any nest-directed behavior and *a fortiori* could not acquire the nest-representation, if it had never interacted with nests.

Since the ant can only have a nest-representation if it has had some interaction with nests, the supervenience-base of the nest-representation has to involve external features: It is always possible that there is an ant with the same internal state as the ant that is displaying homing behavior, but which is living in a world without nests and thus does not have a nest-representation. Therefore, the internal state of the ant cannot function as the *G* in the definition of supervenience, since there are some worlds in which the last implication is false. So, the supervenience base *G* has to contain past interactions with nests. Hence, the supervenience base has to include nests (at least past interactions with nests) and so extends beyond the boundaries of the ant.

Independent of the specific story about perceptual states told by Noë (2005) (but fully compatible with it), mental representations are externallistically individuated (both the contents and the vehicles) in the following sense: It is not possible to display a certain behavior and, *a fortiori*, to acquire a certain mental representation without having had interactions with the represented object. The point is not that the represented object has to be there every time (there are cases of misrepresentation), but that there must be a certain history of the representing system in which it successfully acquired the representation in question. I will call this version of externalism "historical externalism" (authors like Dretske, Lycan, and Tye have defended such an externalism for contents), which follows from the supervenience thesis of content and vehicles if we take into account the acquisition conditions for content, which corresponds to evaluating all possible worlds to interpret the necessity operators in the supervenience definition.

4.2 Entertaining Content

If we aim at fully accounting for the acquisition conditions for mental representations, then it seems that we are forced to adopt historical externalism. However, if we abstract from the acquisition conditions, assuming that the behaving system in question already possesses the ability to entertain a this-and-that-representation, we can focus on the entertaining conditions: What is required of a behaving system to entertain a certain representation (given that it has acquired the ability to do so)? For this end, let us restrict the set of possible worlds which we take into account in evaluating the definition of supervenience. The basic idea is that we consider only worlds in which the system in question has had interactions with the object in question. Since we are dealing with temporal relations, I will use standard temporal modal logic in which a possible world is a world at a certain time-point (whatever a time-point may be).

Let us introduce the notion of an α-world for a behaving system *s*, which is a world in which *s* possesses the ability to entertain an *A*-representation. Such a world has to be preceded by at least one world in which *s* has had interaction with *A*s (within a certain time range).[12] Thus, for every α-world of *s*, the following must be true: $\exists t_i \left[t_{\alpha-\epsilon} < t_i < t_\alpha \wedge V\left(t_i, \exists x\left(Ax \wedge I(s,x)\right)\right) = True\right]$.[13] At least in one world i_i within a certain temporal range ϵ before t_α, the sentence "some *A* exists with which *s* interacts" has to be true. We can now define the set of α-worlds for *s*:

$$\mathcal{W}_\alpha^s = \{t_j \mid \exists t_i\, [t_{j-\epsilon} < t_i < t_j \wedge \mathcal{V}\left(t_i, \exists x\left(Ax \wedge I(s,x)\right)\right) = True]\}$$

Abstracting from the acquisition conditions now means to consider only α-worlds for evaluating the definition of supervenience. (Of course, due to the use of temporal logic to define the sets of possible worlds, we are not allowed to use implicit temporal predicates like "…has had interactions with …" in the object language.) If we do so, the argument for historical supervenience does not work anymore, since now there is no possible world left in which the system fails to have the right acquisition history. And, as we know from cases of misrepresentation, the actual entertaining of a representation does not imply the existence of the represented object. So, if we only look at possible worlds (i.e. worlds at a certain time-point) in which *s* has acquired an *A*-representation (by reducing the set of relevant worlds to those with the right kind of acquisition history), then the presence or absence of *A*s does not change the content of the representation anymore. Thus, there are possible α-worlds in which there is no *A* although *s* entertains an *A*-representation. Accordingly, *A*s cannot be included in the minimal supervenience base if we abstract from the acquisition conditions. Thus, if we focus on the conditions of entertaining a representation, we will have to take cases of misrepresentations seriously and we will thus conclude that both the vehicle and the content are individuated internally. This point seems to be made by Block (2005) when he distinguishes between

12 The limitation to at least one world is only for simplicity of the formula; in fact, it is plausible to assume that the system has to interact with the object several times.

13 I employ the standard notation of temporal logic, where t_i is the world at time-point t_i, and $\mathcal{V}v$ is the evaluation function that maps worlds and formula onto truth values. Moreover, the two-place predicate $I(\nu,\mu)$ is used to express that ν interacts with μ.

causal influence and constitutive factors. He claims that everything that is constitutive for *entertaining* a certain mental representation is internal. At the same time, he does not contradict the idea of historical externalism, presumably because it is not concerned with the constitutive factors for *entertaining* a representation but with the constitutive factors for *acquiring* a representation.

4.3 Metaphysical or Nomological?

The debate about externalism is no doubt a debate about the metaphysics of mental representations. So, advocates of externalism take their claim to be a metaphysical claim (understood as a claim concerning the nature of mental representations) and are thus confined to a metaphysical reading of the necessity operator. In this section, some core arguments will be discussed regarding this claim.

Noë (2005) cites the cat experiments from Held and Hein (1963) to support his externalistic claim. In this experiment, two kittens were put in a vertically striped cylinder. One was allowed to actively explore its surroundings, while the other was passively moved around in the same way the active kitten moved (such that the visual input for both was identical). They could show that the passive kitten had severe deficits in perception compared to the active kitten. Thus, it seems that active exploration of (interaction with) the environment is necessary for the acquisition of perceptual representations. Block (2005) argues that this experiment does not show that active exploration is metaphysically necessary: His counter-argument is that if in the passive kitten the same brain processes would be going on (however they came about), it would be absurd not to believe that the passive kitten had perceptions.

Let us assume that the argument is a good argument against the metaphysical necessity of a specific kind of acquisition history. However, does this mean that it is also a good argument against nomological necessity, i.e. the claim that given the "natural laws" of our world, representations can only be acquired in this way? How could it be that the passive kitten has the same brain states, if not by chance? We could imagine that the passive kitten's brain is connected by wires with the active kitten's brain in an appropriate way. However, in this case it (presumably) would have to be so wired that the motor commands of the passive kitten are also able to control the active kitten. Then, at least two problems arise: (1) Which kitten does effectively control the movements of the active kitten? (2) If the passive kitten could be said to control the movement of the active kitten and to have the perceptual input of the active kitten, then all requirements for having perceptual representations are met, according to the "enactive approach"—we would just be confronted with the weird situation that the brain which processes the relevant signals is at a physical place different from that of the acting body. Thus, Block's argument (Block 2005) cannot show that active exploration of the environment is not nomologically necessary for acquiring representations. (It probably was not intended to show that anyway.)

If the passive kitten had the same brain states by chance, then this thought experiment is essentially the same as the famous "swampman" experiment by Davidson (1987), in which by coincidence a physical twin of Davidson comes into being. The question is whether this swampman, which of course behaves exactly like Davidson, has mental representations or not. According to the picture of mental

representations drawn above, we would have to admit that swampman indeed has mental representations, since he shows flexible behavior which we could not explain otherwise. So, if swampman is a serious metaphysical possibility (and there is not much reason to doubt that, although there is much reason to doubt that it is nomologically possible), then the specific way of acquiring a certain representation cannot be a metaphysical necessity for entertaining this representation. The necessity operator in the supervenience definition therefore has to be understood in terms of nomological necessity, at least for the case of historical externalism. (And in the case of entertaining conditions, the nomologically necessary acquisition condition is already built into the definition of the possible worlds under consideration.)

Thus, as far as the metaphysical conditions for being a representation are concerned, I think that there are always defeating arguments against externalism. It is always metaphysically possible that a physical duplicate of some behaving system is created by chance. There is, by presupposition, no reason to think that this newly created system would behave differently from the "original" system. And, since mental representations are to explain flexible behavior, we have to attribute to these newly created systems the same representations which we ascribe to the original system. Therefore, metaphysically spoken, externalism is wrong. However, it can still be a nomological necessity. The conclusion that we draw from this result depends on the question we ask. If we are interested in the metaphysics of mental representations, the final result is internalism. However, I think that there are good reasons to be even more interested in the results of the nomological discussion: If we want to understand how empirical research works and how it can possibly shed light on mental representation, we should care much more about nomological necessity than about metaphysical necessity. Insofar historical externalism is nomologically true, evolution and psychological development play a major role in the explanation of the acquisition of mental representations. And since the entertaining conditions for mental representations are internal, we can study mental representations in behaving systems by neuroscientific methods that do not account for external factors. The most important conclusion of this paper is hence that the most fruitful debate for the understanding of mental representation is not the metaphysical debate about externalism and internalism, but the debate about the nomological conditions of mental representation, which has direct impact on our way of investigating mental representations in the empirical sciences.

5 Conclusion

Representations are assumed in order to explain flexible behavior. The idea that contents supervene on vehicles is defended on the grounds that otherwise, the explanatory role of representations and/or the praxis of experimental behavioral sciences could not be secured. It then turns out that content externalism implies vehicle externalism. Although vehicle externalism does not logically imply content externalism, it is nevertheless implausible to defend vehicle externalism and content internalism at the same time, since this position (1) has no individuation criteria for vehicles to offer and (2) is not able to do justice to the explanatory role of vehicles in the explanation of behavior. I conclude that arguments for or against externalism are always arguments for or against both vehicle and content externalism.

Taking the debate between Noë (2005) and Block (2005) as an example, I have shown that possibly the two are talking about different things. My proposal is to distinguish acquisition conditions for mental representations from conditions of entertaining a certain representation. While the arguments of Noë (2005) are (plausibly) concerned with acquisition conditions, the counter-arguments of Block (2005) (notedly) discuss entertaining conditions. While it is very plausible that the acquisition of representations does indeed depend on a certain learning history that essentially involves (parts of) the environment, the entertainment of a certain representation does not rely on the environment (provided that the system in question has acquired the ability to entertain such a representation). If we look at all possible worlds, there might always be a behaving system having the same internal state as one that has a certain representation *r*, but that fails to have *r* as well, because it has never interacted with instances of the represented object and so cannot even exhibit the same kind of behavior. Therefore, we have to include external factors in the minimal supervenience base for contents (historical externalism). However, if we restrict the set of possible worlds that we evaluate to those worlds in which the system in question has the right acquisition history, then there is no reason to include external factors into the minimal supervenience base. The reason is that there are always cases of misrepresentation in which the represented object (or represented aspects of objects) are not present. Hence, when we focus on the acquisition conditions for mental representations, we should adopt a form of historical externalism; however, if we focus on the entertaining conditions for a certain representation, then we have good reasons to defend an internalistic view. Moreover, historical externalism is plausible only if we interpret the necessity operator in the supervenience definition as indicating nomological necessity, which might be even more interesting with respect to empirical work.

I have tried to approach some of the debates about externalism from a point of view that takes seriously the experimental praxis of behavioral (cognitive) science. In particular, I have argued that some notions of content and vehicle that we find in the literature on externalism are not apt to do justice to this praxis and to scientific explanation of animals' behavior. If we concentrate, however, on this praxis, a lot of possible confusions about contents and vehicles can be avoided. Moreover, the discussion of the kind of necessity that is involved in supervenience claims showed that understanding externalism as a claim about how mental representations work in our world (as opposed to a claim about the metaphysics of representation being true in every possible world) is much better suited to connect to empirical research. In particular, it is then possible to explain how and why the experimental methods used today are able to provide insight into behavior generation in cognitive systems.

References:

Block, N. (2005), "Review of Alva Noë, Action in Perception", *The Journal of Philosophy* 102: 259–272.

Burge, T. (1979), "Individualism and the Mental", *Midwest Studies in Philosophy* 4: 73–122.

Clark, A. (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, Oxford University Press: Oxford.

Cummins, R. (1989), *Meaning and Mental Representation*, The MIT Press, Cambridge MA: London.

Davidson, D. (1987), "Knowing One's Own Mind", *Proceedings and Addresses of the American Philosophical Association* 60: 441–458.
Fodor, J. (2009), "Where is my Mind? Review of Supersizing the Mind: Embodiment, Action and Cognitive Extension by Clark, A.", *London Review of Books (Online)* 31: 13–15; http://www.lrb.co.uk/v31/n03/jerry–fodor/where–is–my–mind, last access: 2016–03–31.
Frege, G. (1892), "Über Sinn und Bedeutung", *Zeitschrift für Philosophie und philosophische Kritik* 100: 20–50.
Gallistel, C. (1993), *The Organization of Learning*, The MIT Press, Cambridge MA.
Haugeland, J. (1991), "Representational Genera". *In:* W. Ramsey, S. Stitch and D. Rumelhart (eds.), *Philosophy and Connectionist Theory*, Hillsdale: Erlbaum, pp. 61–89.
Held, R. and Hein, A. (1963), "Movement-Produced Stimulation in the Development of Visually Guided Behavior", *Journal of Comparative and Physiological Psychology* 56: 872–876.
Hoffmann, V. and Newen, A. (2007), "Supervenience of Extrinsic Properties', *Erkenntnis* 67: 305–319.
Hurley, S. (1998), "Vehicles, Contents, Conceptual Structure, and Externalism", *Analysis* 58: 1–6.
Jackson, F. and Pettit, P. (1993), "Some content is narrow".In: J. Heil and A. Mele (eds.), *Mental Causation*, Clarendon Press: Oxford, pp. 259–282.
Jacob, P. (2006), "Why Visual Experience Is Likely to Resist Being Enacted", *Psyche* 12: 1–12.
Kim, J. (1984), "Concepts of Supervenience", *Philosophy and Phenomenological Research* 45: 153–176.
Newen, A. and Vosgerau, G. (2007), 'A Representational Account of Self-Knowledge', *Erkenntnis* 67: 337–353.
Noë, A. (2005), *Action in Perception*, Cambridge, MA: MIT Press, London.
Putnam, H. (1975), "The Nature of Mental States". In: *Mind, Language, and Reality*, Cambridge: Cambridge University Press, pp. 429–440.
Rupert, R. (2004), "Challenges to the Hypothesis of Extended Cognition", *The Journal of Philosophy* 101: 389–428.
Soom, P. (2011), *From Psychology to Neuroscience. A New Reductive Account*, Frankfurt: Ontos Verlag.
Sprevak, M. and Kallestrup, J. (2014), "Entangled Externalisms". In: M. Sprevak and J. Kallestrup (eds.), *New Waves in Philosophy of Mind*, New York: Palgrave Macmillan, pp. 77–97.
Stalnaker, R. C. (1989), "On What's in the Head", *Philosophical Perspectives* 3: 287–319.
Tolman, E. C. (1948), "Cognitive Maps in Rats and Men", *Psychological Review* 55: 189–208.
Tolman, E. and Honzik, C. (1930), "'Insight' in Rats", *University of California Publications in Psychology* 4: 215–232.
Vosgerau, G. (2009), *Mental Representation and Self-Consciousness. From Basic Self-Representation to Self-Related Cognition*, Paderborn: Mentis.
Vosgerau, G. (2010), "Memory and Content", *Consciousness and Cognition* 19: 838 – 846.
Vosgerau, G., Schlicht, T. and Newen, A. (2008), "Orthogonality of Phenomenality and Content', *American Philosophical Quarterly* 45: 309–328.
Vosgerau, G. and Synofzik, M. (2010), "A Cognitive Theory of Thoughts", *American Philosophical Quarterly* 47: 205–222.
Walter, S. (2010), "Cognitive extension: the parity argument, functionalism, and the mark of the cognitive", *Synthese* 177: 285–300.

Gotfrid Fosgerau

Posrednici, sadržaji i supervencija

Apstrakt

U ovom radu predstavljam argument u prilog pretpostavci da sadržaji supervenišu na posrednicima koji je zasnovan na eksplanatornoj ulozi predstava u kognitivnim naukama. Potom pokazujem da teza o superventnosti, zajedno sa eksplanatornom ulogom, implicira da su kriterijumi za individuaciju sadržaja i posrednika blisko povezani na takav način da je internalizam (eksternalizam) o sadržaju zapravo ekvivalentan internalizmu (eksternalizmu) o posrednicima. U ostatku rada, pokazujem da neki od različitih pristupa u debati proističu iz različitih istraživačkih pitanja, tačnije, pokazujem da neki od pristupa proističu iz pitanja koje se tiče uslova sticanja mentalnih predstava a drugi iz pitanja koje se tiče uslova za razmišljanje o mentalnim predstavama. Najzad, pokazujem da je teza o eksternalizmu daleko zanimljivija ako je razumemo kao tezu o tome na koji način mentalno predstavljanje funkcioniše u našem svetu a ne kao tezu o tome na koji način mentalno predstavljanje funkcioniše u svim metafizički mogućim svetovima. Tačnije, pokazujem da „nomološko“ razumevanje teze može da objasni kako i zašto eksperimentalne metode koje se koriste u savremenim kognitivnim naukama mogu da pruže uvid u obrazovanje ponašanja.

Ključne reči: eksternalizam, semantički eksternalizam, eksternalizam o posrednicima, supervencija, empirijske metode

UDK: 165.194
https://doi.org/10.2298/FID1804489V
Original Scientific Article
Received: 10.09.2018. Accepted: 14.11.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Karina Vold

OVERCOMING DEADLOCK: SCIENTIFIC AND ETHICAL REASONS TO EMBRACE THE EXTENDED MIND THESIS

ABSTRACT
The extended mind thesis maintains that while minds may be centrally located in one's brain-and-body, they are sometimes partly constituted by tools in our environment. Critics argue that we have no reason to move from the claim that cognition is *embedded* in the environment to the stronger claim that cognition can be *constituted* by the environment. I will argue that there are normative reasons, both scientific and ethical, for preferring the extended account of the mind to the rival embedded account.

KEYWORDS
extended cognition, intracranialism, embedded mind, ethics, theory selection, cognitive rehabilitation

1. Introduction

Andy Clark and David Chalmers's extended mind thesis maintains that while minds are centrally located in one's brain-and-body they are sometimes partly constituted by tools in our environment. Some critics argue that we have no reason to move from the claim that cognition is *embedded* in the environment to the stronger claim that cognition can actually be *constituted* by the environment. In this paper I argue there are normative reasons, both scientific and ethical, for preferring the extended view to the embedded view. In the first place I appeal to the scientific values of simplicity, usefulness, and explanatory power, to argue that our best scientific theory of the mind will include extended mental states. In the second place, reviewing the literature, I appeal to three ethical reasons to prefer the extended view: it better protects against harm to the mind Levy (2007a,b); it better accounts for compensatory rehabilitation as a way of repairing the mind (Drayson and Clark, forthcoming); and it offers a better assessment of the capacities of learning disabled individuals (King 2016). All of these reasons, especially taken together, I argue provide some support for preferring the extended mind thesis to the rival embedded mind thesis.

2. Clark and Chalmers's Extended Mind Thesis

The dominant view in the brain sciences is that the brain alone constitutes, or realizes, the mind. But several philosophers have argued that the mind is sometimes

Karina Vold: Research Associate Ph.D, Leverhulme Centre for the Future of Intelligence, Faculty of Philosophy, University of Cambridge; kvv22@cam.ac.uk.
This work was supported by the Leverhulme Centre for the Future of Intelligence, Leverhulme Trust, under Grant RC-2015-067.

partially constituted by more than just the brain. While this idea has long roots in philosophy (Martin Heidegger, Maurice Merleau-Ponty, and John Dewey have all defended versions of this claim, for example), various contemporary philosophers have given new life to the idea, often under the title of *cognitive extension* or the *extended mind thesis*. In this paper I will focus on one particular version of this claim: the "extended mind" thesis as defended by Andy Clark and David Chalmers (1998). Clark explains, "[p]roponents of the extended mind story hold that even quite familiar human mental states (e.g., states of believing that so and so) can be realized, in part, by structures and processes located outside the human head." (Clark 2008: 76) In other words, the extended mind thesis maintains that mental states (and processes) sometimes *extend* beyond the brain in the sense that they are partially constituted by extra-bodily states (or processes) working together with brain states (or processes).

To defend the extended mind thesis Clark and Chalmers argue,

> [i]f, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. (Clark and Chalmers 1998: 8)

This is now often referred to as their 'parity principle', while others call it the 'fair treatment principle' (e.g. Sprevak 2009; Drayson 2010) as it maintains that we should regard equivalent processes in similar ways, irrespective of whether they are internal or external to the skull. Motivated by this principle, their parity argument can be summarized:

(P1) A physical state (or content-bearing structure) *p* is constitutive of a mental state of type *m* when *p* plays the causal role characteristic of *m* in the system.

(P2) A physical state (or content-bearing structure) *p* located beyond (or partially beyond) an agent's biological body can play the same causal role as physical states of the biological body that surely constitute an ordinary mental state of type *m*.

(C) Therefore, physical states (or content-bearing structures) located beyond (or partially beyond) the biological body can be constitutive of an agent's mental state.

The first premise expresses a commitment to a coarse-grained common-sense role functionalism.[1] To support the second premise, Clark and Chalmers describe a case in which, they argue, an object in the environment does play exactly the same role for one agent that neurons in the brain (something we would surely count as part of the supervenience or realization base of the mind) do for another. The case involves two people: Inga and Otto. Inga decides to go to an exhibition at the museum and to do so, "[s]he thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum." (Clark and Chalmers 1998: 12) Meanwhile, we imagine that Otto suffers from Alzheimer's disease and has to rely on information he stores in a notebook to help structure his life. When he decides to go to the same exhibition he consults his notebook, where he

1 For the sake of brevity I will bypass an explanation of what this view maintains exactly. It is a popular, though not uncontroversial, view in contemporary analytic philosophy of mind.

has written the address and directions for how to get there. He then walks to the museum and heads inside (Clark and Chalmers 1998: 12–13). Clark and Chalmers argue that in the "relevant respects" the information in Otto's notebook "functions just like" the information in Inga's brain that constitutes an ordinary belief and thus both should count equally as part of the constitutive machinery of his mind (Clark and Chalmers 1998: 13). In other words, the information stored in the notebook *really is* a part of Otto's mind, just like the information stored in Inga's brain *really is* a part of her mind. More specifically, the information in the notebook is meant to be an example of an extended standing (i.e. not currently being entertained), non-conscious belief.

3. The Coupling-Constitution Objection

Importantly, the claim made by extended mind theorists is a *constitutive* one – that mental states and processes can be partially *constituted* by objects located beyond the brain and body. One of the major objections confronting the extended mind thesis maintains that we have no reason to move from the claim that cognition is *causally reliant* on the environment to the claim that cognition is partially *constituted* by the environment. Frederick Adams and Kenneth Aizawa argue that Clark and Chalmers commit a coupling-constitution fallacy by mistaking the mere causal dependence, or *coupling*, of extra-neural resources with neural activity for their *constitutive* involvement in unconscious mental states (or processes), such as Otto's belief about the location of the museum.[2] Adams and Aizawa maintain that the *mere* coupling of a resource to a system does not imply the resource is partially constitutive of that system. The circulatory system is coupled to the cognitive system in the sense that circulation *causally supports* cognition in a crucial way, but this does not imply that circulation is partially constitutive of cognition (Adams and Aizawa, 2008: 10–11). Thus, to say that an object, x, is coupled to another, y, does not imply that x constitutes (or is a part of y). So to show that external objects are coupled to our cognitive states or processes does not imply that these partially constitute our cognition.

3.1 The Embedded Mind Thesis

The distinction between the mere coupling of inner and outer resources and the constitutive involvement of outer resources is what distinguishes the *embedded* theory of cognition from the stronger *extended* mind theory. The embedded account maintains that a cognitive system depends, sometimes *crucially*, on the complexity of its environment, but that the environment is not an actual part of the mind. Herbert Simon, for example, argues that much of the apparent complexity of cognitive systems is actually external to the agent, residing in the environment. On this view cognitive systems lean heavily on this worldly complexity without internalizing it (Simon 1969: 51–52). For example humans sometimes structure their own environment to store information and then rely on these external structures instead of

2 See Adams and Aizawa (2001), (2008).

relying on internal resources. The *mise en place* method of lining up one's ingredients in the correct order for cooking, for instance, is widely used by chefs to save them from having to remember the ordering of their recipes while cooking (Clark 2008). The embedded view of cognition tells us that in order to understand and explain cognitive processes, such as the chef's use and processing of information while cooking, cognitive science cannot just study the internal processes of computation instantiated in the brain. Instead, we must study the way that structures in the local environment of an agent facilitate the success of the agent's internal processes. Thus, the embedded view offers an explanatory, or epistemic, reason to look beyond the brain. But the embedded view does not make any substantial constitutive claim; it does not challenge the (metaphysical) view that the brain wholly constitutes the mind (see Rowlands 2010, Chapter three for further discussion). Let us call this position, i.e. the view that the brain wholly constitutes the mind, 'intracranialism'. The key difference then is that embedded mind theorists accept intracranialism, while extended mind theorists reject it.

Extended mind theorists tend to think that embedded claims risk triviality. Almost everyone agrees that the mind is in some sense causally reliant on the body and the extra-bodily world. Many even agree that the body and tools in our environment can work as scaffolds, causally contributing to the development of certain higher-level cognitive capacities or the execution of cognitive tasks, such that these capacities or tasks would not have developed or could not be executed (or may not even confront the agent in the first place) were it not for these extra-cranial contributions. This view about 'cognitive scaffolding' was advocated by the developmental psychologist Lev Vygotsky (1978). But scaffolding falls short of the constitutive claim that the extended mind thesis makes (for more on this issue see Adams and Aizawa 2001, 2008; Shapiro 2008; Clark 2008, 2010).

Those who prefer the embedded mind theory, on the other hand, tend to think that there is no good reason for preferring the stronger constitutive claim made by the extended mind thesis to the more conservative coupling claim. In response to Adam and Aizawa's coupling-constitution objection, Clark (2010) argues that the burden of proof lies with those who reject the constitutive claim. They need a principled reason for maintaining that all mental states are entirely constituted by neural resources and only causally supported by extra-neural ones. To this end that Adams and Aizawa (2005) argue that *original*, or non-derived, content is the distinguishing 'mark of the cognitive. *Original*, or *non-derived*, content is meant to contrast with the *derived* content that non-mental objects can display. So, while non-mental objects, such as the words in a book, e.g. Otto's notebook, may carry *derived* content, original content is uniquely a feature of minds.[3] Thus, Adams and Aizawa argue that the notebook fails to be partially constitutive of Otto's mind because it lacks this 'mark of the mental' that — just as a matter of current contingent fact, on their view — only brains can realize.[4]

3 This distinction seems to be inspired by a distinction between derived and original intentionality made by Searle (1992).

4 I will not respond to Adams and Aizawa's objection here, thought I believe there are several responses one could give (for some responses see Clark 2008b, 2010). Notice that even if the objection works against the example of Otto and his notebook, their distinction

But some who reject the extended mind theory have also rejected Clark's insistence that the burden of proof lies with them. What reasons do we have for preferring the more radical idea that technology, such as pens and papers and even smart phones, can really be partly constitutive of our mental life? Rupert (2004), for example, argues that there is no obvious reason for preferring the extended account. He argues, first, that adopting the embedded account is enough to recognize the indispensability of studying an agent's environment for understanding his cognition without conceding that the environment is actually partially constitutive of cognition. Furthermore, Rupert argues that we can explain all of the relevant phenomena that cognitive scientists study with the embedded account and, thus, moving to the stronger extended account is unjustified and unnecessary (Rupert 2004: 8–9). If the extended account does not offer us anything more than the embedded account, then following the methodological principle of conservatism, Rupert argues, we should endorse the embedded view over the extended mind view (Rupert 2004: 9). In this case, those defending the more radical extended view need to justify their proposed revision of our common-sense views about where the mind is.[5] Since this debate took off, several normative reasons for preferring the extended account over the rival embedded account have been suggested. I will now survey and evaluate these reasons as well as offer my own.

4. Scientific Virtues in Theory Selection

Thomas Kuhn (1977: 321–323) identifies several virtues or preferable characteristics that provide the basis for choice between competing scientific theories. These include accuracy, consistency, scope, simplicity, and fruitfulness. Kuhn thinks that accuracy is the most important of these virtues, but we've seen that there is currently a stalemate in the debate between which view—embedded or extended—is accurate. Thus, I will make the case that by appealing to other scientific values, including simplicity, usefulness and explanatory power, we are compelled to prefer the extended account to the embedded account. While not identified by Kuhn, usefulness and explanatory power are arguably accepted scientific virtues (especially the latter), and should be distinguished from Kuhn's other virtues, such as simplicity and fruitfulness, as they will sometimes be traded off against them.

4.1 Simplicity in Theory Selection

Simplicity is widely accepted as a norm of theory formulation in a wide range of disciplines, both humanistic and scientific. Occam's razor, for example, has long

would not block all possible cases of extension. We can imagine that instead of using a notebook Otto uses the mind of another agent, e.g. his long-time partner, to store the information that forms his beliefs. In this case the external resource is itself a brain, so it would be capable of original content. The result would be an instance of a *socially extended mind* — where one agent's mind has extended into another's brain ('social' because two agents are involved).

5 Especially since Clark and Chalmers' argument appeals to a version of 'common-sense' functionalism.

been employed by philosophers as a way of guiding our preferences when choosing between two competing hypotheses. This principle maintains that, in the case where all other things are equal, for example, where both hypotheses can account equally as well for the data, we ought to opt for the simplest hypothesis, that is, the one that posits the fewest metaphysical entities. With respect to competing theories about where the supervenience base of the mind is, Clark and Chalmers (1998) suggest that we can use simplicity as a way of assessing and arbitrating between the extended view and the embedded view.

One reason folk psychology has endured is, quite plausibly, its simplicity (it of course also has other virtues, including explanatory power, which I discuss below). The belief-desire-intention (BDI) model of our practical reasoning, developed by Michael Bratman (1987), is a way of explaining how we perform everyday actions. It is also an example of the method we commonly use in explaining the actions of others in everyday life, which relies, fundamentally, on folk-psychological concepts. Clark and Chalmers use this model to argue that the extended account yields the simplest explanation of Otto's action:

> Certainly insofar as beliefs and desires are characterized by their explanatory roles, Otto's and Inga's cases seem to be on a par: the essential causal dynamics of the two cases mirror each other precisely. We are happy to explain Inga's action in terms of her occurrent desire to go to the museum and her standing belief that the museum is on 53rd street, and we should be happy to explain Otto's action in the same way. (Clark and Chalmers 1998: 13)

In fact, according to the BDI model, we would explain Inga's action in terms of her occurrent desire to go to the museum, her standing belief about where the museum is located *and* her intention to take action. The extended mind thesis allows for the simplest application of the BDI model to Otto: we would explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief about where the museum is located—which happens to be stored in his notebook, instead of in his brain—and his intention to take action. Thus, the extended view of the mind allows for the simplest application of the BDI model, a method of folk psychology that we regularly use. Consider the alternative explanation, as Clark and Chalmers describe it:

> The alternative is to explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief that the Museum is at the location written in the notebook, and the accessible fact that the notebook says the Museum is on 53rd Street; but this complicates the explanation unnecessarily. If we must resort to explaining Otto's action this way, then we must also do so for the countless other actions in which his notebook is involved; in each of the explanations, there will be an extra term involving the notebook. We submit that to explain things this way is to take one step too many. It is pointlessly complex, in the same way that it would be pointlessly complex to explain Inga's actions in terms of beliefs about her memory. ...In an explanation, simplicity is power. (Clark and Chalmers 1998: 13–14)

We do not explain Inga's action in terms of her belief about her own memory, e.g. that she has stored information about the location of the museum in her brain, which she then accesses in order to take action. This would be "pointlessly complex". In normal cases, such as Inga's, we use our own memory transparently:

we do not have to form beliefs about what is stored in our memories. But this is also how Clark and Chalmers describe Otto: he is so accustomed to relying on his notebook that he uses it transparently. He does not have to reflect on where the relevant information is stored; he simply reaches for the notebook. And, all things being equal, we should prefer the explanation that is simplest. Thus, because the extended mind thesis allows us to give a simpler folk-psychological explanation of Otto's actions it is preferable to the embedded account, which requires us take additional explanatory steps (further discussion in Drayson and Clark, forthcoming).

4.2 Usefulness in Theory Selection

One reason to prefer one theory to a rival is its usefulness. As part of a larger discussion on addiction and responsibility, Levy argues that the extended mind thesis is more *useful* than the rival embedded account insofar as it enables us to better control ourselves (Levy 2007a: 220). He argues that research on ego-depletion suggests that addicts have depleted self-control and thus they experience more difficulty in resisting their cravings than one who craves but is not addicted. Levy suggests that "[i]t may be literally impossible for the addict to refrain from taking their drug… when it is immediately available and their self-control resources are depleted." (Levy 2007a: 219). Nevertheless, he maintains that there are some things that addicts can do both in the short and the long-term to overcome their addiction—namely, they can take steps to control their environment.

According to Levy, the traditional view that the brain wholly constitutes the mind (what we've called intracranialism) works against those suffering from addiction. This view promotes the idea that the only way for one to recover is to change their mind—that is, their brain. In other words, addiction is entirely a matter of "will-power" and the addict needs to just "say no" to their cravings (Levy 2007a: 219–220). Levy argues "[t]o the extent to which we promote the view that giving up a drug, whether it is tobacco or heroin, is all a matter of "will-power," we direct them away from the kinds of environmental modifications they need to make if they are to regain control." (Levy 2007a: 220) Thus, Levy reasons that the extended mind view is more useful in so far as it enables us to better control our own behavior, as well as the behavior of others. We should prefer the theory of the mind that yields the most successful strategies with respect to repairing our minds, Levy argues, and thus the "real-world success" of the extended mind thesis is evidence of its truth. He explains, "[k]nowledge is power: if the [extended mind] hypothesis were false, then it would not yield successful strategies." (ibid.) The stronger constitutive claim that the extended mind makes insists that the just "say no" view about addiction must be rejected.

I think the usefulness of the extended view over the embedded view in this context might be debated. Both accounts point us to the agent's wider environment, beyond the brain, in order to fully *explain* the agent's decisions and reasoning. So perhaps the constitutive claim is not necessary. Thus, this reason on its own may not be sufficient for preferring the extended view to the embedded view—it is also not obvious that Levy intends it to be. But if we accept Levy's point, then the usefulness of the extended account gives one reason (even if not an indefeasible one) to prefer it to the competing theory.

4.3 Explanatory Power in Theory Selection

I argue that the embedded view suffers from explanatory impotence, while the extended mind thesis has explanatory power and that this gives us reason to prefer the latter to the former. Explanatory power refers to the ability of a theory to effectively explain phenomena that pertain to its subject matter. Consider an example from Drayson and Clark (forthcoming) of a sub-population of inner-city Alzheimer's sufferers in St. Louis who scored dismally on standard tests, such as the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) protocol. Based on their test scores, the patients should have been living in full-care hospitals. Yet, they were able to cope with the demands of daily life and to successfully live alone in the city. The patients puzzled doctors and Alzheimer specialists. After visiting their homes, however, it was revealed that they had transformed their living environments with many personalized cognitive tools, props, and aids: from message centers, open notes about what to do and when, to labels and pictures on the walls, including labelled photos of family and close friends, and 'memory books' that recorded new events, meetings, and plans. Some had open storage spaces that kept crucial items, e.g. kitchen tools or chequebooks, visible, rather than requiring memory of where these things were kept.

The problem is that the standard tests for Alzheimer's disease rest on the assumption of intracranialism: they only evaluate one's internal memory. The extended mind thesis is able to explain how these patients continued to effectively function in the world. While the (likely implicit) assumption of intracranialism left experts confused about how to explain this phenomenon.

What is more, if we took the tests of internal memory as the only standard, these patients would have likely been forcefully removed from their homes and re-located to controlled hospital settings much sooner than might be necessary. Drayson and Clark point out that the re-location of Alzheimer's patients is often a fateful turning point in which their conditions become more severe. On the extended view this is explained by the fact that this kind of re-location is on par with the infliction of new brain damage upon the patients: in one fell swoop it removes them from the cognitive tools, props, and aids which were supporting their intracranial cognitive capacities. The embedded view, on the other hand, struggles to explain why the change in environment so often leads to a dramatic degradation in the condition of Alzheimer's patients.

In this section I have pointed to three preferable characteristics of the extended mind view—simplicity, usefulness, and explanatory power—that together provide us a basis for preferring it to the competing embedded view of the mind. In the next section I consider several ethical reasons for preferring the extended account.

5. Ethical Reasons for Preferring the Extended Mind Thesis

In this section I discuss three ethical reasons that have been put forth for preferring the extended account. First, that the extended mind thesis better protects against harm to the mind. Second, that it better accounts for compensatory rehabilitation

as a way of repairing the mind. And third, that it better addresses concerns about the way we assess the capacities of learning disabled individuals.

5.1 Protection of the Mind

Adopting the extended mind thesis can help us better protect our minds from harm. Levy (2007a, b), for example, argues that insofar as we view the tools and technologies we use as a part of the mind, we are more likely to protect those tools from being harmed or removed, in just the way we protect the biological agent, especially the brain, from harm. Levy advances the 'Ethical Parity Principle' (EPP) to capture this idea. His principle is meant to complement Clark and Chalmers's (1998) original parity principle. In its strongest form, EPP is stated as follows:

> EPP (strong): Since the mind extends into the external environment alterations of external props used for thinking are (*ceteris paribus*) ethically on par with alterations of the brain. (Levy 2007a: 61)

According to this principle, any alteration or *harm* to external cognitive tools should be treated as ethically equivalent to alterations of internal cognitive tools—namely, the brain. Stealing Otto's notebook would be morally wrong insofar as it would be theft. But, according to EPP we should view this action as much worse than that, ethically speaking, precisely because the notebook plays an important role in Otto's cognitive life. The notebook stores important information that Otto regularly relies on: it contains his memories, his beliefs, and so on. Thus, the notebook has a cognitive status, and moral status, equivalent to Otto's brain. For this reason, stealing Otto's notebook would be more on par with kicking Otto in the head, causing him serious cognitive damage, than it would be to stealing some property, like his gym bag. According to EPP, we should view theft of a cognitive tool much more seriously than theft of property more generally.

Blitz (2010) argues for a *legal* parity principle, which maintains that when external processes are functionally equivalent to internal cognitive ones we should also treat them as legally equivalent. On his view, the law should protect our cognitive tools from harmful alterations in just the way that it has protected our brains against harm. Just as we have helmet laws that mandate protective devices for our brains, for example, we should also have protections against kinds of external cognitive damage.

As before, we might take the "real-world success" of the extended mind thesis to better protect our minds from harm as evidence of its truth. However, although Levy (2007a) endorses the extended mind thesis, he concedes that the defender of the embedded mind is able to offer equal protections of the mind. The embedded mind theorist might reject the strong version of EPP, but she could account for ethical parity with a weaker principle:

> EPP (weak): Alteration of external props are (*ceteris paribus*) ethically on par with alterations of the brain, to the precise extent to which our *reasons* for finding alterations of the brain problematic are transferable to alterations of the environment in which it is embedded. (Levy 2007a: 61)

The weaker version of the principle maintains that alterations to external cognitive tools are ethically on par with alterations to the brain only insofar as the reasons we have for objecting to the latter are applicable to the former. According to Levy, weak EPP would allow the defender of the embedded account to offer equal protections to external cognitive tools without conceding the stronger claim that external props are actually partly *constitutive* the mind. Weak EPP still requires that we treat interventions and alterations to internal (brain) and external operations on a par, unless we can find ethically relevant reasons for drawing a distinction between them. Thus, according to Levy, appealing to protecting the mind does not itself establish sufficient normative reasons for preferring the extended mind to the embedded mind, precisely because both frameworks offer equal protections. Simply recognizing the importance of external resources for cognition, as the embedded theory does, is sufficient to establish the same protections (King 2016: 47–50).

5.2 Cognitive Rehabilitation: Repairing the Mind

Drayson and Clark try to respond to Levy's claim that the embedded mind would offer the same protections and would have the same ethical implications more generally as the extended mind thesis. They argue that the extended mind thesis has more significant ethical implications than the embedded view, focusing on how these views offer differing accounts of the cognitive rehabilitation of neuroatypical individuals and of our understanding of cognitive impairment.

Cognitive rehabilitation refers to the process of improving an individual's impaired ability to process and use information. Drayson and Clark explain that there are two dominant strategies of cognitive rehabilitation: *restorative* and *compensatory*. *Restorative* strategies aim at restoring the damaged neural area or circuits. *Compensatory* strategies, on the other hand, aim to achieve the same functional results as restorative strategies but in different ways, for example, by using photos or labels to assist memory. This typically involves adaptive strategies that rely on both internal and external resources to improve information processing and use. Some neuroscientists favor restorative strategies and view compensatory strategies only as a recourse, necessary because of our limited understanding of and access to the brain. On this view, as neuroscience advances, compensatory strategies will eventually be replaced in favor of restorative ones. Thus, while compensatory rehabilitation can be an effective substitute for neural restoration, it does not truly restore the mind—only repairing the damaged or affected neural areas can *restore* the mind. This follows from what I will call the *principle of intracranialism*, which holds that neural activity entirely determines mental activity (from King 2016: 55.) Given its commitment to the position of intracranialism, the embedded account is committed to this principle, while the extended account rejects it. And it follows from the principle of intracranialism that only by restoring neural activity can we restore mental activity.

As a result, Drayson and Clark argue that the extended mind thesis and the embedded mind thesis take different positions on rehabilitative strategies. Due to her intracranialist commitment, the embedded mind theorist must maintain that compensatory strategies cannot achieve true restoration of mental function. So she must maintain a distinction between restorative and compensatory rehabilitation: if

the mind is realized only by neurons, then the only way to repair a damaged mind is to repair the neural areas that bring it about. The extended mind thesis, on the other hand, offers a different picture. It allows us to view compensatory strategies as on par with restorative ones. For the extended mind theorist, both strategies are legitimate ways of repairing cognitive ability—neither is a second-best option.

The oddity of what the embedded mind view is committed to in maintaining the traditional distinction between restoration and compensation is brought out by an example that Clark gives (in response to Jerry Fodor's (2009) critique of the extended mind thesis):

> [I]magine a case in which a person (call her Diva) suffers minor brain damage and loses the ability to perform a simple task of arithmetic division using only her neural resources. An external silicon circuit is added that restores the previous functionality. Diva can now divide just as before, only some small part of the work is distributed across the brain and the silicon circuit: a genuinely mental process (division) is supported by a hybrid biotechnological system... (Clark 2009)

In this case Diva's damaged neural circuit has been restored with silicon-based functional replacements of neurons. But, if one maintains the traditional distinction between restoration and compensation, not even this would count as true restoration. If neurons alone can constitute cognition, then even this rehabilitative strategy falls short of true restoration. Even though Diva's doctors have addressed the structural integrity of her neural circuits they have had to rely on silicon structures, rather than biological structures, to do so and thus this has to be seen as merely compensatory. So, while the extended mind theorist can describe Diva's rehabilitative strategy as truly *restoring* her cognitive functioning, the embedded mind theorist cannot.

As another example, consider again the Alzheimer's patients discussed before. These patients developed their own compensatory strategies that allowed them to live successful lives despite their impaired intracranial functioning. The extended mind thesis allows us to see the compensatory strategies developed by these patients as genuinely restoring their cognitive functioning. It offers a new way of conceptualizing the distinction between restoration and compensation, suggesting that we should evaluate rehabilitation based on the functional capacities of the extended cognitive system, rather than restricting cases of "true" rehabilitation to only those that involve restored neural circuits.

Drayson and Clark's argument also offers a response to Levy's view that the embedded mind thesis has the same ethical implications and affords the same protections of the mind as the extended mind thesis. Drayson and Clark argue that if we view the cognitive tools and aids of the cognitively impaired as mere scaffolding, rather than as legitimately a part of their minds, we are *less likely* to protect those tools from harm in the same way that we protect the biological brain from harm. On the extended mind view, the minds of people who rely on external tools are more vulnerable to harm than those who do not rely as heavily on them precisely because the dominant intracranialist view of the mind has failed to recognize and protect those tools. Thus, resisting the extended mind thesis may lead us to neglect what ought to be protected and, as a result, place cognitively impaired people in greater danger of cognitive harm (further discussion in King 2016: 46–48).

5.3 Capacities of Learning Disabled Individuals: Improving the Mind

Finally, King (2016) offers yet another reason to favor the extended account: that adopting the embedded account commits one to problematic views about the cognitive capabilities of learning disabled individuals, while the extended account avoids these commitments. Researchers working on learning disabilities draw a distinction between two kinds of strategies aimed at addressing learning that, according to King, roughly maps on to the distinction between restorative and compensatory strategies discussed in the previous section. Researchers of learning disabilities distinguish between "remedial" strategies and compensatory strategies (e.g., Garner and Campbell 1987). Remedial strategies aim to directly address a learning-disabled individual's impairment by improving their ability to perform tasks in just the same way that a non-disabled individual would (King 2016: 54). Compensatory strategies, on the other hand, attempt to circumvent learning impairment by helping the individual perform the same tasks by using assistive technologies (ibid). King argues that remedial strategies are analogous to restorative strategies in cognitive rehabilitation and (again, because of the commitment to intracranialism) the embedded mind theorist has to say that remedial strategies are the only "true" way to enhance a learning-disabled person's cognitive capabilities. Compensatory strategies, as before, are only a second-best option, employed when remedial strategies are not possible: they might help an individual *compensate* for her impairment but they do not *restore* cognitive capacities. The extended mind theorist, on the other hand, can view both strategies as genuinely restoring cognitive capabilities.

To bring out this difference, King describes a "paradigm case" of a learning disabled person who uses assistive technologies as a compensatory strategy for her disability:

> Consider someone with a learning disability, Dana, who requires a graphic organizer of potential decisions in order to evaluate which decision is best. In this example, let us imagine that Dana has a very difficult time comparing the relevant factors when she must evaluate them solely "in her head," but when allowed to create and utilize a visual diagram of the various possibilities, her decision making skills are just as good as anyone's. In this case, she needs a particular physical configuration of information in order to be able to perform a cognitive process like comparing and choosing among potential courses of action, and without the aid of these external resources, it would appear that she is incapable of performing this cognitive action. (King 2016: 49)

According to King, environmental tools, such as graphic organizers (i.e. visuospatial ways of representing information such as cognitive maps, venn diagrams, flowcharts) are typically employed as assistive technologies in compensatory strategies for learning, problem solving, and planning, for example. The embedded mind thesis would say that Dana couldn't make complex decisions. King argues that we should resist this conclusion. Dana may need to rely on graphic organizers, but she is quite capable of making complex decisions. She just requires a particular sort of environmental scaffolding that a non-learning disabled individual would not. In other words, Dana is only unable to make complex decisions when she is denied access to the external cognitive aids that she requires (King 2016: 50). Thus, the extended account better accommodates our intuitions about learning-disabled individuals than the embedded account.

5.3.1 Response from the Embedded Mind Theorist

Is this an accurate portrayal of the embedded mind theory? One might object that the embedded mind thesis can allow that compensatory strategies genuinely repair and improve the mind. After all, the embedded mind theorist holds that the environment plays a crucial and sometimes indispensable role in supporting cognition. Thus, just because the embedded mind theorist denies the constitutive claim that the extended mind theorist endorses does not mean that she would also resist compensatory strategies that make use of assistive technologies (King 2016: 54–55). But given the commitment to the principle of intracranialism, the embedded mind theorist is committed to an inverse relation between the extent to which an individual relies on external tools and the extent to which we ought to say that she, or her mind, is really doing x, where x is some cognitive process (ibid.). This means that Dana is only "doing" as much, cognitively speaking, as her neurons are doing (King 2016: 56). And therefore, she only merits "cognitive credit" for what her neurons do: the cognitive work that is done by whatever assistive devices she employs is not being done by her and so she should not get credit for the achievements of these devices. King points out that this means that the more heavily integrated assistive devices are—the more one relies on these tools—the less cognizing one is actually doing. The embedded mind theorist is, thus, forced to say that Dana has less cognitive capacities and deserves less cognitive credit than someone who could perform the same task 'intracranially'. So, the embedded mind thesis really does commit one to saying that learning-disabled individuals, such as Dana, who rely on assistive technologies are cognitively capable of less than non-disabled learners.

On the extended view we need not draw equivalence between neural capacity and cognitive capacity, and, thus, diminished neural capacity does not entail diminished cognitive capacity. This means that compensatory strategies are not just acceptable (as the embedded mind theorist would also accept them), but they are equally legitimate ways of improving and repairing cognitive capacities as restorative strategies are (ibid.). King argues that the extended view better captures our intuitions about learning disabled individuals: "[a]ssistive technologies are tools that help [learning-disabled] individuals do more, not less" (King 2016: 57). But furthermore, citing various testimonial reports, she also argues that the extended view better captures how learning-disabled individuals view themselves and their own relationship to the assistive technologies that they use:

> [Learning-disabled] individuals who have consistent access to assistive technologies in the classroom report feeling less anxious, more independent, and more confident in their own abilities (Day and Edwards 1996). Rather than making them feel as if the more they use technologies, the less they can do "themselves" (the picture of cognitive capability that the embedded thesis predicts), well-assisted [learning-disabled] individuals report feeling as if they are capable of doing more, and with an increased sense of independence and self- reliance. Testimonial reports of the impact of access to assistive technology on [learning-disabled] individuals' self-concept further supports the suggestion that [learning-disabled] individuals see themselves in this way. (King 2016: 57).

While this is hardly an indefeasible reason to favor the extended view, the first-person accounts from learning-disabled individuals do seem to align with our

own intuitions that assistive technologies help them to become better at processing information and using information.

The embedded mind defender might also maintain that Dana's graphic organizer helps her arrive at better decisions than had she not used the device. But, there is a key difference between what, precisely, the two accounts mean when they say that the technology allows Dana to "do more": the embedded view would say that the graphic organizer enables Dana to achieve more qua competent tool user, while the extended view asserts that the graphic organizer allows Dana to do more qua cognizer (ibid). King uses the analogy of a runner who would fatigue after just a few miles when running barefoot but when given a pair of quality running shoes (an assistive device) could run for significantly longer distances and at greater speeds. The shoes could be seen as merely an assistive technology—a "tool" for running—but what we want to say, and what we are more likely to say, is that their use improved the runner's abilities qua runner (King 2016: 60). Likewise, the extended thesis views the cognitive agent and the tool as a single (wide or extended) system: it draws no divide between the agent and the tool, while the embedded thesis would have to attribute any increased capabilities that result from the tool-use to the tool itself rather than the individual. Because the embedded theorist clings to intracranialism, she must insist on a clear distinction between the cognitive capacities of the agent and the non-cognitive capacities of the tool (King 2016: 59–60). King's argument suggests that while the extended mind thesis might seem radical and less intuitive, on a closer analysis of how the view explains rehabilitation strategies employed by learning-disabled individuals it seems it is the embedded thesis that struggles to accommodate our intuitions about the use of assistive technologies. Furthermore, the embedded view must reject the beliefs that learning-disabled individuals self-report about their own relationships to their assistive devices.

I conclude that King's argument offers another normative reason to prefer the extended account. The ethical reasons discussed in this section should also serve as caution that our biological capacities should not condition our notions of rehabilitation, well-being, or decision-making, at the risk of alienating, marginalizing, and even harming those who rely on technologies for their cognition. The picture that I hope emerges from this paper is that technology can, and often does, allow our minds to transcend our biological capacities.

6. Conclusion

The debate between the embedded and the extended mind views has seemed to reach a stalemate. Is the mind merely embedded in the world, coupled to the various tools that it uses, or is it partly constituted by these tools? And what hinges on this? Insofar as one agrees that we have reached a stalemate, I have argued that there are normative reasons that we can appeal to that make that extended mind thesis preferable to the embedded mind thesis. First, in terms of choosing between these views as two scientific theories of the mind, we can appeal to various norms of theory selection. I have argued that the extended mind thesis has three virtues—simplicity, usefulness, and explanatory power—that make it preferable to the embedded mind view. Second, looking at the ethical import of these views, I argue that

there are again good reasons to prefer the extended mind thesis: it better protects against harm to the mind, it better accounts for how we repair the mind, and finally, it offers a better assessment of the capacities of learning disabled individuals. While some of these are not on their own decisive reasons to prefer one view to the other, I believe that taken together, they provide good reason for preferring the extended view. Thus, taken as a whole these reasons can be used as a response to the coupling-constitution deadlock. We should, I argue, view external tools as enabling us to transcend our biological capacities: our minds are extended, not just embedded.

References:

Adams, F. and K. Aizawa. (2001), "The Bounds of Cognition." *Philosophical Psychology* 14(1): 43–64.

Adams, F. and K. Aizawa. (2008), *The Bounds of Cognition*. Blackwell Publishing.

Blitz, M. J. (2010), "Freedom of Thoughts for the Extended Mind: Cognitive Enhancement and the Constitution", *Wisconsin Law Review* 4: 1049–1117.

Bratman, M. E. (1987), *Intentions, Plans, and Practical Reasons.* Center for the Study of Language and Information Publications.

Clark, A., and D. Chalmers. (1998), "The Extended Mind." *Analysis* 58(1): 7–19.

Clark, A. (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* New York: Oxford University Press.

Clark, A. (2009), "Response to Jerry Fodor's Review 'Where Is My Mind?'", Review of Supersizing the Mind. *London Review of Books*: Letters 31(6).

Clark, A. (2010). "Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa". In: R. Menary (ed.), *The Extended Mind.* Cambridge: MIT Press, pp. 81–100.

Drayson, Z., and A. Clark. (Forthcoming). "Augmentation, Agency, and the Spreading of the Mental State." <https://sites.google.com/site/zdrayson/papers>

Drayson, Z. (2010). "Extended Cognition and the Metaphysics of Mind", *Cognitive Systems Research* 11(4): 367–377.

Garner, J. B., and P. H. Campbell. (1987), "Technology for Persons with Severe Disabilities: Practical and Ethical Considerations", *The Journal of Special Education* 21(3): 122–132.

Levy, N. (2007a), "Chapter Six: The 'Self' of Self-Control", *Neuroethics: Challenges for the 21st Century*. Cambridge: Cambridge University Press, pp. 197–221.

Levy, N. (2007b), "Rethinking Neuroethics in the Light of the Extended Mind Thesis." *American Journal of Bioethics* 7(9): 3–11.

Levy, N. (2011), "Neuroethics and the Extended Mind". In: J. Illes et al (eds.), *Oxford Handbook of Neuroethics,* Oxford: Oxford University Press, pp. 285–294.

King, C. (2016), "Learning Disability and the Extended Mind" *Essays in Philosophy* 17(2): 38–68.

Kuhn, T. (1977), *The Essential Tension. Selected Studies in Scientific Tradition and Change*, Chicago: University of Chicago Press.

Rowlands, M. (2010), *The New Science of the Mind.* Cambridge, MA: MIT Press.

Rupert, R. D. (2004), "Challenges to the Hypothesis of Extended Cognition", *Journal of Philosophy* 101(8): 389–428.

Shapiro, L. A. (2008), "Functionalism and the Boundaries of the Mind", *Cognitive Systems Research* 9(5): 5–14.

Simon, H. (1969), *The Sciences of the Artificial,* Cambridge, MA: MIT Press.

Sprevak, M. (2009), "Extended Cognition and Functionalism", *Journal of Philosophy* 106: 503–527.

Vygotsky, L. (1930), *Mind in Society: The Development of Higher Psychological Processes*. M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (eds), Cambridge, MA: Harvard University Press.

Karina Vold

Prevazilaženje zastoja: naučni i etički razlozi za prihvatanje teze o proširenom duhu

Apstrakt

Prema tezi o proširenom duhu iako se naši umovi prevashodno nalaze u našem mozgu i telu, njih ponekad delimično ustanovljuju delovi naše sredine. Kritičari tvrde da nemamo razloga da od teze da je naša kognicija uronjena u našu sredinu napravimo korak ka prihvatanju jače teze prema kojoj je naša kognicija konstituisana našom sredinom. U ovom radu, pokazujem da postoje normativni razlozi, naučni i etički, da prihvatimo tezu o proširenom duhu umesto suparničke, uronjene, teze.

Ključne reči: proširena kognicija, intrakranializam, uzglobljen duh, etika, selekcija teorija, kognitivna rehabilitacija

UDK: 165.412
https://doi.org/10.2298/FID1804505R
Original Scientific Article
Received: 10.09.2018. Accepted: 14.11.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Andrea Roselli

THE MIND BEYOND THE HEAD: TWO ARGUMENTS IN FAVOUR OF EMBEDDED COGNITION

ABSTRACT
In this paper I defend situated approaches of cognition, and the idea that mind, body and external world are inseparable. In the first section, I present some anti–Cartesian approaches of cognition and discuss the intuition they share that there is a constitutive interaction between mind, body and external environment. In the second section, I present the fallacy of the *Cartesian theater of the mind* and explain its theoretical premises. In the third section, I present a spatial argument against it, and argue that some case studies could give support to the idea of the mind stretching over the boundaries of the skull. In the fourth section, I present a temporal argument, and argue that even in this case the idea of an interaction between our cognitive life and the external world has *at least* a very strong intuitive palatability.

KEYWORDS
Extended Cognition, Mind, anti–Cartesianism, Situated Cognition, Embedded Cognition

1. Anti–Cartesian Approaches of Cognition

What is the relation between mind, body and external world? According to situated approaches of cognition[1], they are inseparable; knowledge is the achievement of the whole body in its relation with the environment. Instead of seeing learning as an accumulation of knowledge, this cognitive process is seen in terms of the increase of effective performances in different situations, co–determined by the brain and the environment. A strong Cartesian dualism, between mind and body, is rejected, in favor of an emphasis on the symbiosis of perceptions and actions; acting is in a certain sense thinking, and thinking is in a certain sense acting. This *certain sense* opens a spectrum of different possibilities[2]; in particular, mental processes could be described as 'Embodied' (depending on the entire physical state), 'Embedded' (with the accent put on the interaction between the agent and the world), 'Enacted' (with the accent put on the actions of the agent and their relevance for cognition),

1 See e.g. Greeno (1989) for a presentation of this topic, or Suchman 1987 for a research in cognitive science – where cognition is taken to be continuous with processes in the environment.
2 See Rowlands (2010): 51.

Andrea Roselli: Leverhulme Trust Post Doctoral Research Fellow in Philosophy, Durham University; andrea.roselli@durham.ac.uk.
The research that led to this publication was supported by a research project grant from the Leverhulme Trust (D/Ref. 75647).

and 'Extended' (when external objects act as parts of our minds). In general, we can say that this spectrum of positions is 'Anti–Cartesian', in that it contrasts with the view that there is, in our heads, an independent and pure cognitive 'control center' of the body's behavior, as if an homunculus lived there and used the body as his personal automatic puppet. The debate is complicated by the fact that the aspects listed above overlap (at least partly) in many theories, the fact that even more different positions could be distinguished and, finally, the fact that they are not always thoroughly distinguished by philosophers in this field. A taxonomy of all the different theories present in current literature falls beyond the scope of the present paper; however, it is important to sketch some features of at least the four possibilities listed above to point out the anti–Cartesian element they have in common, since the two original arguments I will present in this paper are related to it.

Embodied cognition refers to the idea that an agent's cognition is shaped and determined by aspects of the entire body (e.g. hormonal states influence higher cognitive processes), that go beyond the brain itself. Cognition, then, depends upon having a body with a motor system and a perceptual system, and upon the fact that these systems interact constructively with the surrounding environment. The goal of successfully surviving in the external environment is thus seen as something which is constitutively built into the cognitive structure of the agent, and that modifies – and, in a sense, *build* – the reality that surrounds him/her[3]. As Varela et alii (1991: 172–173) put it,

> By using the term embodied we mean to highlight two points: first that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological and cultural context.

This approach is then Anti–Cartesian in that it rejects the idea that perception and the motor systems are mere peripheral tools, like input and output devices; mind and body, on the contrary, as seen as acting as a bigger united entity, and the brain is on a par with the body and the external environment in explaining intelligent behaviors. Embedded cognition, while sharing a lot of features with Embodied cognition, refers however much more to the physical interaction between the agent and the external environment, and how this drives the behaviors of the subject and, ultimately, his cognitive life. Enacted cognition puts an accent on the actions and their relevance for cognition; organisms are seen as cognitive systems that have a central role, through interaction with the external world, in the generation of meaning – they *enact* a world[4]. Extended cognition, finally, claims that cognition extends beyond the boundaries of our heads, to include features of the external environment. When we use our fingers to count, our mind extends to our hands and the whole system produces a cognitive act; when we have to remember a number, we can store it in our middle–long term memory or write it down[5] – in the latter case, the piece of paper becomes an extension of our mind; etc. Intuitively

3 N.B.: This passage will be crucial for the two original arguments developed in this paper.
4 Di Paolo et alii (2014): 33.
5 Or, as in McClelland et al. 1986, the use of pen and paper to perform long multiplication.

the idea, then, is that in certain conditions parts of the external world can substitute functions usually performed by (parts of) our brains. In *The Extended Mind* (1998), Andy Clark and David Chalmers famously defended this active externalism[6] – arguing that the external world can participate in the birth of a cognitive act; there are cases in which external objects can perform just the same tasks performed by internal brain structures. As they put it:

> If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process. [...] If we remove the external component the system's behavioural competence will drop, just as it would if we removed part of its brain. [...] The re-arrangement of tiles on the tray, while playing Scrabble, is not part of action; it is part of thought. (Clark & Chalmers 1998: 8–9–10 *passim*)

This view of an extended cognition may also be relevant to scientific investigation, in that the analysis of what have always been considered the result of inner processes may, if this idea is accepted, involve some external objects / events / factors in general. What will become relevant for the two arguments developed in the present paper, however, is another aspect that Clark and Chalmers highlight very well, when they claim that the brain has evolved "in ways which factor in the reliable presence of a manipulable external environment. It certainly seems that evolution has favoured onboard capacities which are especially geared to parasitizing the local environment so as to reduce memory load, and even to transform the nature of the computational problems themselves" (1998: 11). Our visual system, for example, has been shown[7] to rely on the external environment, exploiting contingent facts about the structure of natural scenes.

Many of the problems in the field of philosophy of mind arose from addressing some of the basic assumptions of Cartesianism, such as the dualism between mind and body. In the present paper, I will describe (in the next section) the fallacy of the *Cartesian theater of the mind* and present two arguments against it, one from a spatial point of view and another one from a temporal point of view. Intuitively, what these arguments will show is that there are cases in which the coupled system "brain–environment" is essential to understand how our brain works (in the spatial case) and why do we have a certain phenomenology (in the temporal case). There arguments are meant to show that the Cartesian theater of the mind is indeed a fallacy, and that the idea of the mind that stretches over the boundaries of the skull has *at least* a very strong intuitive palatability. Extended Mind and Extended Cognition are two conceptually closely related hypotheses, and I take it that they can both benefit from the two arguments here presented.

---

6 "Some accept the boundaries of skin and skull, and say that what is outside the body is outside the mind. Others are impressed by arguments suggesting that the meaning of our words 'just ain't in the head', and hold that this externalism about meaning carries over into an externalism about mind. We propose to pursue a third position. We advocate a very different sort of externalism: an active externalism, based on the active role of the environment in driving cognitive processes." (Clark & Chalmers 1998: 7) Their active externalism is opposed to a 'passive externalism' as, for example, in Putnam (1975) or Burge (1979).

7 See e.g. Ullman and Richards 1984.

2. The Cartesian Theatre of the Mind

Daniel Dennett (1991) argued that, in some cases at least, there is no sharp dividing line between memory and experience, and the project of trying to ascertain the temporal microstructure of consciousness is misconceived. If two visual stimuli are presented in rapid succession, for example, most subjects will be able to identify the second stimulus far more reliably than the first. The standard interpretation of this sort of experiment, according to Dennett, is that in such cases the subject do not experience the first stimulus; they have no visual experience of it, since the occurrence of the second stimulus somehow interferes with the normal perceptual process. However, there is a second possible interpretation of the subject's responses. Perhaps the subjects do experience the first stimulus, but something interferes with their memory of this experience, and so when subsequently queried they deny having seen it[8]. Dennett maintains that in the case described (and others like it) neither interpretation is correct; there is simply no fact of the matter as to whether or not the subject's experiences the first stimulus: "the boundary between perception and memory [...] is not perfectly sharp" (Dennett & Kinsbourne 1992: 192). The assumption that there must be a determinate answer is grounded, in his opinion, in a sort of Cartesian conception of experience; according to such a conception, the question 'what is currently appearing on the stage of your consciousness?' always has a fully precise answer, an answer determined by the experiential contents present in the relevant subject's consciousness. Dennett's arguments, it has been claimed[9], have a 'verificationist slant', relying as they do on the principle that if there isn't evidence that P obtains or not, there is not fact of the matter as to whether or not P obtains. In my opinion, however, the argument is – more radically – that there simply is not a point at which external data become a conscious experience, or a conscious experience becomes a memory; it is not only difficult to discover – there is not such a region in our brain.

Inspired by Dennett's argument, Simon Prosser (2016) argued that the differences between the models of our temporal experience seem to concern the point at which external data become conscious[10] – where to put the line between memory and present consciousness. If, for example, it is a temporal extension of 500ms that, all together, is presented by our ocular nerves to our brain – let me say consciousness – so that we get to know the motion happened in front of us in the last 500ms, we have a direct experience of a temporal extended atom – and Extentionalists and Retentionalists are right; if, on the other hand, what happens is that different atoms – say, snapshots taken every 30ms – are presented in succession to our brain, we have a direct experience of a snapshot taken in the last 30ms by our eyes, and a short–term memory of the preceding snapshots – and Cinematists are

8 He calls the first mechanism *Stalinesque*, in that the experiences our perceptual systems produce do not accurately reflect the objective facts – in a manner reminiscent of Stalin's show-trials, and the second mechanism Orwellian, in that a false version of recent events is being rewritten, in a manner reminiscent of Orwell's dystopian societies.

9 See for example Dainton (2017).

10 This is Dennett's position, not mine: I think that the debate between the different models is genuine; simply, there are many ways to formulate it. I present Dennett's objection, however, because it gives an interesting insight in what I have to say in the following.

right. Such a debate, Prosser claims, seems to presuppose the fallacy of the Cartesian Theatre of the mind: the idea that there is a place and a moment in which the mere perceptual external data become conscious experience – as if a consciousness homunculus lived inside our head and watched the data presented to him. Besides reasonable worries of an infinite regress, the point is that our conscious experience is much more diversified and complex than this and, most of all, there isn't a finish line – a modern pineal gland, so to speak.

I agree on this with Dennett and Prosser: there isn't any fully conceptualised experience, happening at a definite time, as opposed to the process of obtaining it; it is the process itself that constitutes our experience in its different degrees of consciousness. Our cognition of the external world begins in the eye, in the ears, in the fingers; there is a process of rising consciousness, of course, but it would be vane to look for a precise locus where *we* come to meet an external phenomenon. It is pointless to try to distinguish between the real consciousness, the real person, the one that knows and understands, and the mere senses and nerves that, like tools and wires, bring information to the person, and to ask ourselves when the real self come to know something, when it is directly perceiving it or only remembering. If the difference between the models, then, consist simply in where the finish line should be placed, the non–existence of a finish line should deflate the whole debate. There is, of course, a phenomenal character associated to our processing of the external input, but – as Prosser (2016: 154) puts it – "we need not think that there is an answer to the question: '*when* is it like that for the subject?' [...] finer–grained questions about what the subject was experiencing at some specific time simply have no good answers".

At this point, one could start doubting that there is a genuine debate between the models of our temporal phenomenology. How is our phenomenological introspection supposed to give support to one or the other model if every model have the same capability to distinguish between different kind of experiences – simply, with different names? I agree with Prosser (2016: 136): "it is not really clear that we must choose between a theory that combines instantaneous contents with a short–term memory and a theory according to which there is a short–lived Specious Present". Will we be ever able to understand if our visual experience of a car going at 50mph results either from a comparison between the last snapshot and the preceding snapshots (short–term memory), or from an extended experience, or from a retention of the past experiences? But even more radically: is there a difference at all? It seems that all the models tell the same story about information processing; some preceding data must be combined with the last acquired data in order to produce our experience of motion. Of course, every model has a specific line of defense, but it doesn't seem that any of them contain elements capable of explaining phenomenological features that the other models can't; maybe they simply tell the same story with different names. If this is the case, the Specious Present would merely be another name to convey an idea, but nothing specific, nothing that we would be able to clearly individuate in our phenomenology.

A real, human experience takes time to be formed. Within this extension of time, it is not clear at all what is a direct perception and what a short–term memory; it is not even clear if there is, or should be, such a distinction. At which moment a

conscious experience ends and become a memory? The mathematical description of the present as a point on the temporal line should not be confused with our phenomenological present, which consists in a cognitive blend of the last apprehended data and – fact that is often underestimated – an anticipation of the future; our understanding of external environment is almost totally focused on our capability to intervene on it, or escape from it – everything, in our spatial and temporal basic observation of the world, is centred on our possibility to act.

This line of thought, however, sheds a new light on the debate about the Specious Present. If the dispute is merely about the point at which a visual information become a conscious experience and then become a memory, there isn't much hope. But if we intended it, instead, as a debate about the experiential 'here–now' – much nearer to our phenomenological and practical life – there is the possibility of a new dawn for it. In this case, the debate would be a genuine dispute about the best model to account for a certain phenomenal intuition, which has even a definite physiological counterpart, as I am going to show.

3. A Spatial Argument

Think of *what* are our senses, and *why* we have them; animals are the structured organisms that can move. The evolutionary reason of the functional and integrated role of our eyes, our ears, our nervous system, is to permit us to move in, intervene on or escape from the external environment or other animals. Our cognition of space and time is not unrelated to this logic; it would be an error to think of us as organisms with such and such characteristics, such and such temporal and spatial phenomenology, which are then lowered in a particular world, as Adam and Eve, shaped in Eden and then fallen on Earth. It is the world itself that shaped us and our evolution in it – our understanding of it, and the possibility to act, to move, is a central part of the project.

The mechanism that underlies our capability to grab objects is a perfect example of that. Before the discovery of brain neurons, it was natural to think that when we have an object in front of us – 20 centimetres or 2 metres away – we can decide to take it or not; if we decide to take it, our brain tells our arm to move and take it – or, if it is too far, tells our body to walk there and our arm to take it. Nowadays, however, we know that what really happens is much more complex[11]; there are motor neurons firing for every object in our proximity, and an inhibitory mechanism blocking the communication between them and the nerves; the motor neurons firing are not only continuously repeating to the arm *how to* coordinate to grab the object, they are literally telling it to take it; it is only thanks to the inhibiting role of the motor cortex – region of the cerebral cortex in the frontal lobe – that we don't actually take every object within our reach. Experiments[12] with the fMRI demonstrated how, if we move the object out of our possibility of reach, the motor neurons stop firing; of course we can still see the object and think 'I want to take it', but it is a completely different cognitive action. The curious fact is that,

11 See Rizzolatti et al. (1996), Rizzolatti et al. (2000), Sinigaglia (2008), Sinigaglia (2008 B).
12 See for example Bear et al. (1996).

if we give the person doing the experiment a stick (with which she could reach the object) the neurons start to fire again.

The moral of the story is that our possibility to directly and immediately intervene on the external environment is something that makes a great difference for us, it is the way we are built; we have senses *for* that reason. The 'here–now', related to our particular possibilities (how long are our arms, if we carry a stick or not, etc.), is central to our way to experience the world. At first glance, we could have thought that there isn't a clear sense in which an object is 'here'; whether it is 20 centimetres or 2 metres far, it is always 'here' in some sense. A debate regarding the exact point at which an object is 'spatially present' for us would have been meaningless; there is not a point at which the object changes its status and becomes present, we could have argued. But we are not Adam and Eve, the 'spatial here' – intended as 'what I can directly and immediately act on', 'what I can reach' – makes a great difference for us, both from a neurophysiological and a phenomenological point of view. There is an extended spatial 'here' clearly distinguished and individuated, and the debate regarding different models trying to describe the situation would be meaningful. I think that a similar point could be made in the temporal case; before turning to the temporal version of this reasoning, however, let me push the argument a little further.

Think of the famous *phi phenomenon* (the phenomenon of apparent motion). If two immobile spots of light on a screen are turned on and off at certain moments (generally the interstimulus interval must be around 30 frames per second), we see – instead of the two dots – one dot moving; Dainton's comment is that "evidently our brains are more than happy to supply us with experiences of motion at the least opportunity" (Dainton 2017 B, 1); but why? Again, I think that the reason is that we have been built by nature; if we see, in the sky or in a field, a black dot disappearing and very briefly another black dot appearing 30 centimetres at its left, the best explanation of that is that something is moving – and not that the first dot simply vanished in the sky, while another one miraculously came into existence; as Hoerl (2013: 162) puts it, "temporal features of reality can enter into the content of perception in the light of the immediate implications they possess for actions"; a similar point is also made by Morgan (2003: 61): "we are not normally conscious of a blur in moving objects: nor do we see them frozen in space–time. Instead, we see recognisable objects in motion". For the same reason, we see the leaves of the trees of the same green in the morning and in the evening – while, in reality, the two perceived colours are totally different, and mostly *not* green. All our conscious perceptions contribute to a successful and homogeneous experience of the world we live in; our brain continuously tries to connect every perception to familiar experiences, experiences that it knows how to react to, and this is why it is so easy to artificially create perceptual illusions in a laboratory using vanishing and appearing objects. My point, then, is that we can't think of our experiences without thinking of the way we are built; I agree with Hoerl (2013: 168) when he claims that "in perception [...] features of reality are represented in the light of their immediate relevance for the subject's actions".

My argument, then, is that the possibility to act on particular objects or events is something that makes a great difference for our experience of the external world;

just as in the spatial case there is a distinct sense in which the present is the 'reachable here', I believe there are good arguments to claim that in the temporal case 'the present' is the extension of time in which we can react to what's happening without the sensation that it is already 'too late'.

4. A Temporal Argument

I am proposing here an argument according to which it is possible to avoid the 'Cartesian theatre of the mind' fallacy in the temporal context; therefore, an intuition that could help Extended approaches to cognition to reject the strict Cartesian dualism, while establishing a definite meaning for the 'phenomenological immediacy' behind the notion of Specious Present. Prosser (2017: 154) claims that there is not any sensible difference between a sequence of two very close visual or auditory stimuli and two which are, on the contrary, separated by many seconds, above and beyond the platitude that one sequence takes more time. Are we able to say that there do really is a difference there?

Imagine, this time, that the the subject is asked to express a preference, an aesthetic judgement for example, between the colours showed in a visual display. In the case of the two colours separated by many seconds, the subject has the time to consider how much she likes the first colour, and some seconds later she sees the other colour and considers it, in turn. When she sees the second colour, she has already aesthetically judged the first one, which then *feel past*. From a phenomenological point of view, there is a clear, distinct sense in which the first visual experience is past – the subject already experienced it. In the second case, vice versa, the subject has not the physiological or phenomenological temporal space to experience the first colour and then the second; she knows that one preceded the other, but in a clear sense she is presented with two colours, she has one experience, and she has to pick her favourite colour out of that only experience. It is legitimate to have different intuitions on that, but it seems there is a strong case to argue that, at least in some specific cases, there is a very definite sense in which the two visual stimuli showed in a fraction of a second *feel equally present* and in which the two separated by many seconds don't. I take it to be a promising approach to understand what the Specious Present is. It is always possible, strictly speaking, to break down an experience and arrive at an atomic snapshot, but this is not what we have in mind when we think of an analysis of our temporal experience; we look for a model that is capable to translate phenomenological differences, instead of merely chronometric. The question of whether our temporal experience has a temporally extended content should be rejected, rather than answered, only if we think of it as the search of a definite moment at which information 'enters' or 'leaves' consciousness, or at which conscious experience starts and ends; if we, instead, consider the temporal content of our experience as the now with which we have a particular interactive role, just as the spatial here that our motor neurons are so good at individuating, then the question becomes interesting again. This is why I think that the interaction between mind, body and external environment should be a factor in our cognitive models. We could ask which is the extension of our present temporal experience just as we can ask which is the extension of our reachable here; in this case, some

sort of Specious Present seems an indispensable element of our temporal phenomenology. It is only *after* such a philosophical analysis that a neurophysiologist can define the nature of the Specious Present and measure its duration.

Finally, it is possible to see another link between our cognitive acts and the external world when we consider the role of anticipation in our experiential lives. The anticipation of the near–to–come future has an important role in the definition of our temporal present window; let me make an example, before turning to some experimental data. It is exactly talking of action and reaction that the role of anticipation came to the surface in such a strong way: not only, in fact, we react to something that we have seen, or touched, or smelled – that is gone, happened, in the past –, but we also act according to what we want to do in the future. It is only keeping in mind the combination of these two elements, past and future, reaction and action, that we can understand the window of the present, and not only because they always coexist in that window, but even and more significantly because one influence the other, as some very interesting experimental results confirm. One that I find significant, in this context, and that stresses the role of anticipation in our temporal experience of the present, is the phenomenon described by psychologists as backward masking: when, for example, we listen to music, the phenomenal character of our experience of a note is affected by the properties of the notes immediately prior to that note and after that note. Laurie Paul (2014: 186) comments:

> How can we 'see into the future' in this way? What is the basis for this experience of foreshadowing? There is debate about the mechanism involved in the cognitive processing of these temporally clustered events. Some have argued that it is a predictive effect [...]. Others have argued that it is what is called a 'postdictive effect' [...]. What matters here is that these foreshadowing and filling–in psychological effects are empirically well–documented, even if their source is not fully understood.

What I find particularly meaningful, in this context, is that this backward masking effect obtains only when the stimuli are close to one another; I see it as an experimental confirmation of the fact that there is indeed a phenomenological difference in the two sequences of notes 'do–re' beyond the mere fact that one sequence takes less time. When two notes are experienced in the same Specious Present, we have one reaction to two sounds – the second sound could even affect the phenomenal character of the first one, while the same thing does not happen when two sounds are more temporally separated. A similar reasoning can be extended to the other senses[13].

At this point, however, it is very easy to see a possible counterargument. If we don't specify the length of the specious present, it can be argued, we haven't actually brought about any kind of improvement to the debate. A stronger version of this argument could be that if we don't specify the length of the specious present, we don't even have arguments to maintain its existence, not even from a phenomenological point of view; affirmanti incumbit probatio, it could be said. If we are interested in maintaining the importance of the specious present, we should be prepared to answer the very reasonable worries about its extension. If the specious

13 See for example Saccuzzo et al. (1996), Herzog et al. (2013) for the studies of visual backward masking in schizophrenic patients.

present is our phenomenological window of presentness, how extended is that window? Is it something that should be decided by a phenomenological investigation, or is it the field of neurophysiological studies? I see the difficulty here, but I do not think it actually is a counterargument to what I am proposing, which is merely a change of perspective. A lot of experimental work should be done on this topic, and something more precise about the extendedness of that window could be said. My point, however, was that the important thing to understand here is that the experiments should not be conducted in the belief that we are going to discover the border between present consciousness and memory; instead of focusing on what particular cognitive acts are going on in our mind, the experimenters should focus on our possibility to act and react. My intuition is that the results (the extendedness of the specious present) will vary depending on the particular task set by the experimenters, but I don't see it as a problem. It is quite natural to think that the window of our phenomenal presence has a different extension depending on what we are doing or trying to do. The important point, in my opinion, is that once we have absorbed this change of perspective there is a definite way to decide if a Specious Present is part of our temporal phenomenology, to understand what it is, to measure what is its duration in a particular situation, and to decide which models of our temporal understanding are more apt to describe our phenomenal temporality.

Conclusions

In the first section, I have presented anti–Cartesian approaches of cognition, and the intuition they share that there is a constitutive interaction between mind, body and external environment. In the second section, I have presented the fallacy of the *Cartesian theater of the mind* and explained its theoretical premises. In the third section, I have presented a spatial argument against it, and argued that some case studies could give support to the idea of the mind stretching over the boundaries of the skull. In the fourth section, I have presented a temporal argument, and argued that even in this case the idea of an interaction between our cognitive life and the external world has *at least* a very strong intuitive palatability.

References:

Bear, Mark, Connors, Barry and Paradiso, Michael (1996), *Neuroscience. Exploring the Brain*. Baltimore: Williams and Wilkins.

Burge, Tyler (1979), "Individualism and the Mental", *Midwest Studies in Philosophy* 4: 73–122.

Clark, Andy, and Chalmers, David (1998), "The Extended Mind", *Analysis* 58 (1): 7–19.

Dainton, Barry (2000), *Stream of Consciousness. Unity and continuity in conscious experience*. London: Routledge.

Dainton, Barry. 2017. The Specious Present: Further Issues, in *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/consciousness–temporal/specious–present.html

Dainton, Barry (2017b), "Some Relevant Empirical Findings (Psychology, Psychophysics, Neuroscience)". In: *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/consciousness–temporal/empirical–findings.html

Di Paolo, Ezequiel, Marieke Rhohde and Hanne De Jaeger (2014), "Horizons for the Enactive Mind: Values, Social Interaction, and Play". In: J. Stewart, O. Gapenne, E. Di Paolo (eds.), *Enaction: Toward a New Paradigm for Cognitive Science*. Cambridge: MIT Press, pp. 29–48.
Greeno, James G. (1989). "A Perspective on Thinking", *American Psychologist* 44 (2): 134–141.
Herzog, Michael, Roinishvili, Maya, Chkonia, Eka, Brand, Andreas (2013), "Schizophrenia and Visual Backward Masking: A General Deficit of Target Enhancement", *Frontiers in Psychology* 4: 254.
Hoerl, Christoph (2013), "Husserl, The Absolute Flow, and Temporal Experience", *Philosophy and Phenomenological Research* 86: 376–411.
Hoerl, Christoph (2014), "Time and the Domain of Consciousness"., *Annals of the New York Academy of Sciences* 1326: 90–96.
James, William (1890), *Principles of Psychology*. London: Macmillan.
Lee, Geoffrey (2014), Extensionalism, Atomism, and Continuity. In: L. Nathan Oaklander (ed.), *Debates in the Metaphysics of Time*. London: Bloomsbury Academic, pp. 149–174.
Le Poidevin, Robert (2007), *The Images of Time: An Essay on Temporal Representation*. Oxford: Oxford University Press.
Lockwood, Michael (2005), *The Labyrinth of Time: Introducing the Universe*. Oxford: Oxford University Press.
McClelland, J. L, D. E. Rumelhart, and G. E. Hinton (1986), "The Appeal of Parallel Distributed Processing". In: J. L. McClelland and D. E. Rumelhart (eds.), *Parallel Distributed Processing* (Vol. 1). Cambridge, MA: MIT Press, pp. 3–44.
Paul, Laurie (2004), "Experience and the Arrow". In: Alastair Wilson (ed.), *Chance and Temporal Asymmetry*. Oxford: Oxford University Press, pp. 175–193.
Pöppel, Ernst (1997), "The Brain's Way to Create 'Nowness'". In: Harald Atmanspacher and Eva Ruhnau (eds.), *Time, Temporality, Now. Experiencing Time and Concepts of Time in an Interdisciplinary Perspective*. Berlin: Springer–Verlag, pp. 107–120.
Prosser, Simon (2016), *Experiencing Time*. Oxford: Oxford University Press.
Prosser, Simon (2017), "Rethinking the Specious Present". In: Ian Phillips (ed.), *Routledge Handbook of Philosophy of Temporal Experience*. London: Routledge, pp. 146–156.
Putnam, Hilary (1975), "The Meaning of 'Meaning'". In: K. Gunderson (ed.), *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press, pp. 131–193.
Rizzolatti, Giacomo, Gallese, Vittorio, Fogassi, Leonardo and Fadiga, Luciano (1996), "Premotor Cortex and the Recognition of Motor Actions", *Cognitive Brain Research* 3: 131–141.
Rizzolatti, Giacomo, Gallese, Vittorio and Leonardo Fogassi (2000), "Cortical Mechanisms Subserving Object Grasping and Action Recognition: A New View on the Cortical Motor Functions". In: Mauro Gazzaniga (ed.), *The Cognitive Neurosciences*. Cambridge: MIT Press, pp. 539–552
Rowlands, Mark (2010), *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge: MIT Press.
Saccuzzo, Dennis, Cadenhead, Kristin and Braff. David (1996), "Backward versus forward Visual Masking Deficits in Schizophrenic Patients: Centrally, not Peripherally, Mediated?", *American Journal of Psychiatry* 153: 1564–1570.
Sinigaglia, Corrado (2008), "Enactive Understanding and Motor Intentionality". In: F. Morganti, M., A. Carassa, G. Riva (eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective to Study of Interactions*. Amsterdam: IOS Press, pp. 17–32.
Sinigaglia, Corrado (2008b), "Mirror Neurons: This Is the Question", *Journal of Consciousness Studies* 15: 70–92.
Strawson, Galen (2009), *Selves*. Oxford: Oxford University Press.
Suchman, Lucy (1987), *Plans and Situated Actions*. Cambridge: Cambridge University Press.
Szelag, Elzbieta (1997), "Temporal Integration of the Brain as Studied with the Metronome Paradigm". In: Harald Atmanspacher, Eva Ruhnau (eds.), *Time, Temporality, Now. Experiencing Time and Concepts of Time in an Interdisciplinary Perspective*, pp. 121–131.

Ullman, Shimon and Whitman Richards (1984), *Image Understanding*. Norwood, NJ: Ablex.
Varela, Francisco, Rosch Eleanor and Evan Thompson (1991), *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.
Wittmann, Marc (2011), "Moments in Time", *Frontiers in Integrative Neuroscience* 5: 66.

Andrea Rozeli

Duh izvan glave: dva argumenta u prilog uronjene kognicije

Apstrakt

U ovom radu branim situirane pristupe kogniciji i ideju da su duh, telo i spoljašnji svet nerazdvojivi. U prvom odeljku predstavljam nekoliko anti-kartezijanskih pristupa razumevanju kognicije i ispitujem deljenu intuiciju da postoji konstitutivna interakcija između duha, tela i spoljašnje sredine. U drugom odeljku predstavljam logičku grešku kartezijanskog teatra i objašnjavam njene teorijske pretpostavke. U trećem odeljku predstavljam prostorni argument protiv kartezijanskog teatra i pokazujem da neke studije slučajeva mogu da daju potporu ideji da se duh proteže izvan granica lobanje. U četvrtom odeljku predstavljam vremenski argument i pokazujem da čak i u ovom slučaju ideja o interakciji između našeg kognitivnog života i spoljašnjeg sveta ima barem veoma snažnu intuitivnu prihvatljivost.

Ključne reči: proširena kognicija, duh, anti-kartezijanizam, situirana kognicija, uronjena kognicija

II

STUDIES AND ARTICLES

STUDIJE I ČLANCI

UDK: 111.1 : 14 Heidegger M.
https://doi.org/10.2298/FID1804519W
Original Scientific Article
Received: 19.06.2018. Accepted: 12.09.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Hongjian Wang

ONTOLOGISIERUNG DER PRAXIS UND VORTHEORETISCHER PRAXISBEGRIFF. WEITERE ÜBERLEGUNGEN ZUM PROBLEM DER HEIDEGGER'SCHEN PRAKTISCHEN PHILOSOPHIE

ZUSAMMENFASSUNG
Obwohl in der Forschung nicht selten vorkommt, dass Heideggers Gedanken im Lichte des Verhältnisses zwischen Theorie und Praxis betrachtet werden, nimmt Heidegger selbst eine Position gegen diese Dichotomie ein. Um die Kritik Heideggers an dieser Dichotomie und seine Konzeption der ursprünglichen Einheit zwischen Theorie und Praxis zu verstehen, muss Heideggers Lektüre der aristotelischen praktischen Philosophie genauer betrachtet werden. Hier wird versucht zu zeigen, dass Heideggers Ontologisierung der Praxis darauf abzielt, einen vortheoretischen, zur Einheit der Theorie und Poiesis unterschiedlichen Praxisbegriff herauszuarbeiten. Die Ontologisierung kommt durch die Methode der formalen Anzeige zustande, die in diesem Sinne im Mittelpunkt Heideggers Philosophie steht.

SCHLAGWÖRTER
Praktische Philosophie, Ontologisierung, vortheoretisch, Formale Anzeige

In seinem Lexikon-Artikel zum Begriff der Praxis hat A. Schmidt bemerkt, dass er „zu den wichtigsten wie umstrittensten Kategorien zeitgenössischen Philosophierens zählt" (Schmidt 1973: 1111). Der Praxisbegriff, mit dem normalerweise die *menschliche* Tätigkeit bezeichnet wird, ist vor allem insofern kontrovers, als er auf verschiedenen Ebenen Verwendung finden kann. Im weitesten Sinne kann er durch den Anthropomorphismus für alles Seiende gelten, während er im engsten Sinne ausschließlich auf die menschliche Interaktion und die gemeinschaftliche Tätigkeit des Menschen verweist. (Vgl. Ritter/Gründer 1989: 1278, 1284) Ein weiterer Grund dafür ist, dass der Denkrahmen des Verhältnisses zwischen Theorie und Praxis vieldeutig ist. Im Allgemeinen wird die Praxis als die konkrete Anwendung der Theorie angesehen. Gleichwohl ist Gadamer zufolge der klassische Gegensatz „letzten Endes ein Gegensatz des Wissens, nicht der Gegensatz zwischen Wissenschaft und Anwendung der Wissenschaft" (Gadamer 1972: 326). Diese Einsicht stammt vermutlich von Heidegger, der in *Sein und Zeit* bereits indiziert, dass die Praxis auch ihre eigene „Sicht" hat. (Vgl. Heidegger 1977, 93)

Diese aufschlussreiche Ansicht verdankt Heidegger seiner Aneignung des griechischen Denkens. Folgt man ihm darin, dann sollte vor allem auf die Quelle seiner Anschauungen eingegangen werden, um die ursprüngliche vorphilosophische Bedeutung von Praxis zu entdecken. Das griechische Wort πρᾶξις, dessen Verbform

Hongjian Wang: Assistant Professor, Yuelu Academy, Hunan University, China; hongjianhongjian@gmail.com.

πράσσειν als „einen Weg zurücklegen", dann mit „vollbringen" oder „vollführen" übersetzt wird, ist etymologisch mit πέρα („darüber hinaus", „jenseits") eng verbunden. (Vgl. Ritter/Gründer 1989: 1277) Bemerkenswert sind ja die folgenden beiden Charaktere des Praxisbegriffs im vorphilosophischen Sinne: Erstens besitzt er noch keinen Bezug auf den Wert, sondern stellt sich als die neutrale Bezeichnung des menschlichen Handelns dar; zweitens kommt die Dichotomie zwischen Theorie und Praxis noch nicht zum Vorschein. Es lässt sich fragen, wie diese beiden Charaktere Heideggers Auffassung der Praxis beeinflussen und bestimmen. Im Folgenden wird vor allem gezeigt, wie Heidegger diese Dichotomie kritisiert und die ursprüngliche Einheit zwischen Theorie und Praxis (im Sinne der Poiesis, also der herstellenden Praxis) feststellt. Danach wird anhand der Lektüre von F. Volpi über die Aristoteles-Interpretation Heideggers erläutert, wie der Freiburger Philosoph durch die Neutralisierung, Formalisierung und Ontologisierung der traditionellen praktischen Philosophie den vortheoretischen, eigentlichen Praxisbegriff entwickelt.

Heideggers Kritik an der Dichotomie zwischen Theorie und Praxis

Heideggers Ablehnung der Dichotomie zwischen Theorie und Praxis wird zuerst in seiner Vorlesung aus dem KNS 1919 deutlich, in der die Vorherrschaft des Theoretischen heftiger Kritik unterzogen wird: „Diese Vorherrschaft des Theoretischen muß gebrochen werden, [...] weil das Theoretische selbst und als solches in ein Vortheoretisches zurückweist." (Heidegger 1999: 59) In diesem Zitat kommt das Stichwort des „Vortheoretischen", das dem Theoretischen vorangeht, zum Vorschein. Gleichwohl ist das mit diesem Ausdruck Gemeinte nicht mit dem „Praktischen" gleichzusetzen. Zu beachten ist, dass sich Heidegger außerdem auch gegen diejenigen Interpreten wendet, die nur der Praxis und dem Praktischen Beachtung schenken mögen. Mit Blick auf Heideggers Denken soll gezeigt werden, dass darin weder die Theorie noch die Praxis über ein Vorrecht verfügt, sondern es auf die radikale Ablehnung des Rahmens des Verhältnisses zwischen Theorie und Praxis im Ganzen angelegt ist. Diese Sichtweise bleibt bis zur Daseinsanalyse in *Sein und Zeit* unverändert, wo das Theorie-Praxis-Verhältnis zwar mehrfach, aber jeweils nur am Rande erwähnt wird.[1]

1 Die folgenden drei Zitate ermöglichen ein besseres Verständnis dieses Sachverhalts: „Künstlich dogmatisch beschnitten ist der thematische Gegenstand, wenn man sich ‚zunächst' auf ein ‚theoretisches Subjekt' beschränkt, um es dann ‚nach der praktischen Seite' in einer beigeführten ‚Ethik' zu ergänzen." Heidegger 1977, 418. „Das Phänomen drückt daher keineswegs einen Vorrang des ‚praktischen' Verhaltens vor dem theoretischen aus. Das nur anschauende Bestimmen eines Vorhandenen hat nicht weniger den Charakter der Sorge als eine ‚politische Aktion' oder das ausruhende Sichvergnügen." Heidegger 1977, 257. „Sorge aber als besorgende Fürsorge umfaßt das Sein des Daseins so ursprünglich und ganz, daß sie in der Scheidung von theoretischem und praktischem Verhalten je schon als Ganzes vorausgesetzt werden muß und aus diesen Vermögen nicht erst zusammengebaut werden kann mit Hilfe einer notwendig grundlosen, weil existenzial ungegründeten Dialektik." Heidegger 1977, 398. Durch diese drei Zitate werden jeweils mit verschiedenen Betonungen der Vorrang der Theorie, der Vorrang der Praxis und der nachträgliche Zusammenschluss von Theorie und Praxis kritisiert. Dem Anschein nach erschöpft Heidegger alle Möglichkeiten bezüglich dieses Rahmens; deshalb kann gefolgert werden, dass seine Daseinsanalyse außerhalb dieses Rahmens liege.

Gleichwohl ist es nicht überzeugend, zu behaupten, dass Heidegger durch seine Daseinsanalyse die Problematik der Scheidung von Praxis und Theorie völlig zu lösen vermocht habe. Fraglich ist, ob es zwischen Dasein bzw. Sorge einerseits und Praxis andererseits eine unüberwindbare Differenz gibt. Diese Frage stellt F. Volpi und er kritisiert Heidegger folgendermaßen: „Er [Heidegger] beteuert, die Sorge gehe jeder Unterscheidung von Theorie und Praxis voraus. [...] Doch die Tatsache, daß er diese Beteuerung für nötig hält, verrät, daß am Ende eine thematische Verwandtschaft zwischen der Sorge und dem aristotelischen Begriff der ὄρεξις besteht." (Volpi 1989: 237) Ein weiterer Einwand gegen Heidegger liegt darin, dass das Dasein als Praxis in einem anderen Sinne verstanden werden kann.[2] Es kommt darauf an, wie Praxis, die bei Heidegger auf Ablehnung stößt, zu verstehen ist. Es ist unschwer, darauf hinzuweisen, dass Heidegger nachdrücklich gegen eine theoretisierte Praxisauffassung Stellung nimmt. Gleichwohl kann man nicht ohne weiteres behaupten, dass bei ihm auch der vortheoretische Praxisbegriff zurückgewiesen wird. Diese vortheoretische Praxis lässt sich nun aber anders bezeichnen, so entwickelt Heidegger etwa die Begriffe „Dasein" bzw. „Sorge", um den Praxisbegriff zu ersetzen. Es soll in diesem Zusammenhang geklärt werden, wodurch diese Ersetzung motiviert ist. Der Klärung dieser Frage wird aber ein Exkurs vorausgeschickt, zu zeigen, wie Heideggers Gedanke im Blick auf das Verhältnis zwischen Theorie und Praxis betrachtet wird und woran es dieser Interpretationslinie mangelt.

Die Betrachtung des Gedankens Heideggers im Lichte des Verhältnisses zwischen Theorie und Praxis am Beispiel der Untersuchungen von Prauss und Gethmann

Es ist zu eruieren, dass die Betrachtung von Heideggers Denken im Lichte des Verhältnisses von Theorie und Praxis eine der wichtigsten Interpretationslinien in der Forschung ausmacht. G. Prauss stellt die Frage, ob die Untersuchung in *Sein und Zeit* auf die Enthüllung des Vorrangs der Praxis gegenüber der Theorie abziele, oder eher darauf, im Grunde den Gegensatz zwischen Theorie und Praxis zu eliminieren. (Vgl. Prauss 1977: 9f.) Ähnliches äußert auch E. Tugendhat, was sich in den von ihm herausgestellten beiden Thesen verkörpert:

> Was ich im gegenwärtigen Zusammenhang mit der schwächeren These meine, ist die Behauptung, daß der Sinn von Sein verschieden ist, je nachdem, ob es theoretisch, als zu konstatierendes, oder im vorhin beschriebenen Sinn praktisch, als zu vollziehendes, gemeint ist. Die stärkere These ist, daß „Sein" im Sinn von Vorhandensein nicht nur nicht der einzige Sinn von „Sein" ist, sondern daß dieser Seinssinn gegenüber dem des Zu-seins abkünftig ist. (Tugendhat 1979: 181)

Damit lenkt Tugendhat den Schwerpunkt auf das Sein und auf diese Weise interpretiert er „Praxis" und „Theorie" als zwei Seinsweisen, die jeweils dem Dasein (im Sinne des Zu-seins) und dem Vorhandensein entsprechen. Die schwächere These versteht die beiden als nebeneinander bestehende Hinsichten, die zu vereinbaren sind, während die stärkere These durch die Reduktion der Vorhandenheit auf die

2 Beispielsweise unternimmt B. Sitter den Versuch, eine Ethik aus der Analytik der Fürsorge in *Sein und Zeit* abzuleiten. Vgl. Sitter 1975: 153–173.

Zuhandenheit den Vorrang der Praxis gegenüber der Theorie behauptet. Freilich ist auch zu sehen, dass beide Thesen nicht nur gegeneinanderstehen, sondern auch ihre Gemeinsamkeit haben, die sich in der Ablehnung der einfachen Ausrichtung des Seins auf das Vorhandensein bekundet. Im Folgenden wird zu zeigen versucht, wie die beiden Thesen in der Forschung zum Ausdruck kommen und was durch sie dennoch übersehen wird.

Die Auffassung von C. F. Gethmann, „die Philosophie von *Sein und Zeit*" als „die im deutschsprachigen Bereich früheste Konzeption eines konsequenten Pragmatismus" (Gethmann 1993: 285) zu erachten, geht bereits auf K. O. Apel zurück, von dem Heidegger ein „implizite[r] Pragmatismus" (Apel 1973: 267) zugesprochen wird. Das zeigt sich nach Gethmann vor allem in der Betonung des Umgangs und seiner Umsicht von Heidegger. Nach der Entdeckung der Differenz zwischen Umsicht und Erkennen versucht Gethmann zu zeigen, dass das Erkennen der Umsicht entstammt, was sich im Verhältnis von Auslegung und Aussage und in der Genesis des wissenschaftlichen Erkennens ausdrückt. Das theoretische Erkennen, also die Abstraktion des Handelns aus einer bestimmten Perspektive, stellt sich als thematisierend und objektivierend dar und in diesem Sinne entspringt es „aus seinen ursprünglichen operativen Bezügen" (Gethmann 1993: 298). Mit der Enthüllung des Vorrangs der praktischen Umsicht gegenüber dem theoretischen Erkennen geht einher, dass die Theorie ihre Grundlage in der Lebenserfahrung und Handlungssituation hat.

Nach Gethmann hat Heidegger darüber hinaus das Modell „Umgang - Umsicht" als Ersatz für „Handeln - Erkennen" bzw. „Praxis - Theorie" herausgearbeitet, weil „die Umsicht im Unterschied zum Erkennen nicht an der Wahrheit, sondern an der Zweckhaftigkeit, also an Kategorien des Erfolgs orientiert [ist]". In diesem Sinne stellt die Umsicht kein reflexives Erkennen dar, sondern „ein[en] sich selbst nicht thematisierende[n] Vollzug". (Gethmann 1993: 289) Es ist auch zu beachten, dass der Umgang immer mit der Umsicht verbunden ist, während sich Praxis und Theorie nach ihren traditionellen Bedeutungen voneinander unterscheiden, sofern Theorie hier als „unumsichtiges Nur-Hinsehen", bloße Praxis aber als „sichtloses Handeln" (Gethmann 1993: 290) betrachtet wird.

Während die pragmatische Interpretation von Heidegger darauf angelegt ist, das Modell „Umgang - Umsicht" zu entwickeln und den Vorrang der Praxis gegenüber der Theorie zu betonen, versucht Prauss das Parallele zwischen Umsicht und Theorie aufzuzeigen. Vor allem weist er auf zwei Schwierigkeiten in Heideggers Behandlung des Verhältnisses von Praxis und Theorie hin. Erstens ist die Praxis keineswegs blind, sondern sie erweist sich als der mit Umsicht versehene Umgang, eine Umsicht, die jedoch eine für den Vollzug wegweisende Theorie ausmacht. Daraus ergibt sich, dass Praxis ohne „Theorie" nicht möglich ist. Zweitens ist das theoretische Erkennen auch nicht rein betrachtend, sondern es „muß selber praktischen Einschlag besitzen", und in diesem Sinne lässt sich folgern: „Praktisch soll Theorie als solche selbst sein und nicht lediglich dadurch, daß Praxis mit ihr einherläuft." (Prauss 1988: 179) Gerade von den vorliegenden beiden Gesichtspunkten her muss das Theorie-Praxis-Verhältnis neu interpretiert werden. Prauss verdeutlicht das Parallele zwischen Umsicht und Theorie, indem er aufdeckt, dass sich die Umsicht versehen kann, wie auch die Theorie falsch sein kann. Das Besorgen zeigt

sich dann als „gestörtes", obwohl es auf das „Ungestörte" abzielt (Prauss 1988: 182), das in der Umsicht funktioniert wie das „Wahre" in der Theorie.

In dieser Einsicht von Prauss kommt die Einheit zwischen Erkennen und Handeln zum Vorschein. Das Handeln setzt den Erfolg zum Ziel und kann erfolgreich oder erfolglos sein. Dementsprechend hat das Erkennen, dessen Ergebnis wahr oder falsch sein kann, die Wahrheit zum Ziel. In diesem Zusammenhang lässt sich die Wahrheit auch als Erfolg am Erkennen auslegen, wodurch eine Brücke zwischen Handeln und Erkennen geschlagen wird.[3] Prauss ist darin wohl zuzustimmen, da seine Behauptung impliziert, dass Heidegger nicht so sehr das Erkennen auf das Handeln reduziert als das Erkennen auch als eine Art von Handeln resp. Besorgen ansieht.

Es lässt sich jedoch erkennen, dass die Aufmerksamkeit sowohl von Prauss als auch von Gethmann ausschließlich auf die Behandlung des Theorie-Praxis-Verhältnisses gerichtet ist und somit ein wichtiger Unterschied vergessen wird, nämlich der innerhalb des Praxisbegriffs selbst. Dieser Unterschied geht bereits auf Aristoteles zurück, der die Poiesis von der Praxis im engeren Sinne trennt. Die beiden Interpreten schenken zwar, um mit Heidegger zu reden, dem Unterschied zwischen Vorhandensein und Zuhandensein genug Beachtung, der Unterschied zwischen Zuhandensein und Dasein wird jedoch beiseitegeschoben.[4] Dieser Unterschied bei Heidegger entspricht der aristotelischen Unterscheidung zwischen Praxis und Poiesis, zwischen φρόνησις und τέχνη. Deswegen ist im Folgenden zunächst diese Unterscheidung ausführlich zu betrachten.

Betrachtung des Verhältnisses zwischen θεωρία und ποίησις bei Aristoteles

Im Vergleich zur weitverbreiteten Dichotomie zwischen Theorie und Praxis wurde bei Aristoteles ursprünglich eine Dreiteilung der menschlichen Tätigkeiten zwischen θεωρία, πρᾶξις und ποίησις herausgearbeitet, welche das weitere Denken über dieses Thema in großem Maß bestimmt hat. Erstens unterscheidet Aristoteles gemäß dem Gegenstand der Tätigkeit die θεωρία von den anderen beiden. Den Gegenstand der θεωρία macht das unveränderliche Seiende aus, während die πρᾶξις und ποίησις das, was sich anders verhalten kann, zum Gegenstand haben. Demzufolge behauptet er, dass die θεωρία hinsichtlich des Gegenstandes in seiner Seiendheit höher stehe als die anderen beiden. (Aristoteles 1894: 1141b3-4) Daraus folgt für ihn, dass die

3 Das verbindet Prauss außerdem mit der Radikalisierung des Begriffs der Subjektivität bei Heidegger, die seine kritische Auseinandersetzung mit der Erkenntnistheorie bekundet: „Sich von jener unverhofften Einsicht bis zu dieser Selbstkritik und Radikalisierung seiner Theorie des Besorgens bestürzen zu lassen hätte Heidegger mithin dazu geführt, des Subjekts Subjektivität durch eine Theorie derselben so weit zu entfalten, daß sie alles überboten hätte, was die Neuzeit an Vergleichbarem bis dahin schon geboten hatte: Nicht erst Handeln, sondern schon Erkennen ist gar nichts anderes als Erfolgsintention des menschlichen Daseins als eines Subjekts." Prauss 1988: 183.

4 Beispielsweise stellt es für Prauss einen wichtigen Beleg seiner Behauptung dar, dass die Umsicht sich versehen kann. Aus Heideggers Interpretation von Aristoteles ist aber zu konstatieren, dass ausschließlich die τέχνη fehlbar sei, während sich die φρόνησις, die dem Dasein entspricht, nicht versehen könne. Vgl. Heidegger 1992: 53–55.

θεωρία die höchste, eigentlichste πρᾶξις sei. Außerdem wird bei der Klassifikation der Tätigkeiten ein anderes Kriterium eingeführt, nämlich, ob sich der Zweck der Tätigkeit innerhalb ihrer selbst befindet. Demgemäß wird die ποίησις von den anderen, insbesondere von der πρᾶξις unterschieden. Im Anfangskapitel der *Nikomachischen Ethik* ist von den verschiedenartigen Zwecken der Tätigkeiten die Rede: die einen sind Tätigkeiten selbst, die anderen sind darüber hinaus Werke der Tätigkeiten. (Aristoteles 1894: 1094a4-5) Die ποίησις bezieht sich auf das Werk, das die Tätigkeit zu Ende bringt und sich danach von ihr entfernt, während die Tätigkeit um ihrer selbst willen πρᾶξις genannt wird. Daher kommt Aristoteles zum Schluss, dass „Handeln und Herstellen zu unterschiedlichen Gattungen gehören." (Aristoteles 1894: 1140b3-4) Zudem lässt sich die ποίησις als die uneigentliche Form der πρᾶξις ansehen. Gemäß den vorliegenden beiden Kriterien entwickelt Aristoteles die Dreiteilung zwischen θεωρία, πρᾶξις und ποίησις und kommt dann zu der folgenden Auffassung: Die θεωρία ist die eigentliche πρᾶξις, während die ποίησις die uneigentliche πρᾶξις darstellt. Gerade das führt dazu, dass das Moment der πρᾶξις bei den Nachfolgern von Aristoteles verloren geht. (Vgl. Backman 2007: 209f.)

Fraglich ist, wie die θεωρία nach dem Verlust der πρᾶξις mit der ποίησις im Zusammenhang steht. Darum gilt es zunächst einmal den Begriff der θεωρία zu erklären. Zwar wird θεωρία durch die lateinische Übersetzung als *speculatio*, reines Betrachten, verstanden, aber ursprünglich bedeutet es einfach das Sehen, das die Natur des Menschen ausmacht. Die theoretische Weisheit, also die σοφία, erweist sich als eine der eigentlichen Sichtweisen. (Vgl. Heidegger 2005: 376, 387) Es ist darauf aufmerksam zu machen, dass den Ausgangspunkt vom Weg zur σοφία nichts anderes als das faktische Leben darstellt: „Aristoteles nimmt aus dem faktischen Leben, aus der Weise seines eigenen umgänglichen Sprechens die Rede von σοφώτερον - verstehender sein als - auf." (Heidegger 2005: 387) Aus dem Wort σοφώτερον, das im Text als „verstehender" bzw. „mehr an Verstehen" übersetzt wird und als die komparative Form von σοφός gilt, ist zu erkennen, dass sich der Übergang vom faktischen Lebensumgang zum reinen Verstehen durch eine alltägliche komparative Perspektive ereignet. Die σοφία, die im Höhepunkt des Prozesses der Übersteigung liegt, verwurzelt sich eben in der Lebenserfahrung, genauer in der Herstellungserfahrung.[5]

In der *Sophistes-Vorlesung* aus dem WS 1924/25 hat Heidegger diese Analytik ausführlicher wiederholt. Die fünf Stufen, also αἴσθησις, ἐμπειρία, τέχνη, ἐπιστήμη und σοφία, lassen sich in zwei Ebenen zuordnen, wobei die τέχνη, die Verhaltensweise des Herstellens, eine wichtige Rolle spielt: Die erste Ebene läuft von αἴσθησις über ἐμπειρία bis zu τέχνη, während die zweite von τέχνη über ἐπιστήμη bis zu σοφία geht. Unter αἴσθησις versteht man das rein sinnliche Sehen, in dem der λόγος, das Allgemeine, noch nicht zum Vorschein kommt, während in der Erfahrung der Verweisungszusammenhang hervortritt, der als „Sobald - dann" bezeichnet wird. Dieser Zusammenhang wird aber durch seine Wiederholung weiterhin hervorgehoben und über „Wenn - so" zu „Weil - deshalb" entwickelt, was ein Indiz dafür ist, dass

5 Hier ist darauf zu verweisen, dass die σοφία in dem vorphilosophischen Sinne „immer auf technisch-praktisches Wissen des Sich-Auskennens, auf eine Kenner- und Meisterschaft in einem Gebiet bezogen" ist, weshalb der σοφός „Könner und Sachkundiger" ist. Elm 2007: 258.

das εἶδος, der *Was*-Zusammenhang, nunmehr in Erscheinung tritt. Das geschieht in der τέχνη, welche die Vereinigung von ποίησις und εἶδος bedeutet. Wird darin das εἶδος weiter betont, dann wird τέχνη zu ἐπιστήμη. In diesem Sinne liegt der Unterschied zwischen τέχνη und ἐπιστήμη darin, inwieweit das εἶδος hervorgehoben wird. Die letzte Stufe, also die σοφία, welche die rein verstehende, theoretische Lebensweise bedeutet, wird erst unter der Voraussetzung der radikalen Ausrichtung auf das εἶδος erreicht. (Vgl. Heidegger 1992: 69-77, 91–94)

Aus der vorliegenden Analytik lässt sich folgern, dass diese fünf Stufen die vollständige Beschreibung des Übergangs vom faktischen Leben zur Theorie ausmachen. Das Verhältnis zwischen den verschiedenen Stufen bezeichnet Heidegger als „Mehr" (μᾶλλον), dessen Grundzüge folgendermaßen zusammengefasst werden können: die Hervorhebung des Gegenstandes in seinem εἶδος und die Ausblendung des faktischen Umgangs mit der Umwelt. Das Womit des Umgangs wird schrittweise zum Worauf des Hinsehens, indessen tritt das Theoretische, das Formale und Allgemeine, allmählich hervor. (Vgl. Vetter 2007: 88) Durch die vorliegende Erörterung ist zu erkennen, dass θεωρία und ποίησις bei Aristoteles kontinuierlich sind, indem jene als die Radikalisierung von dieser angesehen werden kann. Um Aristoteles aber zu übersteigen, stellt sich nun die Frage, wo die echte, nicht auf die θεωρία zurückzuführende Praxis zu finden sei.

Die Ontologisierung des Praxisbegriffs und der praktischen Philosophie. Weitere Überlegungen zur These Volpis

Um Heideggers Vertiefung vom traditionellen Praxisbegriff zu verdeutlichen, muss vor allem auf seine Interpretation von Aristoteles hingewiesen werden. Dabei trifft man aber auf die These von F. Volpi, dass sich die frühe Bestimmung der Philosophie Heideggers gewissermaßen aus der Neutralisierung, Formalisierung und Ontologisierung der praktischen Philosophie von Aristoteles ergibt. Im Folgenden ist zunächst einmal diese These in ihren Grundzügen zu skizzieren:

1) Heideggers Daseinsanalytik gilt als die Assimilierung sowie Umformung aristotelischer praktischer Philosophie. Insbesondere gilt es zu zeigen, dass die Begriffe von Vorhandenheit, Zuhandenheit und Dasein den Begriffen von θεωρία, ποίησις und πρᾶξις bei Aristoteles entsprechen.
2) Die Umformung vollzieht sich auf dreierlei Ebenen: Ontologisierung, hierarchische Verschiebung und einheitliche Zuordnung.
3) Darin bedeutet Ontologisierung die Hinwendung von der *Verhaltensweise des Menschen* zur *Seinsweise des Seienden (Daseins)*. (Vgl. Volpi 1989: 230–238)

Es lässt sich zeigen, dass der Leitfaden der frühen Philosophie Heideggers nichts anderes als das Lebensverstehen ist (vgl. Volpi 2007a: 168), welches in der Konzeption „der phänomenologischen Ontologie des faktischen Lebens" (Heidegger 2000a: 44) zum Ausdruck kommt. Zugleich ist darauf zu verweisen, dass die Aufgabe des Lebensverstehens im Kontext der praktischen Philosophie verständlicher ist. Dies ist so zu verstehen, dass die praktische Philosophie hier als eine Deutungsrichtung gilt, von der aus das Lebensverstehen zu vollziehen ist. (Vgl. Volpi 2007a: 170)

Trotz der Tatsache, dass Volpi die existenziale Analyse Heideggers mit der praktischen Philosophie verbunden sieht, gibt er doch zu, dass diese durch Heidegger einer Ontologisierung unterzogen wird, die wiederum mit ihrer Neutralisierung einhergeht. Allerdings hat er an einer anderen Stelle geäußert: „Die existenziale Analyse ist nicht ‚wertneutral', sondern impliziert in ihrem Vollzug die Entschlossenheit zur Eigentlichkeit." (Volpi 2007b: 50) In diesem Sinne scheint Volpi mit sich selbst in Widerspruch zu kommen: Ob Heideggers existenziale Analyse wertneutral ist oder nicht, dazu kann bei ihm keine konsequente Antwort gefunden werden. Tatsächlich ist nicht zu übersehen, dass Volpi bei der Formulierung seiner These sehr vorsichtig ist: „Gewiß, bei seiner Wiederaufnahme radikalisiert Heidegger in ontologischer Hinsicht die aristotelischen Bestimmungen, und nachdem er diese Ontologisierung einmal vollzogen hat, nimmt er von Aristoteles kritisch Abstand." (Volpi 1989: 239) Hier gilt es jedoch, das Zauberwort „Ontologisierung" weiter zu verdeutlichen. Dafür wird Volpi zufolge auf ein Beispiel verwiesen, das auf den Unterschied zwischen ποίησις und πρᾶξις bezogen ist:

> [Es ist zu beachten,] daß es sich dabei nicht um eine ontische Unterscheidung handelt, also um eine Unterscheidung, die sich auf einzelne Handlungsvollzüge bezieht, wovon die einen ποίησις und die anderen πρᾶξις sind; diese Unterscheidung hat vielmehr ontologischen Charakter, sie grenzt zwei verschiedene Seinsweisen ab, die ontisch voneinander nicht abgehoben werden: eine Rede halten kann z. B. die Seinsweise einer ποίησις haben (etwa im Sinne der Herstellung von λόγοι vonseiten eines Redners), sie kann aber auch die einer πρᾶξις haben (etwa im Sinne einer politischen Rede). (Volpi 1989: 231)

Daraus wird deutlich, dass sich die sogenannte Ontologisierung in Volpis Augen schon bei Aristoteles anbahnt, da diese durch die Unterscheidung zwischen ποίησις und πρᾶξις indiziert sei. Das liegt daran, dass beide von Aristoteles nicht als verschiedenartige Tätigkeiten, sondern als unterschiedliche Perspektiven auf dieselbe Tätigkeit gelten.[6] Anders gesagt, es geht hier um die Umwendung von der konkreten Tätigkeit zur Verhaltensweise oder Sichtweise, also vom *Was* zum *Wie*. Dementsprechend kommt der Unterschied zwischen φρόνησις und τέχνη auf. In seiner Betrachtung der φρόνησις zeichnet diese sich nach Heidegger im Vergleich zur τέχνη dadurch aus, dass „die τέχνη das ἔργον nicht hat, die φρόνησις dagegen wohl" (Heidegger 1992: 53).[7] Diese Unterscheidung ist so gesehen nur eine andere Formulierung des Unterschiedes zwischen πρᾶξις und ποίησις: Diese besitzt ein äußerliches Ziel, während jene sich selbst zum Ziel setzt. Gleichwohl ist auch zu beachten, dass Volpi die πρᾶξις als Bewegung von der φρόνησις als Wissensform ausdrücklich unterscheidet. (Vgl. Volpi 2007a: 176) Die φρόνησις gilt als eine Art des Sehens, eine praktische Sichtweise und somit auch als ein anfängliches Moment der πρᾶξις. Durch das oben genannte Beispiel von Volpi wird es klar, dass eine bestimmte Tätigkeit sowohl als πρᾶξις als auch als ποίησις verstanden werden kann, je nachdem, ob die Vollzugsweise dieser Tätigkeit φρόνησις oder τέχνη ist. Somit ist zu

6 Es wird auch in der Aristoteles-Forschung betont: „Aristoteles hat erkannt, daß Herstellen und Handeln nicht disjunkte Tätigkeitsklassen unter sich befassen, sondern daß sie unterschiedliche Aspekte an Tätigkeiten auszeichnen." Ebert 1976: 29.

7 Zur ausführlichen Aufrollung des Unterschiedes zwischen τέχνη und φρόνησις siehe Volpi 2007c: 229–231.

schließen, dass die φρόνησις bei der Umwendung vom Was zum Wie bei Aristoteles eine zentrale Rolle spielt.

Jedoch ist zu beachten, dass Heidegger Volpis Behauptung, dass es bei Aristoteles die Ontologisierung bereits gebe, wohl nicht zustimmen würde. Nach Heidegger gilt die Ontologisierung nicht als die Umwendung von der konkreten Tätigkeit zur Verhaltensweise oder Sichtweise, sondern als die von der Verhaltensweise des Menschen zur Seinsweise des Daseins. Durch die Ontologisierung, welche eine Radikalisierung der aristotelischen Transformation vom Was zum Wie darstellt, wird das Inhaltliche völlig aufgehoben, wodurch das rein Formale zum Vorschein kommt. Im Vergleich zur πρᾶξις lässt sich das Dasein formal anzeigen, insofern das Sein des Daseins dadurch bestimmt ist, dass es ihm „in seinem Sein wesenhaft *um* dieses Sein selbst geht" (Heidegger 1977: 113). Bei der Transformation von der konkreten Verhaltensweise zur formalen Seinsweise spielt sich eine große Umwandlung ab, die den erheblichen Unterschied zwischen beiden Philosophen bekundet: „Wenn man dieses [Heideggers] Verständnis des menschlichen Seins dem aristotelischen Verständnis des menschlichen Lebens gegenüberstellt, dann fällt sofort auf, wie verschieden die beiden Konzeptionen in ihrer grundsätzlichen Ausrichtung sind." (Rese 2007: 181)

Es lässt sich doch noch fragen, wie Heidegger Aristoteles folgend ihn ontologisieren und radikalisieren kann. Tatsächlich handelt es sich beim Zauberwort der Ontologisierung keineswegs um eine vorhandene Antwort auf die Problematik des Verhältnisses zwischen Heidegger und Aristoteles, sondern es bringt diese Problematik ausschließlich zum Vorschein. Um sie zu lösen, ist es erforderlich, auf die methodische Grundlage von Heideggers Philosophie zurückzugehen. Das wird auch von Volpi vorgeschlagen, wenn er die formale Anzeige als die Weise der ontologischen Zuspitzung hinnimmt. (Vgl. Volpi 2007a: 173) Die Formalisierung kristallisiert das Formale aus dem Konkreten heraus und entfernt sich dann von diesem. Von ihr differenziert sich die formal anzeigende Methodik, die Heidegger in seinen frühen Vorlesungen sehr betont.[8] Das Formale in der formalen Anzeige ist Heidegger gemäß im Grunde ein anderes, ja ursprünglicheres als das in der Formalisierung. (Vgl. Heidegger 1995: 59, 63) Die Formalisierung als eine Art von Verallgemeinerung setzt sich das Formale zum Ziel, während dieses bei der formalen Anzeige den Ausgangspunkt der Konkretisierung ausmacht, der für den konkreten Vollzug wegweisend ist. In diesem Sinne wird durch die formale Anzeige die Grundhaltung der Philosophie von der Verallgemeinerung zur Konkretisierung transformiert, sodass sich das Allgemeine mit dem Konkreten verbinden kann. Die Ontologisierung des Praxisbegriffs in seiner formalen Bestimmtheit ist sodann nicht darauf angelegt, das Konkrete zu beseitigen, sondern eher gilt das Formale als etwas Anfängliches, dessen Konkretisierung als die Verwirklichung des Seins im Seienden zu verstehen ist.

Es ist zu erwähnen, dass sich diese Methodik schon bei Aristoteles aufspüren lässt. Volpi verweist darauf, dass der Begriff der φρόνησις nicht nur irgendeine „logoshafte Haltung" ausmacht, sondern „etwas mehr" beinhaltet: „Indem Aristoteles

8 O. Pöggeler bezeichnet die frühe Philosophie Heideggers als „formal anzeigende Hermeneutik", um damit den alten Namen *der phänomenologischen Hermeneutik* zu ergänzen. Pöggeler 1990: 119.

die bis dahin gelieferte Definition der Phronesis merkwürdigerweise für ungenügend erklärt, zugleich jedoch nicht präzisiert, was sie denn genauer sei, fordert er den Leser gleichsam auf, dem nachzugehen, was sie ‚noch mehr' über das von ihm selbst Ausgesagte hinaus sei." (Volpi 2007a: 178) Daraus geht hervor, dass die φρόνησις nicht die vollständige Bestimmung der Praxis ausmacht, sondern nur eine Richtung anzeigt, deren Fortgang jedoch vom Vollzug jedes Einzelnen abhängig ist.

Heidegger zufolge sei die φρόνησις „das *Erblicken des Diesmaligen*, der konkreten Diesmaligkeit der augenblicklichen Lage" (Heidegger 1992: 163). Es ist darauf hinzuweisen, dass sich die φρόνησις durch Heideggers Interpretation nicht nur als praktisch, sondern auch als ontologisch erweist. Sie betrifft nämlich „eine Doppelung der Hinsicht, in die der Mensch und das Sein des Lebens gestellt sind" (Heidegger 2005: 385). Diese Doppelung, die zugleich auf das Konkrete und Allgemeine Rücksicht nimmt, wird erst verwirklicht, indem die formal anzeigende Methodik darauf Anwendung findet; eine Methode, die das Formale und dessen Konkretisierung in den Mittelpunkt rückt. Zusammenfassend lässt sich sagen, dass Aristoteles durch den Begriff der φρόνησις den Charakter der formal anzeigenden Methode in gewisser Wiese schon angedeutet hat, obgleich seine praktische Philosophie Heidegger zufolge aufgrund ihrer Ausrichtung auf das konkrete Seiende im Prinzip ontisch ist und die ontologisierte Umformung erfahren soll.[9] Diese kommt durch die formal anzeigende Methodik zustande, deren Kernpunkt darin liegt, dass das Allgemeine, Formale und Leere als Ausgang nur durch die Konkretisierung zur Entfaltung kommt. Aus der formal anzeigenden Perspektive sind die philosophischen Begriffe beweglich und vollzugsmäßig, sie entfernen sich somit von einem am Gehalt orientierten Ordnungszusammenhang und eröffnen ihrerseits einen neuen Raum, in dem sich das Sein selbst in der konkreten Praxis erscheint. Dadurch sieht sich

9 Hier gilt es, darauf zu verweisen, dass Heideggers Interpretation von Kants Ethik bei der Ontologisierung der aristotelischen praktischen Philosophie eine bedeutsame Rolle spielt. Die kritische Stellungnahme des frühen Heidegger zu Kant erfuhr in den 1920er Jahren plötzlich eine Umwandlung, die sich in einem Brief an K. Jaspers zeigte: „Das Schönste aber, ich fange an, Kant wirklich zu lieben." Biemel/Saner 1990: 57. Diese positive Einstellung hat sich danach vergrößert, was sich in Heideggers Analyse von der formalen Ethik Kants zeigt. In einer Vorlesung aus dem SS 1930 betont Heidegger: „Das Formale ist nicht das unbestimmte Leere, sondern gerade das ‚Bestimmende' (forma, εἶδος)." Heidegger 1982: 279. In diesem Sinne bedeutet Form des Gesetzes „dasjenige, was am Gesetz, an der Regelung, am Ursachesein das Bestimmende, das Eigentliche und Entscheidende ausmacht. [...] Dieses, sich selbst zu wollen, ist das angebliche Leere, im Grunde aber ist es das einzig Konkrete und Konkreteste an der Gesetzlichkeit des sittlichen Handelns." Heidegger 1982: 279f. G. Figal hat über den Auffassungsunterschied der Freiheit zwischen Kant und Heidegger geäußert, dass „Heideggers Begriff der Freiheit ähnlich wie derjenige Kants einen Aspekt hat, den man nun freilich nicht mehr ‚praktisch' nennen darf." Figal 2000: 133. Figal hat zwar die Ähnlichkeit zwischen den beiden Philosophen gesehen, aber den Grund dieser Ähnlichkeit vernachlässigt: Die Beiden sehen das Formale als das einzig Konkrete an. In Heideggers Lektüre hat er das Formale bei Kant nicht als das Abgezogene, das Ende der Verallgemeinerung, erachtet, sondern als einen anfänglichen Augenblick, der sich in einer konkreten Situation erfüllen wird. Dieser Deutungsversuch verdankt sich der Anwendung der Methodik der formalen Anzeige. Auf diese Weise ist Heidegger in der Lage, den ontologisierten vortheoretischen Praxisbegriff zu entwickeln.

Heidegger den aristotelischen Weg einschlagend und zugleich den Vorgänger mit seinen eigenen Denkbemühungen vorschiebend.

Die Zweifaltigkeit der Praxis und der vortheoretische Praxisbegriff

Ausgeführt wurde: Um die ursprüngliche Bedeutung des Praxisbegriffs bei Aristoteles, die nach Heidegger in der philosophisch geschichtlichen Entwicklung verloren geht, wiederaufzunehmen und zu bewahren, unternimmt er den Versuch, den Praxisbegriff zu ontologisieren. Durch die Enthüllung der Verbundenheit der Ontologisierung mit der formalen Anzeige steht fest, dass Ontologisierung nicht als eine Art von Theoretisierung oder Verallgemeinerung zu verstehen ist, sondern eher als Konkretisierung bzw. konkreter Vollzug des formalen Ansatzes, wie es in der Unterscheidung zwischen Formalisierung und formaler Anzeige zu sehen ist. Eben dadurch ist die Konzeption des vortheoretischen Praxisbegriffs möglich.

In diesem Zusammenhang ist es aber nötig, den Rahmen der Dreiteilung zwischen θεωρία, ποίησις und πρᾶξις näher zu klären. Durch die Behandlung der Untersuchungen von Prauss und Gethmann ist deutlich geworden, dass Theorie und Praxis als zusammengehörig zu denken sind. Jedoch erweist sich die hier genannte „Praxis" als Poiesis, also als die uneigentliche herstellende Praxis. Darum ist das, was von den beiden Forschern aufgedeckt wurde, eher die Einheit von Theorie und Poiesis. (Vgl. Volpi 1989: 232; Taminiaux 1991: 119–121) Außerdem wird die Praxis durch die Entdeckung der Differenz zwischen Praxis auf der einen Seite und Theorie und Poiesis auf der anderen Seite ihren eigentümlichen Ort finden. So gesehen lässt sich die These Volpis vereinfacht folgendermaßen zum Ausdruck bringen: 1) Die Praxis entspricht dem Dasein, während die Einheit von Theorie und Poiesis doch mit dem nichtdaseinmäßigen Seienden in Einklang steht; 2) Die Praxis hat einen Vorrang gegenüber Theorie und Poiesis, indem das Dasein dem nichtdaseinmäßigen Seienden (Vorhandensein, Zuhandensein) zugrunde liegt.

Durch die vorliegende Erörterung erscheint es so, als ob auch Volpi in die gewöhnliche Deutungsweise, Heidegger im Lichte der Dichotomie zwischen Praxis und Theorie zu interpretieren, geraten würde. Diese Vermutung erweist sich als unzutreffend, sofern die von Volpi genannte „Praxis", die sich im Grunde von der herstellenden Praxis, welche die implizite Quelle der Theorie ausmacht, unterscheidet, nicht theoretisiert ist. Hierbei kommt die Zweifaltigkeit der Praxis zum Vorschein: Zum einen die uneigentliche herstellende Praxis und zum anderen die eigentliche vortheoretische Praxis.[10] Zusammenfassend lässt sich sagen, dass Heideggers Umwendung von der Theorie zur Praxis sich nicht innerhalb der tradierten Dichotomie zwischen Theorie und Praxis vollzieht, sondern nur dann, wenn vor allem die Einheit zwischen Theorie und Poiesis (der uneigentlichen Praxis) gesehen wird, um danach die ursprüngliche, vortheoretische Praxis zu gewinnen.

10 Diese Zweifaltigkeit der Praxis verkörpert sich ebenfalls in Heideggers Lektüre des Verhältnisses zwischen φρόνησις und τέχνη. Darin tritt eine Schwierigkeit hervor: Woher stammt die τέχνη, wenn ihr Werk, das die τέχνη leitet, jedoch nicht in der τέχνη selbst beibehalten ist? Aus der Behauptung von Aristoteles, dass die ἀρετή der τέχνη weniger φρόνησις als σοφία ist, geht hervor, dass die Praxis (φρόνησις) gründlich unterschieden ist von der Poiesis (τέχνη), die aber mit der Theorie (σοφία) eine Einheit bildet. Vgl. Bernasconi 1989: 136ff.

Folgen wir dem Hinweis von R. Bernasconi, dass die Dreiteilung von „Bauen - Wohnen - Denken" der aristotelischen Dreiteilung der menschlichen Tätigkeit in „Poiesis - Praxis - Theorie" entspricht (Vgl. Bernasconi 1986: 123f.), dann ist klar, dass unsere Behauptung im folgenden Zitat von Heidegger ihre Bestätigung findet:

> Das Wohnen aber ist der Grundzug des Seins, demgemäß die Sterblichen sind. Vielleicht kommt durch diesen Versuch, dem Wohnen und Bauen nachzudenken, um einiges deutlicher ans Licht, daß das Bauen in das Wohnen gehört und wie es von ihm sein Wesen empfängt. [...] Daß jedoch das Denken selbst in demselben Sinn wie das Bauen, nur auf eine andere Weise, in das Wohnen gehört, mag der hier versuchte Denkweg bezeugen. (Heidegger 2000b: 163)

Literatur:

Apel, Karl-Otto (1973), „Wittgenstein und Heidegger. Die Frage nach dem Sinn von Sein und der Sinnlosigkeitsverdacht gegen alle Metaphysik", in: *Transformation der Philosophie.* Bd. 1: Sprachanalytik, Semiotik, Hermeneutik, Frankfurt am Main: Suhrkamp, 225–275.

Aristoteles (1894), *Ethica Nicomachea,* Oxford: Clarendon.

Backman, Jussi (2007), „Für das Wohnen denken. Heidegger, Arendt und die praktische Besinnung", in: *Heidegger-Jahrbuch.* Bd.3, Freiburg/München: Alber, 199–220.

Bernasconi, Robert (1986), „The Face of Distinction between Praxis and Poiesis", in: *Heidegger Studies* 2, 111-139.

Bernasconi, Robert (1989), „Heidegger's Destruction of Phronesis", in: *The Southern Journal of Philosophy* 28. Supplement, 127–147.

Biemel, Walter / Saner, Hans (Hg.) (1990), *Martin Heidegger/ Karl Jaspers. Briefwechsel: 1920 - 1963,* Frankfurt/ München/ Zürich: Klostermann.

Ebert, Theodor (1976), „Praxis und Poiesis. Zu einer handlungstheoretischen Unterscheidung des Aristoteles", in: *Zeitschrift für philosophische Forschung.* Bd. 30, 12–30.

Elm, Ralf (2007), „Aristoteles - ein Hermeneutiker der Faktizität? Aristoteles' Differenzierung von φρόνησις und σοφία und ihre Transformation bei Heidegger", in: *Heidegger Jahrbuch*. Bd. 3, Freiburg/München: Alber, 255–282.

Figal, Günter (2000), *Martin Heidegger. Phänomenologie der Freiheit,* Weinheim: Beltz Athenäum.

Gadamer, Hans-Georg (1972), „Hermeneutik als praktische Philosophie", in: Manfred Riedel (Hg.), *Rehabilitierung der praktischen Philosophie*. Bd.1, Freiburg: Rombach Verlag, 325–344.

Gethmann, Carl Friedrich (1993), *Erkennen und Handeln. Heidegger im phänomenologischen Kontext*, Berlin/New York: de Gruyter.

Heidegger, Martin (1977), *Sein und Zeit*, ed. Friedrich-Wilhelm von Herrmann, Frankfurt am Main: Klostermann.

Heidegger, Martin (1982), *Vom Wesen der Menschlichen Freiheit. Einleitung in die Philosophie*, ed. Hartmut Tietjen, Frankfurt am Main: Klostermann.

Heidegger, Martin (1992), *Platon. Sophistes*, ed. Ingeborg Schüßler, Frankfurt am Main: Klostermann.

Heidegger, Martin (1995), *Phänomenologie des religiösen Lebens*, ed. Matthias Jung, Thomas Regehly, Claudius Strube, Frankfurt am Main: Klostermann.

Heidegger, Martin (1999), *Zur Bestimmung der Philosophie*, ed. Bernd Heimbüchel, Frankfurt am Main: Klostermann.

Heidegger, Martin (2000a), *Reden und andere Zeugnisse*, ed. Hermann Heidegger, Frankfurt am Main: Klostermann.

Heidegger, Martin (2000b), *Vorträge und Aufsätze*, ed. Friedrich-Wilhelm von Herrmann, Frankfurt am Main: Klostermann.
Heidegger, Martin (2005), *Phänomenologische Interpretationen Ausgewählter Abhandlungen des Aristoteles zur Ontologie und Logik*, ed. Günther Neumann, Frankfurt am Main: Klostermann.
Pöggeler, Otto (1990), „Die Philosophie und die Problematik der Interpretation", in: Margot Fleischer (Hg.), *Philosophie des 20. Jahrhunderts. Eine Einführung,* Darmstadt: Primus, 117–136.
Prauss, Gerold (1977), *Erkennen und Handeln in Heideggers „Sein und Zeit"*, Freiburg/München: Alber.
Prauss, Gerold (1988), „Heidegger und die praktische Philosophie", in: Otto Pöggeler/Annemarie Gethmann-Siefert (Hg.), *Heidegger und die praktische Philosophie*, Frankfurt am Main: Suhrkamp, 177–190.
Rese, Friederike (2007), „Handlungsbestimmung vs. Seinsverständnis. Zur Verschiedenheit von Aristoteles' *Nikomachischer Ethik* und Heideggers *Sein und Zeit*", in: *Heidegger Jahrbuch*. Bd. 3, Freiburg/München: Alber, 170–198.
Ritter, Joachim / Gründer, Karlfried (Hg.) (1989), *Historisches Wörterbuch der Philosophie.* Bd. 7, Darmstadt: Wissenschaftliche Buchgesellschaft.
Schmidt, Alfred (1973), „Praxis", in: *Handbuch Philosophischer Grundbegriffe*. Studienausgabe. Bd. 4, München: Kösel, 1107–1138.
Sitter, Beat (1975), *Dasein und Ethik. Zu einer ethischen Theorie der Eksistenz*, Freiburg/München: Alber.
Taminiaux, Jacques (1991), *Heidegger and the Project of fundamental ontology*, Albany: SUNY.
Tugendhat, Ernst (1979), *Selbstbewußtsein und Selbstbestimmung. Sprachanalytische Interpretationen*, Frankfurt am Main: Suhrkamp.
Vetter, Helmuth (2007), „Heideggers Destruktion der Tradition am Beispiel des Aristoteles", in: *Heidegger Jahrbuch*. Bd. 3, Freiburg/München: Alber, 77–95.
Volpi, Franco (1989), „Sein und Zeit: Homologien zur Nikomachischen Ethik?", in: *Philosophisches Jahrbuch.* Bd.96, 225–240.
Volpi, Franco (2007a), „Das ist das Gewissen! Heidegger interpretiert die Phronesis (*Ethica Nicomachea* VI, 5)", in: Michael Steinmann (Hg.), *Heidegger und die Griechen,* Frankfurt am Main: Klostermann, 165–180.
Volpi, Franco (2007b), „Der Status der Existenzialen Analytik", in: Thomas Rentsch (Hg.), *Martin Heidegger: Sein und Zeit*, Berlin: Akademic Verlag, 29–50.
Volpi, Franco (2007c), „Heidegger und der Neoaristotelismus", in: *Heidegger Jahrbuch*. Bd. 3, Freiburg/München: Alber, 221–236.

Hondžijan Vang

Ontologizovanje prakse i preteorijski pojam prakse. Dalja razmatranja o problemu Hajdegerove praktične filozofije

Apstrakt

Iako u istraživanju nije retko to da se Hajdegerove misli posmatraju u svetlu odnosa teorije i prakse, sam Hajdeger zauzima stav protivan ovoj dihotomiji. Da bi se razumela Hajdegerova kritika ove dihotomije i njegova koncepcija izvornog jedinstva između teorije i prakse, mora se vratiti na njegovu lektiru aristotelove praktične filozofije. Pokušaće se pokazati da Hajdegerovo ontologizovanje prakse ima za cilj da izradi preteorijski pojam prakse, koji se razlikuje od jedinstva teorije i poezisa. Ontologizovanje nastaje metodom formalne naznake, koja u ovom smislu stoji u središtu Hajdegerove filozofije.

Ključne reči: praktična filozofije, ontologizovanje, preteorijsko, formalna naznaka.

Ontologization and Pre-theoretical Concept of Practice. Further Reflections on the Problem of Heidegger's Practical Philosophy

Abstract

Although it is not rare in research to consider Heidegger's thoughts in the light of the relationship between theory and practice, Heidegger behaves himself against this dichotomy. In order to understand Heidegger's criticism of this dichotomy and his conception of the original unity between theory and practice, we must trace back to Heidegger's reading of Aristotle's practical philosophy. It is attempted to show that Heidegger's ontologization of practice aims at elaborating the pre-theoretical concept of practice, which is in contrast to the unity of theory and poiesis. The ontologization is realized through the method of formal indication, which in this sense is at the center of Heidegger's philosophy.

Keywords: practical philosophy, ontologization, pre-theoretical, formal indication

UDK: 316.2 : 14 Foucault M.
https://doi.org/10.2298/FID1804533V
Original Scientific Article
Received: 15.02.2018. Accepted: 26.06.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Marija Velinov

INTEGRATED FOUCAULT: ANOTHER LOOK AT DISCOURSE AND POWER

ABSTRACT

This paper argues that there is continuity in Foucault's thought, as opposed to the common division of his work into three phases, each marking a distinct field of research – discourse, power, subject. The idea is that there are no radical turns in his work that justify this division; rather, there is a shift of focus: all crucial concepts are present (more or less [in]explicitly) in all periods of his thought and in all of his undoubtedly differently-toned and oriented works. This is shown through examining the characteristics of archaeology and genealogy, their relation, as well as the relation of discursive practices and strategies of power to knowledge. The retrospective and (re)interpretation intend to shed light on the constant interplay between concepts that demonstrate continuity in Foucault's thought. The viewpoint, based in the integrity of Foucault's work, offers a better starting point for understanding certain aspects of his theories.

KEYWORDS

subject, knowledge, archaeology, genealogy, problematising, continuity

Introduction

One can argue that one of the main features of Foucault's work is constant self-reflection – on numerous occasions he analyses and re-analyses the subjects and methods of his work. The question remains, however, as to whether Foucault had significantly changed his mind over time, or just his focus, i.e. the perspective from which he approached the same subjects.

Foucault's philosophy is commonly divided into three phases, each one dealing with a specific field of research: discourse, power, subject. However, rather than deeming this change in orientation a change in the research field, it seems more plausible to say that Foucault is shedding new light on certain aspects of his previous examinations. One can argue that these three phases are interwoven and their subject-matters present since the very beginning, although emphasized only later. Foucault himself, while talking about his works *Madness and Civilization*, *The Order of Things*, and *Discipline and Punish* in his final interview (Foucault 1990: 243) openly says that some matters have been left inexplicit, because of the way in which he posed problems, in which he approached matters. For example, Foucault clearly states the three main problems he tried to outline: the problem of truth, the problem of power, and the problem of individual behaviour (government of the

Marija Velinov: Research Assistant, Institute for Philosophy and Social Theory, University of Belgrade; marija.velinov@instifdt.bg.ac.rs.

Project cycle of Ministry of Education, Science and Technological Development of the Republic of Serbia "Politics of Social Memory and National Identity: Regional and European Context, Project Number: 179049".

self). He considers them three realms of experience that are properly viewed only in light of their mutual relations, not independently. Foucault thus believes that the shortcoming of the first two phases of his thought and the "shift" in his philosophy are due to the lack of a proper study of the government of the self; nevertheless, that does not stop him from studying truth or power.

With this in mind, can we speak of an unbreakable – or at least an inexplicit – continuity in Foucault's thought, despite the "turns" and "phases" which scholars ascribe to him for analytic reasons? Is it necessary, and appropriate, to speak of the "early and late" Foucault, and thus be left without the segment of his thought one cannot find in discourse, power, or the subject alone, but only in their mutual relations?

The issue is not whether Foucault had changed, because he certainly did not constantly repeat himself. The issue is whether Foucault created a sort of ambivalence and confusion which led to the (far too) sharp division of his generally consistent work. In regard to this, it might be interesting to return to frequently examined subjects of discourse and power, only now through the lens of the idea of a connection between these key concepts, contrary to the conventional division of Foucault's work. The reinterpretation is inspired by the fact that the connection, although Foucault had emphasized it, was either ignored or marginalized – we generally acknowledge it, but nevertheless wonder whether this was also said by the "early" or the "late" Foucault.

For that reason, the focus of this paper should move towards the shift from the first to the second phase of Foucault's work, giving us a better understanding of some characteristics of both archaeology and genealogy (and their mutual relation), the methods most closely associated with these two "periods" of Foucault's thought. The examination needs to invariably reflect on the specific relation between discursive practices and strategies of power towards knowledge, offering thus a (re) interpretation of Foucault's idea regarding the unbreakable connection between power and knowledge. Firstly, one should be reminded of some of the recurring motives and features of Foucault's work as a whole.

Problems I

> "[W]hat I did was designed to bring into the open a series of problems." (Foucault 1981: 103)

While reflecting on his work from the beginning to the first volume of *The History of Sexuality*, Foucault divides his works into books of exploration and book of method, and remarks that while *The History of Madness* and *Birth of the Clinic* are books of exploration, *The Archaeology of Knowledge*, *Discipline and Punish*, and *The Will to Knowledge* (the first volume of the *The History of Sexuality*) are books of method (where the method as such is explicated). The problem is the fact that Foucault does not apply a general method in every work, but constructs a new method of analysis depending on the object of investigation. (Foucault 1981: 28) The prime characteristic of Foucault's approach is that he likes to outline problems. At one point, Foucault even defined his entire work as the history of different problematics – of studying how and why something becomes a problem that needs to be addressed, in what

circumstances and what processes are involved. As he admits, his intention was to circumscribe certain types of problems through his work. (Foucault 1990: 251)

Foucault holds problematising to be a general objective: „I believe that problematising and constant re-problematising what needs to be done. If the effort of thinking has any purpose [...] then it is to radically problematise how we perceive our actions (our sexual activity, punitive practices, attitude toward madness, etc.)." (Foucault 1994a: 612) One should problematise what is believed to be self-evident, what is most self-evident: „It means allegiance to an idea that man is a being that contemplates even its most tacit actions. [...] The effort of thinking does not reveal evil that secretly resides in everything that is, but senses a danger that hides in everything common – what seemed certain becomes problematic." (Foucault 1994a: 612) To question what seems certain, the least problematic, the most evident – that is what Foucault considers the task of thinking.

That being so, the aim of circumscribing certain types of problems is not their solution, but to grasp them in their entirety, to gather them together. That is why his books end by bringing out a certain type of problem. (Foucault 1990: 251) For example, at the end of *Discipline and Punish* there is no answer to the question of the intended and actual function of prison, but a call to explore the power of normalization and the formation of knowledge in modern society. (Foucault 1995: 308)

The task of philosophy, according to Foucault, is not simply describing "contemporaneity" and "ourselves today", but also attempting to understand how that which is, by grasping why and how it is, may no longer be what it is. (Foucault 1998c: 449–450) Foucault's explicit aim is to transform our relationship with the world which so far seemed familiar to us – to change our relationship with *knowledge*, to be precise. (Foucault 1981: 37) "[T]his game of truth and fiction [...] will permit us to see clearly what links us to our modernity and at the same time will make it appear modified to us. This experience that permits us to single out certain mechanisms [...] and at the same time to separate ourselves from them by perceiving them in a totally different form, must be one and the same experience. This procedure is central to all my work." (Foucault 1981: 37-38) But, of course, not only to Foucault's work. He generally regards the work of the intellectual as a possibility to make *that which is* appear as *that which might not be* (or might not be as it is). (Foucault 1998c: 450) As coming to the realisation that something that seems necessary might actually not be so. If we were to attempt to write a history of what seemed necessary, we would discover a network of contingencies; and once we realise that it is something "made" and try to grasp the mechanisms behind it, we have the possibility of unmaking it. (Foucault 1998b: 450) For the way things are ordered does not entail that it is necessary for them to be like that – it is just a contingency. It is also contingent that we question the truth of *some* things and deem *some* matters problematic, but not others.

Foucault calls this approach *thinking in term of the event*. It entails a break with the obvious. Instead of referring to the obvious which is imposed on us, when tempted to refer to it we should try to bring out some "peculiarity". "To show that it is not 'so inevitable'; that it was not obvious that madmen should be deemed mentally ill; that it was not obvious that the only thing one ought do with the offender is to lock him up; that it is not obvious to look for causes of illness in the individual examination of the body, etc. To break with the obvious in which our knowledge,

our consent, our practices are grounded." (Foucault 1994b: 23) Besides, thinking in term of the event reveals connections, encounters, supports, blockages, plays of forces, strategies and so on, which establish what counts as self-evident, universal, and necessary today. (Foucault 1994b: 23)

Practice

Contrary to common interpretation, Foucault's analyses are not aimed at theories or institutions, but practices, in an attempt to grasp the conditions that at a given moment make these practices acceptable. He focuses on practice regimes where what is said and what is done are linked together, on imposed rules and the justifications behind them. Practice regimes entail programming behaviour, prescribing what needs to be done and known. (Foucault 1994b: 22)

In other words, when Foucault examines certain practices, he attempts to grasp how forms of rationality inscribe themselves into practices, and their role within them. For example, the ritual of public torture is no more irrational that putting someone in a cell. It is irrational only in relation to a certain kind of justification that the carceral practice has found for itself. Punitive practices associated with the institution of prison find their justification in discourses that are perceived as true – gaining thereby the right and power to draw the line between true and false – in the dominant scientific discourses of the time: medicine, psychology, criminology. That way, rationality is inscribed into the carceral practice, and it becomes *rational* and *obvious* that imprisonment is the most appropriate punishment for all who violate the law. The establisment of prison as a means of punishment is a discontinuity in relation to previous practices; Foucault attempts to explain this change, to locate and expose *the transformation of „knowledge" that enabled this sudden shift*. (Foucault 1994b: 22) He is interested in how that practice, and every other practice, is inscribed into the field of what is self-evident. For something to even become a practice (in the sense of something constatly conducted) it must have an excuse – a certain regime of rationality as a guarantee that we are doing exactly what we should do.

The idea is, therefore, to examine the play between the law which prescribes forms of conduct and the constitution of *true* discourses which serve as the foundation, excuse, and reason for existence, as well as the principle that transforms these forms of conduct. The problem is discovering how people govern (themselves and others) by constituting the truth, by organizing the field in which the practice of true and false can be determined and adequate. (Foucault 1994b: 27) Thus the issue is always the link between power, discourse, and knowledge.

Archive

> "[W]hat is this specific existence that emerges from what is said and nowhere else?" (Foucault 2002: 31)

Foucault is an archivist, someone who collects archives of something spoken, of a speech. He analyses accumulated discourses in their archival form. (Foucault 1998b: 289–290) To him, speech is not a means for expressing thoughts, with thoughts being the object of attention. On the contrary, he views speech (discourse) as a system that

establishes statements as *events* with its own conditions and domains of appearance. (Foucault 2002: 145) Archive, in Foucault's view, is firstly the law of what can be said.

Speech involve rules; not grammar rules that govern it (which takes us back to analysing speech as a means for expressing thoughts), but the rules of its appearance and existence. Archaeology, being a description of archives, describes *discursive events*. Contrary to language analysis which asks according to what rules has a statement been made, as well as according to what rules can new statements be made, these descriptions pose the question: „how is it that one particular statement appeared rather than another?" (Foucault 2002: 30) Within archaeology, Foucault is interested in the set of rules that define (Foucault 1968: 859-860) in a given time and for a given society what can be said – what is established as a domain of discourse in a given period; which discourses are forgotten, which are remembered, which are censored; which are widely accepted, which are depreciated; what role do previous discourses play in modern society – what old-time discourses have we preserved or attempted to restore, and what kind of individuals and groups accept a certain type of discourse – the institutional nature of the relation between discourse and the people who say it or receive it, but also *how the struggle between different groups for taking over control of discourse is unfolding*. Thus one can see that in the very definition of his research field Foucault already implies that at least part of his analysis will deal with the issues of struggle and power.

Foucault is, therefore, interested in the *conditions for the possibility of the existence* of discourse. But not a discourse that could be spoken but is not – not conditions for a *possible* discourse, but only if a discourse exists or had existed. If the discourse had appeared, Foucault is interested in the conditions that enabled it. If there are things said, or put simply, if something is said, the immediate reason for them is not what is said in them, or the people who said it, but the system of discursivity and the possibilities, i.e. impossibilities it lays down. (Foucault 2002: 145)

Archaeology thus seeks to be the history of spoken speech. As such, it does not deal with what lies behind a discourse, but with speech itself. It is speech that links together, brings into relation, elements that are otherwise dispersed. Only when they merge does something become (e.g. madness) an object – only through speech as a practice do objects one speaks about appear. In Foucault's view, discourses do not speak about objects, but are practices which systematically form the objects of which they speak of and determine the subject's place within that speech. (Foucault 2002: 54) Foucault's interest in a specific aspect of discourse is supposed to shed light on its relation towards both the subject and object: only speech forms the objects and determines the place from which the subjects can speak of these objects.

With this in view, we can safely say that the archaeological analysis calls us to challenge the entire subject-object problematic which implies their separation. If while attempting to understand discourse we presume the creative activity of the sovereign subject or the transcendency of the assumed objective thing – that is, *words* or *things* – then this constitutive problematic is at risk of being subjected to *subjectivist* or *objectivist* reduction. In any case, what is reduced is the specificity of the instance of discourse. (Kozomara 1998: 101) This analysis is an attempt to go beyond words and things and lay the foundations for grasping things said (what is said).

Having in mind that *The Archaeology of Knowledge* presents only methodological specifications of already conducted research, a brief outline of the first phase

of Foucault's work may be as follows: *The History of Madness* deals with the emergence of a group of highly interwoven, complex objects. It was necessary to describe the formation of these objects so as to determine the specificity of the whole of the psychiatric discourse. On the other hand, the point of research of *The Birth of the Clinic* was not so much the formation of the object, but status, institutional siting, the situation, and the modes of insertion used by the discoursing subject. Finally, in *The Order of Things*, research is generally focused on networks of concepts and the rules of their formation. (Foucault 2002: 72) While writing *The Archaeology of Knowledge*, Foucault could not contribute with an analysis of function that a particular discourse must carry out in a field of non-discursive practices, although one can argue that to a certain extent the work *Discipline and Punish* (which supposedly belongs to the second phase of Foucault's work) does exactly that. That analysis shows that neither the process of appropriating discourse, nor its role among non-discursive practices, are extrinsic to the laws of its formation, but are precisely its formative elements. (Foucault 2002: 75)

Origin

> "History becomes 'effective' to the degree that it introduces discontinuity into our very being." (Foucault 1998: 380)

Genealogy is opposed to the search for "origins". This limitation, present in respect to archaeology as well,[1] seems at odds with the common usage of the word "genealogy" (in the sense of a study of origins and roots). Foucault, however, views origin that should be the subject of genealogy differently; he defines genealogy as analysis of origin in the sense of emergence – not as search for meaning or essence. The purpose of history, guided by genealogy, is surely not to discover "origins" or the roots of (our) identity, but the dissipation of identity; it does not reveal our unique homeland, but strives to make visible all discontinuities that cross us. (Foucault 1998a: 386–387) The emergence is always produced in a particular state of forces or power relations – an analysis of origin is supposed to discern the play of these forces. (Foucault 1998a: 377)

The idea of emergence becomes clearer if one explains the idea of *effective* history. Contrary to historical tradition, which views events as if they have continuity, effective history discerns what is unique in a particular event. An example of an event can be a discourse, or a change in how punishment is conducted. The event is thus a discontinuity in relation to previous punitive practices. Genealogy seeks to analyse the origin of this event viewed as emergence.[2]

1 Cf. „Discourse must not be referred to the distant presence of the origin, but treated as and when it occurs." (Foucault 2002: 28) The reason for this is an effort to distance ourselves from the idea that there is another discourse behind what is actually said, as if its meaning was previously conceived, prompting us to search for it; an effort not to analyse figuratively what is said (which is basically an effort not to assume a subject that would serve as an explanation) but to grasp how a particular statement and not some other occurred.

2 One should have in mind that an event should not be viewed as a decision, treaty, reign, or battle, but as the reversal of a relation of forces (as a shift, a domination that grows feeble, enabling the entry of a new one). (Foucault 1998a: 381)

One should stress that Foucault sees discontinuity as a sudden shift, rather than a completely isolated unit. It is not as if a bunch of discourses spontaneously emerged independently of one another. Discontinuity – discourse as an event – appears in virtue of changes which unfold somewhere else (whether at the foundation of a discourse or within some non-discursive practices with which a particular discourse is associated). Besides, if discontinuity does occur within punitive practices, i.e. if a new form of punishment occurs, that is caused by changes which need to be examined. Shifts, events, emergence are not something that happen by themselves. They are a change in reign, although not in the physical sense of replacing one state government with another (although this should not be excluded), but rather specific changes in the play of government (power) and knowledge which serves as the foundation of ourselves and our speech.

History in this view can be characterized as the "history of the present". (Gutting 2005: 10) One can say that one of the features of traditional history (whether aim or consequence) is showing that the current circumstances are inevitable. Contrary to this, Foucault's histories aim to show the contingencies of what history has given us. They aim to remove the "air of necessity" by showing that the past ordered things differently and that the processes leading to our current practices and institutions were in no way inevitable. (Gutting 2005: 10–11) The genealogic method represents the history of the present in the sense that it attempts to problematise the necessity of *dominant* categories and procedures. Foucault does not find a reliable viewpoint from which he could interpret everything prior to the present, but defines our present as the most problematic. The present is where we can free ourselves from ourselves, where we can question what we are. (Kozomara 2001: 104)

An example of the first "genealogically" oriented work by Foucault is *Discipline and Punish*, which primarily focuses on practices and institutions. However, it is disputable whether we can differentiate Foucault's methods, and associate these particular methods strictly with particular books which belong to a particular period of analyses. What is rather the case is that *all* of his books are to a certain degree both archaeology and genealogy, the history of concepts, the history of the present, and the history of problematisations. The only difference between them is their focal point, which largely depends on the topics discussed. With this in view, one can argue that all Foucault's histories are histories of the present, in the sense that the issues they deal with correspond to modern ideas and practices which Foucault for some reason considered to be especially dangerous.

(Dis)continuity

Foucault's switch from archaeology to genealogy is marked by an analysis of the relation between discourse and power, and indicated in his inaugural lecture at College de France in 1970 (a year after he published *The Archaeology of Knowledge*). (Foucault 1981)

There Foucault defines his analysis or, to be precise, the methodology of his future research as having both a *critical* and *genealogic* aspect. The critical aspect attempts to grasp the reversal happening at the level of discourse. As Foucault tells us, these reversals occur due to forms of exclusion, of limitation, and of appropriation

of discourse. He attempts to identify and grasp these forms, which control discourse. (Foucault 1981: 70–73)

On the other hand, the genealogical aspect is concerned with the emergence of discourses, showing how a series of discourses came to be formed. In this context, the emergence of discourse is seen as an event, as a discontinuity, produced by contingent circumstances. He attempts to describe where that event takes place, the domain of its contingent production, and the conditions for its appearance. What is new is that Foucault attempts to grasp the *formation* of discourse in its affirmative power, meaning the power to constitute domains of objects, in the sense that it can affirm or deny true or false propositions.[3] (Foucault 1981: 71–73)

These two tasks cannot be separated. As Foucault stresses, it is not that there are forms of exclusion, of limitation, and of appropriation of discourse on the one hand, and the spontaneous emergence of discourse (which is only later put under control) on the other. The very emergence of discourse entails mechanisms of control. What separates the critical and genealogical enterprise is thus not so much a difference of object or domain, but rather of point of attack and perspective. (Foucault 1981: 71–73)

Power

> "[I]t seems to me now that the notion of repression is quite inadequate for capturing what is precisely the productive aspect of power." (Foucault 1980: 119)

Apart from altering the meaning of genealogy and thus abandoning the search for origin, the great theorist of power Foucault seemingly paradoxically holds the view that "power" does not exist. In this aspect, Foucault is a nominalist, excluding the existence of universal categories that encompass our experience, leaving us only with particular entities which constitute and exhaust a specific category. That means that there is no "power" as such, but only individual instances of domination, control,

3 In his inaugural lecture, Foucault speaks of the will for truth. Simply put, the will for truth is desire to know the truth about certain matters, to divide statements about objects into true and false, to privilege some discourses – or to be precise, some subjects within discourses – assigning them the right to tell the truth. However, the will for truth entails a desire for more truth, for more knowledge. It is an uncontrollable will for ever-more knowledge. For example, the will for knowledge pressures and coerces various disciplines to look for a rational foundation in true discourse. The punitive system firstly sought its foundations in law, switching afterwards to medical and psychological discourse. The law has its authority only in virtue of its connection to true discourse – discourse which at a given moment has power over the distinction between true and false. The will – the desire to seek the truth about matters – brings about a specific relation towards power or in Habermas' words: power is ironically hidden within it. One can argue that all discourses have their origin in some kind of power. This is why Foucault switched from archaeology of knowledge to the genealogical analysis of the origin of discourse. While archaeology of knowledge reconstructs the rules that bring about some discourse, the role of genealogy is to shed light on the "discontinued" series of orders that impose upon man a particular interpretation of the world. Genealogy studies origins (in its specific sense) of discursive products which, formed as orders of discourses, represent power practices which are entwined with one another in the game of overpowering. Cf. Habermas 1990a, 1990b

edification, and similar. (Gutting 2005: 40) "Power" exhausts itself in relations – there are only concrete practices or *strategies of power*. Foucault does not imply a subject who has power which he may or may not exercise over others, but power relations which entwine those who exercise power, as well as those over whom power is exercised.[4] He is interested in the very act of exercising power, how one acts upon the actions of others, or to be precise, how a group of subjects act upon the (present or future) actions of others. (Foucault 1983: 219-220) Power exists only in this act. For a system of power to function, it needs to be exercised over free people, the subjects of action. Accordingly, we can say that power does not entail violence over someone – it does not destroy, but produces actions (as well as practices, speech, knowledge, even identity) as reactions to power.

From this perspective, Foucault persistently attempts to demonstrate that there is nothing necessary about accepted categories and practices. All categories, truths, and knowledge which we deem obvious, owe their self-evidence to a system of power which is not something unchangeable. A history of necessities would unveil a web of contingencies which lie in its foundation. Genealogy is precisely the history of necessities; the web of contingencies is power-government. If one grasps how necessity or self-evidence is established – the strategies of power that guide it – one gains the power to disassemble it to its foundation. That way, we do not only lose self-evidence (the obvious, the necessary) but pave the way for different opinions and actions.

Truth

What is the relation between knowledge and power? How does power inscribe itself into knowledge, and vice versa? The answers to these questions must follow the trajectory of the previous analysis, for it is only in respect to knowledge that we can clarify the relation between discursive practices and strategies of power. Having in mind that Foucault defined knowledge in relation to archaeology as a group of statements formed by discursive practices, it is clear that strategies of power are inscribed into knowledge through discursive practices. Strategies of power, characteristic for a certain period, determine the object of speech, how to speak about it, who has the right to speak, and so on.

For example, in a summary of his lectures at College de France from 1973 and 1974, published under the name *Psychiatric Power* (Foucault 2006), Foucault remarks that the power the doctor exercises over the patient – visible in the right to separate him from his family and friends and consign him to an isolated place – is associated with what is true. The doctor (in this case the psychiatrist) is someone entitled to produce the truth about illness.

4 This paper uses the term "power" because it is the common translation of the word *pouvoir*. However, since Foucault does not use it as a verb (meaning "to be able to"), but as a noun, the term "government" seems to be a more accurate translation. If one bears in mind that power actually means government, it becomes clear that Foucault speaks of it both in respect to those who govern and to those who are governed. Contrary to power which is seen as something per se, as something which someone has and can use, but does not have to, government entails power over someone, i.e. that someone or something is in the power of another. Foucault is interested in this relation, not in power as such.

The psychiatrist is in the position of someone who knows more of the madman than the madman himself, giving him the right to label the patient's behaviour as illness – what is more, it is (the type of) illness of which the patient has no right over. Since science on which the doctor relies gives him the right to name what he sees as illness, this gives him the right to intervene. But (again with respect to knowledge) doctors consider him/her a very specific kind of patient – one that suffers from mental illness. And on that occasion, in that way, given the knowledge, competency and right of the doctor, hence in regard to specific power relations – psychiatry is established. It was this power game that established knowledge, which in turn laid the foundation for the rights of that power, which then paved the way for psychiatry. In these circumstances, the doctor turns out to be (or is established as) an official subject of knowledge, with knowledge "not [being] made for understanding ... [but] for cutting." (Foucault 1998a: 380) The power carried by knowledge is the power to determine, assess, define, and classify another person. (Marinković 2005: 246–248)

Science is undoubtedly undergoing twists and turns, breaks with previously formulated true proposition, with how one speaks or sees, with a whole ensemble of practices, but Foucault holds that these are not new discoveries, but a new regime in discourse and knowledge. He remarks that certain moments and certain orders of knowledge undergo shifts and transformations which are at odds with the "continuist" image of science, that there is discontinuity regarding how some matters are perceived or approached. However, Foucault concludes that these changes are only a sign of other, deeper changes: changes in the rules of formation of statements accepted as scientifically true. It is thus not a change of content, nor theoretical form, but change at the level of what governs statements and the way they govern each other so as to constitute a set of propositions which can be verified, i.e. undergo some sort of scientific procedure. That is what Foucault deems the problem of regime, the politics of scientific statement. When analysing changes within science, he is interested in the modifications of the regime of the power of knowledge which lies in its foundation. This set of problems – the problem of status, conditions, exercise, functioning, the institutionalisation of (scientific) discourse – although not emphasized enough, seems crucial for the first phase of Foucault's work. (Foucault 1980: 112–113)

Foucault therefore does not mean by truth an ensemble of verified propositions that need to be discovered and accepted, but a set of rules which draw the line between true and false, and according to which certain effects of power are associated with what is true. (Foucault 1980: 132) He is interested is finding how *effects of truth* are historically produced within discourses, which in themselves are neither true nor false. (Foucault 1980: 118) Truth as such is a system of ordered procedures for the production, regulation, distribution, circulation, and operation of statements. These procedures, these regulations, are why there is something like a battle for truth, as Foucault calls it. (Foucault 1980: 132–133) History is just a continuation of struggles, strategies, and tactics. (Foucault 1980: 114) These strategies, however, are not mere repressive techniques that limit us; they produce and form knowledge – they produce discourse. They are a productive network that pervades the whole of the social body. This network governs us by virtue of production. In

the words of Foucault: "if it [power] never did anything but to say no, do you really think one would be brought to obey it?" (Foucault 1980: 119)

Truth in this sense is not something external, nor without power. It is circularly associated with systems of power that produce and support it, as well as with the effects of power it produces and preserves and which renew it. Every society has its regime, its general politics of truth. This regime entails the acceptance and the functioning of certain discourses as true; it is a mechanism that enables true statements to differ from false. It is a political and institutional regime of the production of truth which statutorily gives official authority to those who are obliged to say what is true.

Problems II (instead of a conclusion)

If one were to single out one aspect of the relation between knowledge, power, truth, and discursive practices, perhaps it should be *problematisation* as such. As previously mentioned, in his later essays Foucault defines his previous work as a history of problematisations, i.e. of examining how and why something is determined as a problem that needs to be addressed, as well as the circumstances and processes involved. Having in mind that power inscribes itself into knowledge by means of discursive practices – which are systems of rules that determine what can become an object of knowledge – we can conclude that power is inscribed into knowledge through problematisation. Power exists where the question is posed.

If one is reminded that Foucault holds that problematisation is the purpose of thinking, it seems justified to argue that thinking as such opens the possibility for, if not overcoming an omnipresent system structured by means of discursive and practical techniques, then at least for evading its dominance. Only thinking, and not political practice, enables us to distance ourselves from the naturalised struggle for power, the endless circle of overcoming and resistance. The instance of thinking – problematising self-evident categories under which we and others fall – disables the total functioning of those systems and procedures of power within which we nevertheless reside.

References:

Flynn, Thomas (2005), "Foucault's Mapping of History", in: Gary Gutting (ed.), *The Cambridge Companion to Foucault*. New York: Cambridge University Press, pp. 29–48. 1968.

Foucault, Michel (1980), "Truth and Power", in: Colin Gordon (ed.), *Power/Knowledge: Selected Interviews and Other Writings: 1972–1977*. Pantheon Books, New York, pp. 109–133

Foucault, Michel (1981), "The Order of Discourse", in: Robert Young (ed.), *Untying the Text: A Post-Structuralist Reader*. Routledge & Kegan Paul, pp. 48–78.

Foucault, Michel (1981), *Remarks on Marx*. Semiotext(e). 1983.

Foucault, Michel (1990), "The Return of Morality", in: Lawrence D. Kritzman (ed.), *Politics, Philosophy, Culture, Interviews and Other Writings 1977–1984*. New York, London: Routledge, pp. 242–254.

Foucault, Michel (1995), *Discipline and Punish*. Vintage Books, New York.

Foucault, Michel (1998a), "Nietzsche, Genealogy, History", in: Paul Rabinow (ed.), *Aesthetics, Method, and Epistemology*. The New York Press, pp. 369–392

Foucault, Michel (1998b), "On the Ways of Writing History", in: Paul Rabinow (ed.), *Aesthetics, Method and Epistemology*. The New York Press, pp. 279–298.
Foucault, Michel (1998c), "Structuralism and Post-structuralism", in: Paul Rabinow (ed.), *Aesthetics, Method, and Epistemology*. The New York Press, pp. 433–458
Foucault, Michel (2002), *Archaeology of Knowledge*. Routledge Classics.
Foucault, Michel (2006), *Psychiatric Power*. Palgrave Macmillian.
Foucault, Michel (1983), "The Subject and Power", in: Hubert L. Dreyfus and Paul Rabinow (eds.), Michel Foucault: Beyond Structuralism and Hermeneutics, Chicago: The University of Chicago Press, pp. 209-226.
Foucault, Michel (1968), "Response a une question", Esprit 5, 850-874.
Foucault, Michel (1994a), "À propos de la généalogie de l'éthique : un aperçu du travail en cours", Dits et écrits (1954-1988), tome IV (1980-1988). Editions Gallimard, Paris. pp. 609-631.
Foucault, Michel (1994b), "La poussière et le nuage", Dits et écrits (1954-1988), tome IV (1980-1988). Editions Gallimard, Paris. pp. 10-19
Gutting, Gary (2005), "Introduction", in: Gary Gutting (ed.), *The Cambridge Companion to Foucault*. New York: Cambridge University Press, pp. 1–28.
Habermas, Jürgen (1990a), "Some Questions Concerning the Theory of Power: Foucault Again" in: Philosophical Discourse of Modernity. Polity Press. pp. 266-293.
Habermas, Jürgen (1990b), "The Critique of Reason as an Unmasking of the Human Sciences: Michel Foucault" in: The Philosophical Discourse of Modernity. Polity Press. pp. 238-265.
Kozomara, Mladen (1998), *Govor i subjektivnost*. Plato, Beograd.
Kozomara, Mladen (2001), *Subjektivnost i moć*. Plato, Beograd.
Marinković, Dušan (2005), "Nauka kao znanje/moć: od disciplinujuće do diskurzivne prakse", in: Pavle Milenković and Dušan Marinković (eds), *Mišel Fuko 1926-1984-2004 Hrestomatija*. Novi Sad: Vojvođanska sociološka asocijacija, pp. 243–252.

Marija Velinov

Integrisani Fuko: drugi pogled na diskurs i moć

Apstrakt

Nasuprot uobičajenoj podeli Fukoovog rada na tri faze, koje obeležavaju različita polja razmatranja – diskurs, moć, subjekat – u ovom radu se zastupa kontinuitet njegove misli. Ideja da nema naglih preokreta koji opravdavaju oštre podele, već radije promene težišta istraživanja, da su svi ključni pojmovi, manje ili više (ne)skriveni, prisutni u svim razdobljima Fukoovog mišljenja i u svim, nesumnjivo različito orijentisanim i intoniranim spisima, ilustruje se razmatranjem karakteristika arheologije i genealogije, te njihovog odnosa, kao i odnosa diskurzivnih praksi i strategija moći prema znanju. Ta retrospektiva i (re)interpretacija Fukoove misle valjalo bi da ukaže na neprestanu saigru pojmova koja svedoče o njenom kontinuitetu. Takvo stanovište, koje se zasniva na integritetu Fukoovog dela, može da ponudi i bolju osnovu za razumevanje njegovih posebnih aspekata.

Ključne reči: subjekat, znanje, arheologija, genealogija, problematizovanje, kontinuitet

UDK: 82 Celan P.
https://doi.org/10.2298/FID1804545B
Original Scientific Article
Received: 15.06.2018 Accepted: 20.08.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Petar Bojanić

PAUL CELAN ON THE IMPOSSIBILITY OF TESTIMONY: "*ORT MEINER EIGENEN HERKUNFT*"

ABSTRACT
In his poems, Paul Celan does not use words such as territory, border, border crossing, and only very rarely the word space. I would like to reconstruct the traces of "*Heimat*" in Celan (in a number of poems from different periods "Heimat" plays an important role), and perhaps try to describe what Heimat might have meant for the young Paul Antschel (his real name). That is to say, I would like to understand whether "Heimat" is synonymous with what Celan speaks about, many years after his name change, in the address given on the occasion of the Georg-Buechner-Preis: "*Ich suche auch, denn ich bin ja wieder da, wo ich begonnen habe, den Ort meiner eigenen Herkunft.*" In the poems written at the time when Antschel is learning Hebrew as well as reading Martin Buber (Israel Chalfen) for the first time, I look for some basic figures Celan ties to his life in Bukovina at the time, in the environment of Czernowitzer Judentums. Aside from the works by Israel Chalfen, Else Keren and Elke Guenzel, I would like to make use of a book published some ten years ago, a detailed listing of Celan's Paris library. I would like to consult this archive in the coming period, since Celan punctuated the margins of many of those books with evocations of his early creative period.

KEYWORDS
testimony, *Heimat*, home, birth, border

To what does Celan testify? Does he testify at all, and can he at all be a witness? How can we describe his activity and effort to construct a position from which he could potentially testify, that is, be a credible witness?

Allow me to hastily and preliminarily define Paul Celan's engagement (his poetry, thematization of his poetry in fragments, diaries, published speeches) as an attempt to deconstruct the protocol of testimony and the position of the witness – that is to say, the poet as witness and poem as testimony. Although there are difficulties and often rather unclear testimonies, I would be interested to reveal the epistemological value of Celan's testimony and perhaps his original contribution to the reconstruction of the concept or practice of testimony. What is it that Celan lets us know? What do we learn when Celan speaks or writes? What makes Celan's insistence on the reconstruction of the position of witness or testimony different from well-known stereotypes we associate with these protocols? I would like to delve

Petar Bojanić: Professor, Institute for Philosophy and Social Theory, University of Belgrade; bojanic@instifdt.bg.ac.rs.
This paper is a result of the project 43007 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

into Celan's position, use various, tentative darting probes, in order to list some of the essential characteristics or forms of testimony. A first option would refer to the narrativity of testimony. A given content is transferred from one place to another and it should be believed or not believed. When I speak of epistemological status of someone's testimony, what I have in mind is that testimony implies the possibility of becoming meaning (provided it is true), that it is grounded performatively, and therefore accepted by others in a social setting or group.[1] Paul Celan often neglects the narrative potential of his expression (as we often do not know what he is speaking of) for the sake of the importance of his own presence or "*un acte présent*" [a present act] or even presence of act without actor (poem without poet, act of poem writing without the poem itself).[2] Second, Celan is never third (*terstis* or *testis*).[3] He is ever on the move, on the road, in motion and retreat, and thus does not occupy the position of one who is present and who sees[4] (one cannot help but think of the Serbian word *očevidac* – literally one who sees directly with one's eyes; nor indeed of the English *eyewitness*, which might be different from simply *witness*). Celan also does not occupy the position of arbiter (*l'arbitre*) who, according to Benveniste, "also fulfills a testimonial function of the idea of seeing without having seen." Even when he is imagining a completely new association of all those exiled from their native countries [*Der Verband der Heimatvertriebenen*] ("An association of global exile still remains to be founded…" [*Der Verband der Weltvertriebenen wäre noch… ins Leben zu rufen*]), Celan does not assume either collective or individual testimony, but paradoxically a new, silent pseudo-homeland: "In their thoughts of who and what they are and how they are exiled – there is their homeland" [*Im Gedanken, dass und was und wie sie vertrieben wurden, ist die eigentliche Heimat*].[5] It is impossible to testify about one's homeland. Third, Celan never testifies about himself nor is his own witness (I am thinking of the paradigmatic

1 I am ignoring the cases of testimony in which the witness is guilty and revealed because he is a witness and because he is a survivor (if he has carried over knowledge, his function is completed).

2 "*L'essence du témoignage ne se réduit pas nécessairement à la narration, c'est-à-dire aux rapports descriptifs, informatifs, au savoir ou au récit; c'est d'abord un acte présent. Le martyr, quand il témoigne, ne raconte pas d'histoire, il s'offre*." [The essence of testimony does not necessarily reduce to narration, that is to say, to descriptive or informative reporting, to knowledge or to account; it is first of all a present act. The martyr, when he testifies, does not recount a story, he offers himself.] (Derrida 1996: 29). In another, yet similar, context, Emmanuel Levinas speaks of "*le témoignage pur*" [pure testimony], which "*ne thématise pas ce dont il est témoignage*" [does not thematize that to which it witnesses] (Levinas 1993: 220).

3 "*Etymologiquement* testis *est celui qui assiste en 'tiers'* (terstis) *à une affaire où deux personnages sont intéressés* (…)." [Etymologically, *testis* and one who assists as 'third' (*terstis*) where there are two concerned parties (…)]. (Benveniste 1980: II, 277)

4 "*Le* testis *est là au vu et au su des parties*." [The *testis* is present to sight and knowledge of the parties]. (Benveniste 1980: II, 174) The 1990s wars in the Balkans had examples where a man and husband clandestinely watches the rape of his own wife and mother, vicious murders of one's own children by criminals and later gives testimony regarding this. He sees, but is not seen, thus his testimony should not be considered credible.

5 Bertrand Badiou has pointed out to me (perhaps unconsciously) the slight shifts in versions of this sentence, mentioned several times in Celan's fragments. "*Der Verband der Weltvertriebenen wäre noch… ins Leben zu rufen*" (*Mikrolithen*, 46)." "*Der Verband der Weltvertriebenen wäre wohl noch… ins Leben zu rufen*" (*Meridian*, F 85, 6)." "*Der Verband der*

model found in the Gospel according to John) nor his own guarantor, as 'his own' quasi 'I', which always appears in the search for its homeland or in the return to the homeland ("Heimkehr"), is always transformed or a plurality of various 'I's,[6] or else a lost, imaginary 'I'.[7] In that sense, paradoxically, the status of one who is not at all able to testify convincingly, fulfills the famous rule of which speaks Pascal.[8]

In the following sentence from Celan's *Der Meridian*, "*Ort meiner Herkunft*" (The Place of my Origin), which I have used as the title of my paper, a lot is said; yet, I have added a potentially disturbing question by Celan: "*Heimat, und Ich?*".[9] Any conversation or mention of *Herkunft*, 'being at home' or 'at one's place' or 'being home', any mention of native land or homeland, necessarily poses and brings to surface this question, our own and about ourselves: "*Heimat, und Ich?*" [Homeland, and I? or Homeland, and me?]. The emphasis is certainly on the comma and break after the word *Heimat*. A question or narrative (everything usually begins in narration) of *Heimat* immediately poses the question of 'I' or 'me', of birth certificate and passport, of memories and time, a distant place, etc. The word 'identity' (a complicated and difficult word, often unnecessary and trivial, often in a trivial register), which usually harmonizes all these operations and which usually begins with the pronoun 'I' and question 'I?', implies two more protocols on which Paul Celan insists.[10] The first refers to feelings and often comes in the form of insecurity,

---

Weltvertriebenen wäre ja wohl erst ins Leben zu rufen" (*Mikrolithen*, 30)." "*Der Verband der Weltvertriebenen wäre ja wohl noch ins Leben zu rufen*" (*Mikrolithen*, 43)."

6 For example, the poet's homeland, as well as his identity changes from poem to poem.

7 In a 1960 letter to his friend in Bucharest, Celan wonders if it would not have been better to have stayed in his native land.

8 In a passage referring to the history of China, Pascal writes: "*Je ne crois que les histoires dont les témoins se feraient égorger*" [I only believe stories of witnesses whose throats have been slit] (Lafuma 822). There are no witnesses, no narrative, but there is hesitation and detour.

9 On the last, 160th page of Jean Améry's book *Jeinseits von Schuld und Sühne. Bewältigungsversuche eines* Überwältigten (1966) Celan notes "*Heimat, --- Und Ich ? Ich war nicht einmal \ zuhause, als ich daheim \ (zuhause) war*" (in Badiou's translation: "*Pays natal... Et moi ? Je n'étais même pas a la maison, quand j'étais chez moi (à la maison*)." (Celan 2004: 451, 459) We have here Celan reading Améry's essay "Wievel Heimat braucht der Mensch?" in this book, and encounters the sentence "*Ich war kein Ich mehr und lebte nicht in einem Wir*." Celan's question could have also been provoked by Améry's sentence "Die Heimat ist das Kindheits- und Jugendland." Leonard Olschner is one of the rare readers who takes on Celan's reading of Améry and assumes that the 1968 poem "Dein Heim" emerges from this (Olschner 2007: 24–25). "Dein Heim" opens with the verse "in wievel Häusern?". Still, perhaps this indication is already present with the double meaning of *zweihäusig*, which appears a few times in *Die Nimandsrose*.

10 A potential third protocol could be the one mentioned by Améry in the second part of the sentence "*Ich war kein Ich mehr und lebte night in einem Wir*." *Heimat* implies the existence of something more intimate than community as such – the existence of "*Wir*" and the belonging to that "*Wir*." Such a "*Wir*" is a fictitious first entity that implies a future thematization of *Heimat* or loss of *Heimat*. Namely, *Heimat*, as entirely different for all members of a community, becomes the completely specific integrative factor for all of us, making u scloser and more equal as members of the community. We are connected if and only if each belongs to their own, to their own native land, that is to say everyone is at home. For, the condition that we are all together is satisfied if each of us in his own place, in his real place and belongs to their original "*Wir*." One without a native land,

disquiet or inappropriateness (it is unseemly for someone to ask me or even that I ask myself whether I am and where I am at home, even though I myself am always asking myself that) – Celan's answer is: "homeland – and I? I was not even at home when I was at home (at that house)." The second refers to the return, returning, the road, dilly-dallying, the way we return and how we travel back. Returning[11] can often be connected to vertigo, nausea, vomiting (in Serbian, the words *vraćanje* [returning] and *po-vraćanje* [vomiting] are a good indication of the discomfort, but also the need for careful and very delicate travel back or around that happens in *Herkunft* or in Identity). It seems to me that Celan's response about the simultaneous arrival and return, about the road, is quite clear:

> The Landscape from which I – by what detours! But are there such things: detours? – the landscape from which I come to you might be unfamiliar to most of you. It is the landscape that was home to a not inconsiderable portion of those Hasidic tales that Martin Buber has retold for us all in German. It was, if I may add to this topographic sketch something that appears before my eyes now from very far away – it was a region in which human beings and books used to live (*Die Landschaft, aus der ich – auf welchen Umwegen! Aber gibt es das denn: Umwege? -, die Landschaft, aus der ich zu Ihnen komme, durfte den meisten von Ihnen unbekant sein. Es ist die Landschaft, in der ein nicht unbeträchtlicher Teil jener chassidischen Geschichte zu Hause war, die Martin Buber uns allen auf deutsch wiedererzählt hat. (...) es war eine Gegend, in der Menschen und Bucher lebten*). (Celan 2001: 395; 1986: III, 185)

Two years later, in a speech that thematizes the road in numerous places, *Nesselweg*, travel, or *Toposforschung*, Celan supplements the words from Bremen:

> Then does one, in thinking of poems, does one walk such paths with poems? Are these paths only by-paths, bypaths from thou to thou? Yet at the same time, among how many other paths, they're also paths on which language gets a voice, they are encounters, paths of a voice to a perceiving Thou, creaturely paths, sketches of existence perhaps, a sending oneself ahead toward oneself, is search of oneself... A kind of homecoming. (...) I also seek – for I'm back again where I began – the place of my own origin. (*Geht man also, wenn man an Gedichte denkt, geht man mit Gedichten solche Wege? Sind diese Wege nur Um-Wege, Umwege von dir zu dir? Aber es sind ja zugleich auch, unter wie vielen anderen Wegen, Wege, auf denen die Sprache stimmhaft wird, es sind Begegnungen, Wege einer Stimme zu einem wahrnehmenden Du, kreatürliche Wege, Daseinsentwürfe vielleicht, ein Sichvorausschicken zu sich selbst, auf der Suche nach sich selbst... Eine Art Heimkehr. (...) Ich suche auch, denn ich bin ja wieder da, wo ich begonnen habe, den Ort meiner eigenen Herkunft*). (Celan 2001: 412–413, 1986, III: 201–202)[12]

Probably, we ought to very carefully reconstruct these two protocols offered by Paul Celan, as an endlessly poor and laborious answer to the question "*Heimat, und*

---

paradoxically, is not part of the community or not part of us all. One without their "*Wir*" – does not exist or will soon cease to exist.

11 The poem "Heimkehr," written between 1955 and 1956 is translated into Russian, for example, as "Vozvrascenie na Rodinu." Cf. Celan 2014: 94; Celjan 2013: 79.

12 There is also another translation, by Pierre Joris: "Does one take, when thinking of poems, does one take such routes with the poems? Are these routes only re-routings, detours from you to you?" or "The poem is the detour from you to you; it is the route (*Das Gedicht ist der Umweg von dir zu dir; es ist der Weg*)." (Celan 2011: 11, 40)

Ich?". Likely this reconstruction would be successful to the extent that each one of us individually answers this question that affects us all, without blindly following all of Celan's various associations and obsessions about home and way home, as Celan's answer connects all too quickly with some other answers in the histories of wandering and discomfort.[13] Celan's response is above all in resistance (unconscious and certainly intuitive) to nostalgia, that is, to Heidegger and his understanding of the path and the protocol revealed by the word *Umweg*.

Well then, what is nostalgia? In paragraph 4 of his thesis *Nostalgia oder Heimwehe*, defended 22 June 1688 at the Medical School in Basel, Johannes Hofer, describes the difficult injury and agony of a local peasant girl. After a fall from some height, she lay motionless and unconscious in an improvised hospital, slowly coming to only after surgery and various remedies. Awake and seeing unknown women caring for her, she is all of a sudden overcome with nostalgia (*Nostalgia statim correpta*). Hofer tells us she is rejecting all food and answering all questions identically: "*Ich will Heim, Ich will Heim* [I want to go home, I want to go home]. When the parents finally allow her, debilitated, to return to her home (*Tandem ergo a parentibus licet maxime imbecillis domum est delata*) (Hofer 1688: 8),[14] her state improves suddenly and without medication. This case, along with another case presented in the same paragraph, of a young student from a family in good standing, who goes to Basel for his studies and falls gravely ill (his condition improves rapidly when he is ordered home at once), help Hofer construct his argument about the appearance of an entirely new and odd illness, and help explain why there is an epidemic among Swiss soldiers.

For Celan, there is first of all no house where we will be at home (when translating Emily Dickinson, he discovers the vindicating phrase "homeless at home"), primarily because the home is no more or because the home has been abandoned forever.

Ich war nicht einmal zuhause, als ich daheim (zuhause) war.

I am not at home even when it seems that I am at home, at mine. In one way or another, I cannot be a witness – one without a house or *Heimat* cannot testify – at the same time, I testify that I am not a witness, that I cannot be a witness, and

13 One of the main reasons I choose to speak about Celan who explicitly mentions Bukovina as *Heimat* or as *Heiland* (and not Bukovina as Ukraine or as snow, or as mother, etc.) – this is the Bukovina where in 1930 there are 93,101 Jews officially counted, while on 20 May 1942 there were 17,033 (Günzel1995: 24) – refers sometimes to the unsatisfactory readings of Celan by Jacques Derrida (today it seems to me that my doctoral thesis and some texts, such as the ones on C. Schmitt or W. Benjamin, were also to an extent my attempt to correct his oversights). In *Sovereignties in Question. The Poetics of Paul Celan* (a compendium of different texts published in French), Derrida all too quickly and insufficiently justifiably reduces Celan's engagement to Freud's "Unheimlich" (Derrida 2005). It beggars belief that in the seminar "Le Bête et le souverain" (2002-2003) where he interprets Heidegger's "Umweg" in detail along with a philosophy essentially opposed to any "Umweg," Derrida does not oppose this idea by the "Umweg" in Celan or "Methode als Umweg" in Benjamin. This omission is all the greater when we know that in his last seminars, Derrida often analyzes Celan's "Meridian."

14 ; English translation by Carolyn Kiser Anspach, "Medical Dissertation on Nostalgia by Johannes Hofer, 1688," published in the *Bulletin of the Institute of the History of Medicine* 2 (1934).

as such I am a witness without testimony. This naked 'Ich' is never at home, never with itself, and never feels itself in a space that gives the sense of potential closeness. This is not a hidden nostalgia, nor the possibility of "being at home anywhere,"[15] nor does Celan invent some kind of authentic language of the native land (he speaks of poetry that might help show the way, but not about language; poetry supposes translation and transcends language and plurality of language[16]). In other words, there is no trace in Celan of Humboldt's idea of inseparability of *Sprache* and *Heimat*. It seems to me that Humboldt's famous sentence – "*Die wahre Heimath ist eigentlich die Sprache* [The true home is really language] (Humboldt 1848: 322) – from his 1827 letter to Charlotte, with all its various incarnations and repetitions from Hannah Arendt to Derrida or Gadamer, could also be completed well in Heidegger's speech on the occasion of the seven hundredth anniversary of his native city of Masskirch: "*Unsere Sprache nennt den Zug zur Heimat das Heimweh*". (Heidegger 2000: 578)

In the second fictional protocol Paul Celan offers as a response to the question of the position of "*Ich*" in relation to "*Heimat*" ("*Heimat, und Ich?*"), we find a collision between the "arrival from somewhere" and "return back." These two operations conducted simultaneously mean that there can never be 'direct' movement, nor on a straight path (*Weg*); rather, that the "*Ich*" is always within the register of "*Um-wege*." If we remove, or put aside those aspects most interesting to Celan and his readers from his speeches in Bremen and Darmstadt – the idea of *Landschaft* or *Heimat* where books and people live together, as well as the idea that "one travels through poems" back to this magical region where people and books comprise the *Wir* – then we are left with a double perspective that necessarily cancels this ever-wandering or pseudo-wandering "*Ich*" down myriad detours. As if Celan's "*Ich*" (and not only his, of course) – does not move at all.[17]

In the passage delivered by Celan in Bremen, he informs us that he has arrived from somewhere, that this country is unknown to us and that he has arrived "*auf welchen Umwegen!*". The question that remains is as follows: "*Aber gibt es das denn: Umwege*?"[18] In the fragment from Darmstadt, the direction or perspective is the opposite: now we are dealing with "*Eine Art Heimkehr*." The return seems to be performed through poems, but through these detours (*Umweg*). Here too, Celan asks a similar question: "*Sind diese Wege nur Um-Wege, Umwege von dir zu dir*?"[19]

---

15 "*Die Philosophie ist eigentlich Heimweh, ein Trieb, überall zu Hause zu sein*." Heidegger analyzes this statement by Novalis in § 2, at the beginning of the winter seminar in 1929 (Heidegger 1983: 7-10). Celan of course did not know of Heidegger's seminar.

16 In "Meridian," Celan defines poetry as that which is outside of text (*hors texte*). "*Eh bien, me voici, et il le fallait bien, hors texte.*"

17 Levinas formulates this situation as follows: "*La circularité de ce mouvement sans retour*." (Levinas 1976: 64)

18 There are several variants of translation into English ("roundabout ways," "byway," etc.). One of them, "*by what detours! But are there such things: detours?*" (Celan 2001, 395). Jean Launay translates similarly into the French: "*par quels détours! Mais est-ce que cela existe : des détours*?" (Celan 2002: 55), "*kakimi okolnjimi putjami! susestvujut li vobse okolnie puti*?" (Celjan 2013: 363)

19 John Felstiner's translation: "*Then does one, in thinking of poems, does one walk such paths with poems? Are these paths only by-paths, bypaths from thou to thou?*" (Celan 2001:

Once again, if we are able to hold at distance Celan's main assumption, of which he speaks in a passage on antibiography towards the end of 1953 (or beginning of 1954), that the *Heimat* of the poet is indeed the poem itself, a poem that always brings together I and Thou,[20] the only thing that remains unchanged in Celan is his uncertainty and dubiousness regarding the existence of these detours.

Even though the Darmstadt speech, alternating between seeking and finding, ends in the discovery of the Meridian, it would appear that Paul Celan's only effort is to sustain and hold this question or these questions about the paths that come from the abyss and lead to the abyss ("The abyss is their *Heimat*, their language is being-on-the-road"), on a course filled with hesitation, always tardy, never arriving.[21] In the sentence that follows, at the moment *Herkunft* is reached, where everything is discovered and where everything is in place, the only remaining thing is to further ask and always seek anew:

> "*Ich suche auch, denn ich bin ja wieder da, wo ich begonnen habe, den Ort meiner eigenen Herkunft*".[22]

Translated by Edward Djordjevic

References:

Améry, Jean (1966), *Jeinseits von Schuld und Sühne. Bewältigungsversuche eines Uberwältigten*, München: Szczesny.
Benveniste, Émile (1980), *Le Vocabulaire des institutions indo-européennes*, Vol. I-II, Paris: Minuit.
Celan, Paul (2014), *Die Gedichte. Kommentierte Gesamtausgabe*, Frankfurt am Main: Suhrkamp.
—. (2011), *The Meridian*, tr. P. Joris, Stanford, Stanford University Press.
—. (2005), " *'Mikrolithen sinds, Steinchen' Die Prosa aus dem Nachlass*", Frankfurt am Main: Suhrkamp.
—. (2004), *La bibliothèque philosophique*, A. Richter, P. Alac, B. Badiou (ed.), Paris: Editions rue Ulm.
—. (2002), *Le Méridien & autres proses*, tr. J. Launay, Paris: Seuil.
—. (2001), "Bremen Speech," in: *Selected Poems and Prose of Paul Celan*, tr. J. Felstiner, New York / London: Norton.
—. (1986), *Gesammelte Werke*, Vol. III, Frankfurt am Main: Suhrkamp.
Celjan, Paul (2013), *Stihotvorenija. Proza. Pisma*, tr. T Baskakova and M. Belorusets, Moskva: Ad Marginem.

---

412) Pierre Joris' translation: "*Are these routes only re-routings, detours from you to you*?" (Celan 2011: 11). Jean Launay's French translation: "*Seraient-ce simplement des allers-retours, des chemins qui font un détour pour aller de toi à toi-même*?" (Celan 2002: 82) Only in Russian is the translation unchanged: "*okoljnije puti*."

20 "*Echte Dichtung ist antibiographisch. Die Heimat des Dichters ist sein Gedicht, sie wechselt von einem Gedicht zum anderen.*" (Celan 2005: 95-96) True poetry is therefore not biographical; that is to say, for it to be poetry, there can be no testifying about oneself in it.

21 *Einer der Wege: andere ~~führen~~, aus dem Bodenlosen kommend, führen{d} ins Bodenlose zurück; der Abgrund ist ihnen die Heimat; ihre Sprache ist ihnen das Unterwegssein; nicht mehr"* (Meridian, F 40-42).

22 "I also seek – for I'm back again where I began – the place of my own origin." (Celan 2001: 413)

Derrida, Jacques (2005), *Sovereignties in Question. The Poetics of Paul Celan*, New York: Fordham University Press.
—. (1996), "Demeure: filiation et témoignage" in: M. Lisse (ed.), *Passions de la littérature: avec Jacques Derrida*, Paris: Galilée.
Günzel, Elke (1995), *Das wandernde Zitat. Paul Celan im jüdischen Kontext*, Würzburg: Königshausen & Neumann.
Heidegger, Martin (2000), *Reden und andere Zeugnisse eines Lebensweges 1910-1976*, GA 16, Frankfurt am Main: Vittorio Klostermann.
—. (1983), *Die Grundbegriffe der Metaphysik. Welt-Endlichkeit-Einsamkeit*, GA 29/30, Frankfurt am Main: Vittorio Klostermann.
Hofer, Johann (1688), *Dissertatio medica De Nostalgia, oder Heimwehe*, Basel: Typis Iacobi Bertschii [Eng. tr. C. K. Anspach, "Medical Dissertation on Nostalgia by Johannes Hofer, 1688," *Bulletin of the Institute of the History of Medicine* 2 (1934): 376–391]
Levinas, Emmanuel (1993), *Dieu, la Mort et le Temps*, Paris: Grasset.
—. (1976), *Noms propres*, Paris: Fata Morgana.
Olschner, Leonard (2007), *Im Abgrund Zeit. Paul Celans Poetiksplitter*, Göttingen: Vandenhoeck & Ruprecht.
Von Humboldt, Wilhelm (1848), *Briefe von Wilhelm von Humboldt an eine Freundin*, Vol. I, Leipzig: F. A. Brockhaus.

Petar Bojanić

Paul Celan o nemogućnosti svedočenja: „*Ort meiner eigenen Herkunft*"

Apstrakt

Paul Celan u svojim pesmama ne upotrebljava reči kao što su teritorija, granica, prelaženje granice, a veoma retko reč prostor. Namera mi je da rekonstruišem tragove „Heimat" kod Celana (u nekoliko pesama iz različitih perioda „*Heimat*" je važna tema) i eventualno pokušam da opišem šta je Heimat mogao da znači za mladoga Paul Antschel (njegovo pravo ime). Odnosno da li je „Heimat" zaista u sinonimiji sa onim što Celan kaže, puno godina kasnije, u govoru povodom uručenja Georg-Buechner-Preis: „*Ich suche auch, denn ich bin ja wieder da, wo ich begonnen habe, den Ort meiner eigenen Herkunft*". Pokušaću da u pesmama napisanim u vreme kada Antschel uči hebrejski i prvi put čita knjige Martina Bubera (Israel Chalfen), pokažem neke osnovne figure koje Celan vezuje za svoj tadašnji život u Bukowini, a u okruženju Czernowitzer Judentums. U ovom istraživanju ću se koristiti, osim radova Israela Chalfena, Else Keren i Elke Guenzel, i knjigom koja je objavljena pre desetak godina u kojoj je detaljno popisana pariska biblioteka Paul Celana. Namera mi je da konsultujem ovaj arhiv u narednom periodu jer Celan na marginama svojih knjiga na nekoliko mesta evocira svoje prve stvaralačke godine.

Ključne reči: svedočenje, domovina, kuća, rođenje, granica

UDK: 123.1 Hegel G.V.F.
https://doi.org/10.2298/FID1804553J
Original Scientific Article
Received: 08.02.2018. Accepted: 03.05.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Rastko Jovanov

ZWISCHEN PHILOSOPHIE UND STAAT HEGELS DIALEKTIK DER FREIHEITSINSTITUTIONALISIERUNG

ZUSAMMENFASSUNG

Hegel betrachtet in seinem philosophischen System die verschiedenen Bestimmungen der Freiheit; er unterscheidet die subjektive, objektive und absolute Freiheit. In dieser Arbeit wird mich primär die Dialektik der objektiven Freiheit interessieren, die Hegel am Niveau der Staats- und Geschichtsphilosophie einführt, um danach die Problematik der Geschichtlichkeit der objektiven Freiheit auszulegen, und schließlich zu behaupten, dass der Freiheitsbegriff erst am Niveau des absoluten Geistes die Qualität der wahren Geschichtlichkeit bekommt. Damit wird im Denken ein Raum geöffnet um eine These von der dialektischen Spannung, die in der Hegelschen Auffassung der erfüllenden Freiheit in beiden Niveaus seines Systems anwesend ist, aufzustellen: Nämlich, im Staat als Erreichung der konkreten Freiheit innerhalb der Objektivität des Geistes, als auch im scheinbaren a-politischen Freiheitsbegriff in der Sphäre des absoluten Geistes, bzw. der Sphäre des konkreten Denkens, der Sphäre der Philosophie selbst.

SCHLÜSSELWÖRTER

Hegel, Freiheit, Staat, Geschichte, absoluten Geist

Der Begriff der Freiheit ist einer der zentralen Begriffe der neuzeitlichen Philosophie. Die Philosophie der Spätmoderne, insbes. die des deutschen Idealismus und Kants ist mit der Frage bestimmt: Wie denkt man die Freiheit des Subjekts, bzw. wie ist die Freiheit überhaupt zu denken? Die Kantische und Fichtesche Philosophie, mit deren Auseinandersetzung Hegel ein eigener philosophischer Gedanke aufkommt, stehen unter einer explizit anführten Stellung, dass der Begriff der Freiheit der „Eckstein" ihrer Philosophie ist. Mit der Betrachtung, dass die Weltgeschichte der Fortschritt im Freiheitsbewusstsein ist und dass die Freiheit eigentlich die Aufhebung der Entzweiung ist, d.h. Zurückkommen des Menschen zu sich selbst – stellt sich Hegel an die grundlegende Denkweise des deutschen Idealismus und versucht sie im Zusammenhang mit der antiken Auffassung von der Sittlichkeit im „Gedanke seiner Zeit" durchzuführen. Hier handelt es sich um die Vergeschichtlichung der Begriffe der Freiheit und des Rechts und um die Einführung der Idee der Geschichte in den Sichtkreis der europäischen Philosophie.

Rastko Jovanov: Research Fellow, Institute for Philosophy and Social Theory, University of Belgrade;; jovanov@instifdt.bg.ac.rs.
This paper is a result of the project 179049 funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Durch die Wirklichkeit der geschehenen Geschichte, durch die in der Französischen Revolution positivierte Freiheit des Subjekts und durch die Zuschreibung der Vernünftigkeit zur bürgerlichen Gesellschaft – vergeschichtlicht Hegel den ursprünglichen Kantischen Begriff der Freiheit und vollendet Kants kopernikanische Wende im Rahmen der praktischen Philosophie, bzw. Geschichtsphilosophie. Darum bemüht sich Hegel den Kantischen Begriff der Freiheit mittels der Dialektik der *geschehenen* Freiheit in der Geschichte zu retten. Die Freiheit ist daher nicht mehr nur eine Idee im Kantischen Sinne, sondern vielmehr ein Faktum in der geschichtlichen Welt. Freiheit und Geschichte können sich nicht mehr getrennt denken.

Die Freiheit ist nicht nur einer von den Begriffen im Hegelschen System der Philosophie, - sie ist der Begriff, von welchem das dialektische Denken dieses, wie ihn Heidegger einmal genannt hat, „letzten Griechen" anfängt und um welchen es sich dreht. Der Begriff, von welchem der Ausbau des Systems Hegels anfängt und mit welchem es endet.[1] Der Begriff des *Geistes*, dieser Inbegriff der Philosophie Hegels, ist nichts anderes als das Selbstbewusstsein der Freiheit; *das Recht* ist die Freiheit in seinem äußerlichen Dasein; und auch zur *Vernunft* kommt man mittels der Freiheit, weil sie nichts anders ist als *Verstand* der seiner Freiheit bewusst ist, d.h. sie ist jene, welche die Bestimmungen in die objektive Welt einsetzt und sich durch die in sich beinhaltete Intersubjektivität dieses *Wir* verwirklicht.

Hegel operiert in seinem System mit verschiedenen Bestimmungen der Freiheit; er unterscheidet die subjektive, objektive und absolute Freiheit. In dieser Arbeit wird mich primär die Dialektik der objektiven Freiheit interessieren, die Hegel am Niveau der Weltgeschichte einführt, um danach die Problematik der Geschichtlichkeit der objektiven Freiheit auszulegen, und schließlich zu behaupten, dass der Freiheitsbegriff *erst* am Niveau des absoluten Geistes die Qualität der wahren Geschichtlichkeit bekommt. Damit wird im Denken ein Raum geöffnet um eine These von der dialektischen Spannung, die in der Hegelschen Auffassung der erfüllenden Freiheit in beiden Niveaus seines Systems anwesend ist, aufzustellen: Nämlich, im Staat als Erreichung der konkreten Freiheit innerhalb der Objektivität des Geistes, als auch im scheinbaren a-politischen Freiheitsbegriff in der Sphäre des absoluten Geistes, bzw. der Sphäre des konkreten Denkens, der Sphäre der Philosophie selbst.

I

Das Konzept der spekulativen idealistischen Philosophie ermöglicht Hegel die Jahrhunderte dauernde Streitigkeit des Freiheitsproblems zu lösen, die traditionell über das Dasein der Willensfreiheit geführt wurde und die die Gegner auf zwei gegenteilige Lager teilte: auf Deterministen und Indeterministen. Die Freiheit als übernatürlich denkend, als etwas was sich nicht aus der körperlichen Konstitution des Menschen ableiten lässt, stellt Hegel seine These auf, dass die Freiheit nur real ist als der Wille selbst. Hegel reduziert sie dabei aber nicht auf einzelne Willensakte

1 Vgl. Hegels Aussage zu Beginn der *Enzyklopädie* (1817; § 5): „Die Philosophie kann auch als die Wissenschaft der Freiheit betrachtet werden [...] nur in der Philosophie ist die Vernunft durchaus bei sich selbst". Sofern nicht anders bezeichnet, werden Hegels Schriften nach Hegels Werke in zwanzig Bänden (=TW), auf der Grundlage der Werke von 1832-1845 neu edierte Ausgabe, E. Moldenhauer, K.M. Michel (red.), Frankfurt a.M. 1970ff., zitiert.

des Individuums, sondern fasst sie als ursprüngliches Wesen des Menschen, der Gesellschaft, des Staates und der Geschichte, bzw. der Wirklichkeit selbst auf.

Damit wurde das Problem der Freiheitsverwirklichung unmittelbar mit dem Prinzip der gesellschaftlichen Entwicklung verbunden, welches Hegel im Anschluss an Rousseau und Kant, als widersprüchlich aber zugleich als unwiederbringlichen Fortschritt in der Weltgeschichte interpretierte. Es handelt sich um die Emanzipation, wohin man m.E. den Kern des Hegelschen Denkens der Freiheitsdialektik in der Geschichte findet. An einer Stelle der *Philosophie der Weltgeschichte* spricht Hegel davon, dass der Mensch „nicht sowohl *aus* der Knechtschaft befreit worden [ist], als vielmehr *durch* die Knechtschaft" (TW 12: 487). Immer wieder begegnet der Mensch in der Geschichte etwas ihm Gegenüberstehendes, etwas was ihn beschränkt und was ihn, im strengsten Sinne genommen, in neue Knechtschaft bringt. Diesen Begriff der Knechtschaft denkt Hegel auf eine spekulative Weise, was der Dialektik als einer Methode ermöglicht, die Wirklichkeit wesentlich als Prozessualität zu verstehen. Jede Epoche hat daher seine eigene Aufgabe der geschichtlichen Befreiung. Denn „[d]ie Geschichte des Geistes ist immer seine Befreiung: das, was er ist, sich zum Gegenstand zu machen, es zu wissen und dadurch sich von ihm zu befreien und somit eine höhere Stufe zu erreichen"[2]. Die Befreiung des Subjekts von den objektiven Inhalten, zugleich auch ihre Veränderung, ist möglich, weil die Objektivität des Geistes nicht zu etwas Natürlichem gehört: Das Objektive existiert nur, weil „es gewußt wird" (Enz §442). Und es wurde nur durch die Philosophie gewusst, durch die Idee selbst: „die *Vollendung* eines Erfassens ist zugleich seine Entäußerung und sein Übergang"[3].

Das, was vorbereitet und was als Ursprung und Bedingung der Entstehung neuer Befreiung *durch* die Knechtschaft existiert, ist die Philosophie, die höchste Form des absoluten Geistes, die einzig, trotz ihrer „objektiven" Verwirklichung in der Weltlichkeit der Welt, die Geschichte transzendieren und dem Begriff der konkreten (schon verwirklichten) Freiheit die zerstörende „Abstraktheit" und die konstruktive, stiftende „Konkretheit" geben kann. Der Freiheitsbegriff Hegels ist *apolitisch*, der aber – dieses genuine im Begriff – nur innerhalb der Objektivität des Geistes sein Dasein hat. Daher ist die vernünftige Wirklichkeit die Wirklichkeit des verwirklichten Geistes. Der Freiheitsbegriff muss sich also im Objektiven verwirklichen, was immer durch die Gewalt und Negativität geschieht. Der *Rest der Negativität* im „äußeren Staatsrecht" Hegels wird damit zum Ort der zukünftigen Begegnungen der Freiheit und Geschichte.Alles was in der objektiven Welt wird und existiert, stellt das Werk der Negativität dar, bzw. das Werk der Freiheit des menschlichen Willens. Jede Form der Verwirklichung der Freiheit, d.i. jede institutionelle Organisation der Positivität der Freiheit, ist der Sieg der Negativität

2 G.W.F. Hegel, *Die Philosophie des Rechts. Die Mitschriften Wannenmann (Heidelberg 1817/18) und Homeyer (Berlin 1818/19)*, hrsg. von K.H. Ilting, Stuttgart 1983, § 135.

3 TW 7: § 343. Vgl. dazu die Gleichstellung von Vernichtung und das Aufgehen eines neuen Prinzips in der *Philosophie der Geschichte* (TW 12: 104). „The rose in the cross of the present, the tragic irony of Hegel's dialectical apprehension of his world, means that while Hegel saw himself as comprehending the new world of post-1789 (or post-1815) Europe, this by itself meant that this new world, which Hegel heralded in his *Phanomenology*, is already reaching its maturity and is somehow, slowly but surely, on its way out" (S. Avineri, *Hegel's Theory of the Modern State*, Cambridge 2003, s. 129).

über das Vorhandene, das Unmittelbare.[4] Die Weiterentwicklung des Geistes ist das Werk des Besiegten – dessen, wer sich durch die Negativität seiner Freiheit die eigene Befriedigung in der Objektivität der Welt sucht.

Die Versöhnung in der Geschichte geschieht immer in einer bestimmten Form der verwirklichten, d.h. konkreten Freiheit, in der sich die Geschichtlichkeit des Geistes, d.h. das Innere der Geschichte zeigt, – der Gedanke, welcher sich in der Freiheit des Selbstbewusstseins bemüht, jede Objektivität, die ihm im Weg steht, aufzuheben, um bei-sich-selbst-zu-sein. Die Frage der politischen Freiheit wird sich – wie in dieser Arbeit interpretiert wird – als zeitlich bedingt und grundsätzlich unlösbar zeigen, weil jede Gestalt der Verwirklichung der Freiheit, ob es sich um einen Staat oder um einen Völkerbund handelt, ihr Ende hat. Alles Bestehende, so Hegel ein bisschen zynisch, geht mit Recht zugrunde, sodass es keine *Anhaltung* der Dialektik gibt, weil der Gegenstand der philosophischen Betrachtung selbst dialektisch ist. Auf dem onto-logisch-begrifflichen Niveau wird die Idee als Werden aufgefasst, als Einheit des Seins und des Nichts, des Begriffs und der Objektivität bestimmt: „Denn das Wahre – so Hegel in seinem Brief an Duboc vom 30. 07. 1822 – ist nicht ein Ruhendes, Seiendes, sondern nur als sich selbst bewegend, als lebendig; – das ewige Unterscheiden [...] Nur als diese Bewegung in sich, die ebenso absolute Ruhe ist, ist die Idee, Leben, Geist".

Walter Jaeschke betrachtet, dass die Hegelsche Entdeckung der Geschichtlichkeit in der Aussage zu finden ist „was *wir* sind, sind wir zugleich geschichtlich" (TW 18: 21), dessen Ursprung in der Bestimmung der Vernunft als geschichtlich, als Entfaltung dessen, was Vernunft an sich ist, liegt. Die Geschichtlichkeit des Geistes erkannte Hegel zuerst durch die Geschichtlichkeit des absoluten Geistes und in seiner Berliner Zeit erweitert er sie auf das Feld der Weltgeschichte. Nach Jaeschke ist aber „[d]ie Vernunft *in* der Geschichte [...] nicht die Vernunft *der* Geschichte"[5]. Er beschließt daraus, m.E. ganz richtig, dass die Verwirklichung der Freiheit nicht das Ziel der Weltgeschichte sein kann, weil der Fortschritt im Freiheitsbewusstsein nur am defizitären Modus der Freiheit gerichtet ist, nämlich an der politischen Freiheit. Die volle Freiheitsverwirklichung ist das Ergebnis der Geschichte des absoluten Geistes.

Ich werde hier unmittelbar an seiner Auslegung anknüpfen und sie mit der These ergänzen, dass der Begriff der absoluten Freiheit, der in der Weltgeschichte seine Kraft der Zerstörung zeigt, gerade das ist was tief geschichtlich an dem Begriff der *Philosophie* selbst ist und dass die Philosophie als volles Sich- Wissen des Geistes, d.h. als höchste Form des absoluten Geistes, die *tätige* Seite in der Weltgeschichte darstellt.

Das subjektive Denken selbst kann nicht zum objektiven Freiheitsbegriff kommen, weil die „Freiheit im Gedanken, [...] nur *den reinen Gedanken* zu ihrer Wahrheit"

4 Im Gegensatz dazu würde die Rechtsphilosophie Hegels, statt in der „verlangten" Versöhnung, im gewalttätigen System des gegenseitigen Krieges der partikulären Staaten enden. Deshalb schließt Hegel die *Grundlinien* mit der Geschichtsphilosophie, d.h. mit der Beschreibung der vergangenen Gestaltungen der Verwirklichung der Freiheit. Der moderne Staat wird sich daher als ein geschichtlicher zeigen und seine Form als vorübergehende: „Dies ist nun der Standpunkt der jetzigen Zeit, und die Reihe der geistigen Gestaltungen ist für jetzt damit geschlossen" (TW 20: 461).

5 W. Jaeschke, „Die Geschichtlichkeit der Geschichte", in: *Hegel-Jahrbuch* 1995, s. 370.

(TW 3: 158) hat, und nicht „die lebendige Freiheit selbst" (ebd.) ist. Wenn auch Hegel oft anführte, dass der Mensch nur im Denken frei ist, dann hat er die Voraussetzung seines Idealismus selbst im Sinn, denn das Denken ist die Substanz nicht nur des Menschen, sondern auch der Wirklichkeit selbst. Die lebendige Freiheit ist damit eine gegenständliche Freiheit, die ihre Wahrheit in einer bestimmten Zeit hat und wenn sie zugleich mit ihrer Zeit zugrunde geht und als eine Positivität vor dem Subjekt erscheint, dann würde sie ihre Substanz verlieren. Die Philosophie ist daher zugleich die Erkenntnis des Zerfalls einer Welt, als auch die Erkenntnis des kommenden neuen Prinzips, obwohl zuerst nur in seiner abstrakten Form.

In der Berliner Eintrittsrede hebt Hegel hervor, dass das Dasein des *modernen* Staates und der allen anderen sittlichen Institutionen auch die *Rechtfertigung durch das Denken* und die entsprechende philosophische Theorie verlangt.[6] Die *Grundlinien* weisen damit die Tendenz auf, selbst der Teil des Daseins der politischen Institutionen zu werden. Nicht nur ein Teil, sondern ihr wesentlicher Teil, weil nur mittels der Gedanken und der eigenen Gesinnung des Subjekts die politischen Institutionen des modernen Staates zur Anerkennung seitens der Bürger kommen kann. Das ist „im europäischen Sinne Freiheit" (Enz §503 Anm.) der modernen Subjektivität, die notwendigerweise ihre Befriedigung erfinden muss und Hegel betrachtet, dass nur innerhalb des Staates, aufgefasst als das sittliche Ganze, die moderne Subjektivität ihre Substanz schaffen wird.[7]

Daher ist der Philosophie, als Rechtsdiskurs, die Aufgabe gestellt um die Anwendung des Rechts innerhalb des Verfassungsstaates zu rechtfertigen, d.h. die Gewalt als das wesentliche Moment der täglichen Rechtsverwirklichung im Staate zu berechtigen. Der Sittlichkeitsbegriff übernimmt diese Rolle um den Rechtszwang als etwas „gewolltes" zu bestimmen. Von ihren Anfängen in der praktischen Betrachtung der Lebenswelt wurde die Philosophie auf die Untersuchung des *Nomos* im Rahmen des *Polis* (d.i. der sittlichen Gemeinschaft), bzw. auf die Erwägung der gerechten (Platon) oder guten (Aristoteles) Lebensführung angewiesen. Die Philosophie ist institutionalisiert um dem Diskurs den Begriff des Rechts, das einzig die Beziehungen des Mitlebens in Ordnung bringen kann, zu bescheren; – dem mit der Gewalt stiftendem Recht und welches sich mittels der Gewalt – d.i. mittels seiner Verwendung – erhält.[8] Ohne Philosophie würde es weder das Leben im Recht geben, noch könnte die Gewalt gerechtfertigt werden. Das Recht erhält sich (oder stört sich) mittels, oder in dem philosophischen Diskurs über das Recht.

Die theoretische Stelle, die dieses Verhältnis zwischen der Philosophie und dem Recht (des Staates) enthüllt, ist jene Stelle, die den philosophischen Diskurs in den

6 In dem modernen Staat „das was gelten soll, vor der *Einsicht* und dem *Gedanken* sich rechtfertigen muß" (TW 10: 400). Siehe auch Hegels Brief an Niethammer (*Briefe von und an Hegel*, Bd. 2, hrsg. von J. Hoffmeister, Hamburg 1969ff., s. 271), worin Hegel die *Grundlinien* als „ein Buch über Staatspädagogik" darstellt.

7 Vgl. die berühmte Stelle aus der *Politik* Aristoteles' (1252 b 27 – 1253 a), wo steht, dass nur durch die politische Gemeinschaft der Mensch *autarkeia* bekommt. Für die Einwirkung Aristoteles auf die Hegelsche politische Philosophie, vgl. die bedeutungswerte Studie von K.H. Ilting „Hegels Auseinandersetzung mit der Aristotelischen Politik" (*Philosophisches Jahrbuch* 71/1963-64).

8 Vgl. J. Derrida, „Force of law: the 'mystical foundation of authority'", *Cardozo Law Review* 920/1989.

Grundlinien den Schranken seiner Geltung aufstellt. Die zwischenstaatlichen Verhältnisse zeigen somit die territorialen Rechtsgeltungen auf. Zwischen zwei (oder mehr) im Streit gekommene Rechte eröffnet sich ein mit der Negativität erfüllender Durchlass, ein Raum des willkürlichen Spiels der Freiheit. Damit die in der Naturgewalt zurück gefallenen Freiheiten nicht das letzte Wort seiner praktischen Philosophie sein würde, stellt Hegel das Recht des Weltgeistes als einen eigentlichen Richter in diesem Spiel auf. Dieses höhere Recht ist nicht mehr mit der Souveränität, der Territorialität oder mit der eigenen gewalttätigen Verwendung bestimmt. Es hat keinen eigenen Geltungsraum: seine Geltung befindet sich in das zu-sich- selbst-kommen des Geistes, bzw. in der Erkenntnis, dass eine besondere und beschränkte Form der Freiheitsverwirklichung in der Weltobjektivität realisiert ist. Diese Realisierung als endliche ist eine vorübergehende und der Geist versucht bereits eine höhere Befriedigung zu finden. Seine Geltung ist seine Befreiung.

Die auf der Ebene des Weltgeistes gefundene Sittlichkeit ist von den permanenten internationalen Streitigkeiten und Weltkriegen untrennbar. Sie ist, trotz der stetigen Zurückdrängung und Suspendierung, was immer wieder als das zwischen den Staaten geteilte Allgemeine erscheint. In Anbetracht dessen gibt es in der Welt nach Hegel nur eine einzige Philosophie, welche im Wesentlichen die *Weltphilosophie* ist. Institutionalisiert um die Gewaltverwendung zu rechtfertigen, findet sie aber – wie die Epopöe und die Odyssee des absoluten Geistes – ihre Ruhe nur in sich selbst, damit ihre Freiheit – in ihrer a-politischen Gestalt – bei-sich-selbst-sein kann. Die Freiheit daher, bzw. jede Stiftung der gesellschaftspolitischen Gemeinschaft, beginnt mit der Gewalt, mit dem Missbrauch der Freiheit.[9]

II

Der philosophische Gedanke, verwirklicht in der Sittlichkeit der politischen Gemeinschaft, ist nicht lediglich der „unbewegte Beweger" der geschichtlichen Entfaltung; er ist gleichzeitig auch die Kritik der Positivität der bestehenden politischen Institutionen und Arten der intersubjektiven Bindungen der Menschen innerhalb der politischen Gemeinschaft.[10] Der a-politische Freiheitsbegriff, der zum Bewusstsein

9 Vgl. dazu I. Kant, *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*, 6. Satz, AA VIII: 23; als auch *Zum ewigen Frieden* (AA VIII: 371): „... so ist in der Ausführung jener Idee (in der Praxis) auf keinen andern Anfang des rechtlichen Zustandes zu rechnen, als den durch Gewalt, auf deren Zwang nachher das öffentliche Recht gegründet wird; welches dann freilich (da man ohnedem des Gesetzgebers moralische Gesinnung hiebei wenig in Anschlag bringen kann, er werde nach geschehener Vereinigung der wüsten Menge in ein Volk diesem es nun überlassen, eine rechtliche Verfassung durch ihren gemeinsamen Willen zu Stande zu bringen) große Abweichungen von jener Idee (der Theorie) in der wirklichen Erfahrung schon zum voraus erwarten läßt." C. Perelman gibt auch vor (C. Perelman, „Was der Philosoph vom Studium des Rechts lernen kann", in: *Wissenschaft und Weltbild. Zeitschrift für Grundfragen der Forschung*, Nr.4 1966, s. 161ff.), dass der Bruch mit einer früheren Ordnung und die Einführung eines neuen Regimes in der Geschichte nur mit der Gewalt realisiert werden kann, und dass es eine Illusion ist, wenn man glaubt, dass eine neue Ordnung sich allein mit seiner eigenen Rationalität aufdrängen kann.

10 Siehe die hegelianische Auffassung Kojeves' von der Philosophie, die Anteil an der Geschichte nimmt (A. Kojeve, „Tyranny and Wisdom", in: Strauss, L., *On Tyranny*. Including Strauss-Kojève Correspondence, Chicago 2000, s. 152. Er interpretiert die Teilnahme

seiner selbst im absoluten Geist kommt, ist jener der sich in der Objektivität des Geistes als konkrete Freiheit verwirklicht. Hierdurch ist er zugleich nur *scheinbar* apolitisch, weil er sich immer im Anderen-sich-selbst finden muss; er muss sich vergegenständlichen und sich in der Objektivität der Welt, d.h. in der Objektivität seiner selbst anschauen. Daher wird Hegels Stellungnahme, dass es ohne Staat als Verwirklichung der konkreten Freiheit nicht zum wunderschönen Sonnenaufgang in Griechenland kommen würde, verständlich; d.h. die politische Freiheit kommt bei Hegel als Bedingung für das Aufkommen der Philosophie als *freier* Gedanke: „In der Geschichte tritt daher die Philosophie nur da auf, wo und insofern freie Verfassungen sich bilden".[11] Der Staat ist aber keine letzte Entfaltung des Weltgeistes: Die letzte und höchste Stufe des Absoluten ist der absolute Geistes, verwirklicht durch die Kunst, Religion und Philosophie. Das Hegelsche System bekommt mit der Philosophie als der höchsten und letzten Stufe des Systems eine kritische Stellung zum objektiven (und nicht absoluten Welt-) Staat. Der Staat ist allein im Stande die begrenzte Freiheit, die Freiheit der Bürger oder die sittliche Freiheit zu sichern. Während wahre Freiheit etwas Höheres als Recht der Bürger, als Recht überhaupt, ist. Sie ist die Substanz des Geistes, weil „alle Eigenschaften des Geistes nur durch die Freiheit bestehen, alle nur Mittel für die Freiheit sind" (TW 12: 30). An einer anderen Stelle lesen wir: „Eben die Freiheit ist das Denken selbst; wer das Denken verwirft und von Freiheit spricht, weiß nicht, was er redet. Die Einheit des Denkens mit sich ist die Freiheit, der freie Wille [...] der Wille ist nur als denkender frei. Das Prinzip der Freiheit ist aufgegangen, und hat dem Menschen, der sich selbst als das Unendliche faßte, diese unendliche Stärke gegeben" (TW 20: 307f.). In jedem Fall aber entwickeln sich diese höchsten Gestalten der Freiheit innerhalb des Staates, weil er nach Hegel das sittliche *Ganze* in der objektiven Welt darstellt.

Die Thematisierung der „allgemeinen Weltgeschichte" am Ende des objektiven Geistes und im Übergang zum Absoluten hat für die Hegelsche Philosophie eine doppelte Aufgabe. Für das philosophische Denken der *Gegenwart* soll, einerseits der in der Rechtsphilosophie entwickelte Begriff des Staates in seiner modernen Gestalt als *notwendig* legitimiert werden; andererseits aber auch als auf der geschichtlichen Realität gegründet und somit der Zufälligkeit überlassen werden. Als Problem für das zeitgenössische philosophische Denken stellen sich daher die ontologisch-wesentliche Geschichtlichkeit des objektiven Geistes und seine *werdende* Wirklichkeit heraus. In ihrer letzten Gestalt, stellt die *Enzyklopädie* (1830) die *Bewegung* des Weltgeistes mit dem „Weg der Befreiung der geistigen Substanz" (§ 549) in dem Sinne gleich, dass erst die Aufhebung der äußeren Objektivität die wahre *Verwirklichung* der Freiheit darstellt. Diese Befreiung des Geistes, wie man im nächsten Paragraphen lesen kann, verbindet Hegel mit dem Begriff des Geistes des herrschenden Volkes in einer Epoche und verbindet ihn so unmittelbar mit seiner Auffassung der *Geschichte* und der Philosophie, bzw. mit der Geschichte der Philosophie und der Philosophie der Geschichte. Das sagt uns nichts anderes, als

---

der Philosophie an der Geschichte durch die Beschreibung des radikalen hegelianischen Atheismus, nach welchem „Being itself is essentially temporal (Being = Becoming)".

11 TW 18: 117. Vgl. G.W.F. Hegel, *Einleitung in die Geschichte der Philosophie*, hrsg. von J. Hoffmeister, Hamburg 1959, s. 227: „In der Geschichte tritt die Philosophie also da auf, wo freie Verfassungen existieren".

dass erst der Begriff des absoluten Geistes uns die Hegelsche Freiheitslehre und die damit verbundene Geschichtlichkeit des objektiven Geistes aufklären kann.

Die Rechtsphilosophie Hegels erachtet es als ihre Aufgabe, das was ist (*Vorrede*, 25) zu begreifen, wobei das „was ist" als das was dem Begriff der verwirklichten, konkreten Freiheit entspricht, aufgefasst werden soll. Es ist mit dem Selbstbewusstsein der Freiheit aufgesetzt und offenbart sich durch die Französische Revolution und die neuere deutsche Philosophie. Daher kann überhaupt die Rechtsphilosophie Hegels die rechtlichen Institutionen als Daseiende beschreiben, obgleich sie in Preußen nicht voll verwirklicht waren. Soweit aber die konkrete Wirklichkeit der freien Subjektivität sich nur im modernen rechtlichen Staat verwirklichen kann, - so bleibt der Mensch doch in letzter Instanz auf seine Innerlichkeit gestützt. Das Wesen der Freiheit kann sich nicht voll zeigen, weder mittels der Kategorien des Rechts noch in den objektiven Institutionen der äußeren Beziehungen, sondern nur innerhalb der Gedanken. Nach Hegel gibt nur der moderne Verfassungsstaat dem menschlichen Willen, frei (nach außen) und moralisch (nach innen), den wahren Inhalt. Deswegen ist es seit Aufkommen des Christentums die Aufgabe der Zeit, das Prinzip der Subjektivität in die Welt einzubauen. Jedes neue Prinzip, welches seine Verwirklichung in der Weltgeschichte findet, ist nach Hegel immer ein Prinzip der Freiheit, der höheren Freiheit, die sich durch die Knechtschaft befreit.

Diese Dialektik der Freiheit ist immer schon auf das denkende Subjekt gerichtet. Sie legt die Struktur der Bewegung des Seins und des Subjekts dar, worin Hegel die Subjektivität als die Wahrheit der Substanz im Sinn ihres für-sich-sein-kommens denkt. Die Dialektik ist keine Methode, sie stellt nicht unmittelbar auf den Gegenstand ab, sondern sie ist in ihrer Bewegung die Auslegung und die Interpretation des Gegenstands selbst und ist zugleich eine eigene Berichtigung darauf. Sie ist synonym mit dem Geist, d.h. mit einem Ganzen, in dem die endlichen Bestimmungen nur Momente sind (Enz §§ 81, 386 A). Deshalb können wir betrachten, dass der Staat als Wirklichkeit der konkreten Freiheit im postrevolutionären Europa, allein nur eine endliche Bestimmung der Freiheit ist. Denn die Freiheit ist keine bestimmte Eigenschaft, sondern als Bewegung, als Vermittlung mit ihrer selbst – d.i. das bei-sich- selbst-Sein-im-Anderen – damit sie sich auf einem anderen Boden gegenüber der erfahrenden Realität befindet; sie ist das Absolute, „die ewige Unruhe des Begriffs" (TW 2: 487), die zuerst in ihrer Abstraktheit in der objektiven Welt aufkommt. Vom substantiellen-ontologischen Standpunkt sieht man oft voraus, dass die Hegelsche Metaphysik der Freiheit eine Offenheit beinhaltet und diese ist als *die Geschichte einer niemals beendeten Institutionalisierung der Freiheit* sichtbar. Hegel lässt Raum für eine zukünftige Arbeit am Begreifen der Gegenwart der Ewigkeit. Die Richtung der Zeit ist durch geistige Aufgaben bestimmt, welche mit dem Prozess der Wirklichkeitsbildung nicht enden können, sondern stufenweise die *Jenseitigkeit*, auf die die Bewegung der Zeit gerichtet ist, aufstellen. Es ist nicht nur das gültig, was das Denken einer Zeit erfasst, sondern auch das, was noch nicht ist.

Das Verhältnis des Gedankens zur Freiheit, bzw. das Verhältnis zwischen der Philosophie und der Gestaltungen der verwirklichten Freiheit, als der Geschichte zu Grunde liegend, ist in den besonderen Epochen des Überganges und des neuen Anfangs aufgekommen. Darum geht es für Hegel immer um die Epochen in der

Weltgeschichte: Jede endet nämlich mit der Aufhebung der verwirklichten Freiheit und deren Übergang in eine neue Knechtschaft, in eine neue Abhängigkeit vom Herr. Die geschichtliche Veränderung, die da zuerst als *Negation* und weiter als *Vermittlung* aufkommt, nennt R. Bubner den Hegelschen „strategische[n] Kunstbegriff“[12] und fügt mit Recht ein, in Rücksicht auf die Philosophie als das Allgemeine: „Aber es ist nicht das Einzelsubjekt, dessen Räsonnement, Zustimmung oder Protest abgefragt wird. Es ist das Ausmaß tatsächlich gewährter und gesicherter Freiheitserfüllung, an dem der Epochenwechsel hängt“ (ebd. 46). Denn die Freiheit als das Allgemeine, als das *Negative*, die sich immer einen neuen konkreten Inhalt gibt, tritt erst in der Gegenwart in die abstrakte Gestalt (TW 18: 501). Die politische, bzw. objektive Realisation der Freiheit ist damit bereits an-sich beschränkt und Hegel betrachtet noch in seiner späten Jenaer Periode, dass allein das absolute Wissen, die vollendete inhaltliche Bestimmung des Freiheitsbegriffs darstellt, weil die höchste Freiheit das Sich-wissen des Geistes in der Form einer „begriffenen Geschichte“[13] ist.

Das bedeutet nichts anders, als dass der Begriff der *Entzweiung* die gesamte Hegelsche Philosophie beherrscht und auch für seine Rechtsphilosophie, die ihre Geschichtlichkeit insofern zeigt, als „die Geschichte sich in ständig neuen Entzweiungen weiter bilde“, grundlegend bleibt.[14] Es handelt sich nicht um eine Versöhnung der Entzweiung, sondern vielmehr um die Versöhnung *mit* der Entzweiung, die sich wieder in jedem Übergang der Epochen offenbart; insofern zeigt sie sich als die notwendige Folge der, am Anfang abstrakten, Freiheit, wie am Beispiel der Französischen Revolution. Damit ist die politische Gewalt – die zusammen mit der Entstehung der neuen Entzweiung und der Anstrengung sie aufzuheben aufkommt, dass das an-sich zum für-sich wird – ein permanenter Zustand, weil es immer die Möglichkeit gibt, dass die Gewalt, ähnlich wie jeder Krieg, plötzlich ausbricht. Eine Welt die, infolge der Positivität seiner „Institutionen“, zugrunde geht, bietet der abstrakten Freiheit die Macht um die *geschichtliche* Wirklichkeit zu verändern. Deswegen kommt die abstrakte Allgemeinheit zur ihrer Konkretheit, zur Entstehung einer neuen Sittlichkeit. Die Gewalt kann daher nicht aus der geschichtlichen Erfahrung verworfen werden. Die Negativität der Subjektivität zeigt, – gerade am radikalsten in Hegels Auffassung des internationalen Rechts – dass jede konkrete, verwirklichte Freiheit *nicht ohne die Gewalt* werden kann und dass sie nur durch *Individualisation* (in Hegels Zeit durch den besonderen, individuellen

12 R. Bubner, „Hegel am Jahrhundertwechsel“, in: Bubner, R./Mesch, W. (Hg.), *Die Weltgeschichte – das Weltgericht*, Stuttgart 2001, s. 45.

13 TW 3: 591. Vgl. die These von D. Köhler von der „begriffene Geschichte“ als die Realisationsbedingung der Freiheit („Freiheit und Geschichte in Hegels Phänomenologie des Geistes und Schellings Freiheitsschrift“, in: E. Weisser-Lohmann/D. Köhler (hrsg.), *Verfassung und Revolution. Hegels Verfassungskonzeption und die Revolutionen der Neuzeit*, Hegel-Studien Beihefte 42/2000, s. 110ff.

14 Auf diese Weise charakterisiert O. Pöggeler die Hauptrichtung von J. Ritters Interpretation der politischen Philosophie Hegels (O. Pöggeler, „Hegel und die Französische Revolution“, in: E. Weisser-Lohmann/D. Köhler (hrsg.), *Verfassung und Revolution. Hegels Verfassungskonzeption und die Revolutionen der Neuzeit*, Hegel-Studien Beihefte 42/2000, s. 211ff.). Seine Vorlesung über die Ästhetik endet Hegel – während er von der „Befreiung des Geistes vom Gehalt und den Formen der Endlichkeit“ und von der Entfaltung der Wahrheit, die sich in der Weltgeschichte offenbart, spricht – mit der absoluten Subjektivität, die doch „in sich befriedigt“ ist, aber in der Negativität zur objektiven Welt steht (TW Bd. 15: 571-572).

Staat) entsteht. Sodass sich auch der jakobinische Terror in seiner Notwendigkeit aufzeigt, gleich wie die Selbstaufopferung der Bürger im „äußeren Staatsrecht". Die abstrakte Negativität ist die einzige Weise die neue und höhere Form der Freiheit, bzw. das konkrete Allgemeine geschichtlich zu verwirklichen. Das Moment der Negativität ist nicht nur in der politischen Praxis unvermeidlich, - es ist auch konstitutiv für die philosophische Lehre Hegels. Der Widerspruch des Begriffs der (menschlichen) Freiheit ist nicht nur der dialektische Anreger der Geschichte, er ist wesentlich der Widerspruch des Hegelschen philosophischen Systems selbst.[15]

Die Wahrheit der Hegelschen Rechtsphilosophie und der Ort der wahren Freiheit ist lediglich die *Philosophie* als das *Verstehen* des Ewigen in der Gegenwart, so wie die einzige wahre Aufhebung die Aufhebung im Gedanken ist. Wenn diese Aufhebung, wie jeder geschichtlicher Epochenübergang zeigt, als Konsequenz blutige Konflikte hat, die diesen Gedanken in der Objektivität des Geistes verwirklichen, stellt es zugleich den Untergang und die Unwahrheit dieses Gedankens selbst dar. Der neue wahre Gedanke wird geboren und sucht seine Vollendung in der veränderten Objektivität des Geistes: „In solchen Zeiten, wo die politische Existenz sich umkehrt, hat die Philosophie ihre Stelle; und dann geschieht es nicht nur, dass überhaupt gedacht wird; sondern dann geht der Gedanke voran und bildet die Wirklichkeit um. Denn wenn eine Gestalt des Geistes nicht mehr befriedigend ist, dann gibt die Philosophie ein scharfes Auge dazu, dies Unbefriedigende einzusehen".[16]

Die Frage der politischen Freiheit – vor 2000 Jahren auf der Agora geboren – zeigt uns die vielschichtige Dialektik, die diese Frage in sich trägt. Es ist nicht nur die Frage von der Beziehung zwischen dem Allgemeinen und dem Einzelnen oder von der Autonomie des Subjekts innerhalb der Gewalt (der bürgerlichen Gesellschaft) und der freien Entscheidung des Einzelnen (oder des „Zwangs" eigener Natur, was der Mensch an sich ist) sich mit dem Anderen (dem Staat) in Harmonie wiederzufinden; - diese Frage der politischen Freiheit, auch die Negativität und die Natürlichkeit, in die sie durch die Verhältnisse der Staaten untereinander abstürzt sagt uns, dass man die wahrhafte Verwirklichung der Freiheit irgendwo anders suchen muss.[17] Diese Frage – solange sie noch durch die Orientierung der Staaten am Vorrang der Subjektivität zu lösen ist – lässt sich weder lösen, noch wird sie sich je lösen lassen, weil diese Lösung nur, einerseits, zu Interventionskriegen, andererseits, zum Quietismus und zum Rückzug vor der Wirklichkeit führen kann.

---

15 In einer der dunkelsten Passagen seiner gesamten Schriften, schreibt Hegel im Zuge seiner Behandlung des Begriffs der Negativität: „Der Mensch ist diese Nacht, dies leere Nichts, das alles in ihrer Einfachheit enthält – ein Reichtum unendlich vieler Vorstellungen, Bilder, deren keines ihm gerade einfällt –, oder die nicht als gegenwärtige sind. Dies die Nacht, das Innere der Natur, das hier existiert – r e i n e s S e l b s t, – in phantasmagorischen Vorstellungen ist es rings um Nacht, hier schießt dann ein blutig Kopf, – dort eine andere weiße Gestalt plötzlich hervor, und verschwinden ebenso – Diese Nacht erblickt man, wenn man dem Menschen ins Auge blickt – in eine Nacht hinein, die furchtbar wird, – es hängt die Nacht der Welt hier einem entgegen" (G.W.F. Hegel, *Jenaer Systementwürfe III*, hrsg. von R.P. Horstmann, Hamburg 1987, s. 172).

16 *Einleitung in die Geschichte der Philosophie*, s. 286.

17 Ich habe argumentiert, dass Hegel wegen dieser Einsicht, die Verwirklichung der Freiheit in das Gebiet des absoluten Geistes überführt hat, den einzigen Ort der wahren Vereinigung der Menschen. Das ist zugleich der Ort der Morgenröte des Werdens.

Bibliographie:

Aristoteles (1969), *Politik*. Übers. von O. Gigon, Stuttgart, 1961.
Avineri, S. (2003), *Hegel's Theory of the Modern State*, Cambridge.
Bubner, R. (2001) „Hegel am Jahrhundertwechsel", in: Bubner, R./Mesch, W. (Hg.), *Die Weltgeschichte – das Weltgericht*, Stuttgart.
Derrida, J. (1989), „Force of law: the 'mystical foundation of authority'", *Cardozo Law Review* 920.
Hegel, G.W.F. (1959), *Einleitung in die Geschichte der Philosophie*, hrsg. von J. Hoffmeister, Hamburg.
Hegel, G.W.F. (1969ff.), *Briefe von und an Hegel*, Bd. 2, hrsg. von J. Hoffmeister, Hamburg.
Hegel, G.W.F. (1970ff.), *Werke in zwanzig Bänden* (=TW), auf der Grundlage der Werke von 1832-1845 neu edierte Ausgabe, hrsg. von E. Moldenhauer, K.M. Michel, Frankfurt a.M.
Hegel, G.W.F. (1983), *Die Philosophie des Rechts. Die Mitschriften Wannenmann (Heidelberg 1817/18) und Homeyer (Berlin 1818/19)*, hrsg. von K.H. Ilting, Stuttgart.
Hegel, G.W.F. (1987), *Jenaer Systementwürfe III*, hrsg. von R.P. Horstmann, Hamburg.
Ilting, K.H. (1963-64), „Hegels Auseinandersetzung mit der Aristotelischen Politik", in: *Philosophisches Jahrbuch* 71.
Jaeschke, W. (1995), „Die Geschichtlichkeit der Geschichte", in: *Hegel-Jahrbuch*.
Kant, Immanuel (1902ff.), Gesammelte Schriften, Berlin: Akademie-Ausgabe
– *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*, Band VIII.
– *Zum ewigen Frieden*, Band VIII.
Köhler, D. (2000), „Freiheit und Geschichte in Hegels Phänomenologie des Geistes und Schellings Freiheitsschrift", in: E. Weisser-Lohmann/D. Köhler (hrsg.), *Verfassung und Revolution. Hegels Verfassungskonzeption und die Revolutionen der Neuzeit*, Hegel-Studien Beihefte 42.
Kojeve, A. (2000), „Tyranny and Wisdom", in: L. Strauss, *On Tyranny*. Including Strauss-Kojève Correspondence, Chicago.
Perelman, C. (1966), „Was der Philosoph vom Studium des Rechts lernen kann", in: *Wissenschaft und Weltbild. Zeitschrift für Grundfragen der Forschung*, Nr.4.
Pöggeler, O. (2000), „Hegel und die Französische Revolution", in: E. Weisser-Lohmann/D. Köhler (hrsg.), *Verfassung und Revolution. Hegels Verfassungskonzeption und die Revolutionen der Neuzeit*, Hegel-Studien Beihefte 42.

Rastko Jovanov

Između filozofije i države: Hegelova dijalektika institucionalizovanja slobode

Abstrakt

Hegel razmatra u svom sistemu filozofije različita određenja slobode; on razlikuje subjektivnu, objektivnu i apsolutnu slobodu. Mene će u ovom radu prvenstveno zanimati dijalektika objektivne slobode, koju Hegel uvodi u okviru svoje Filozofije prava, da bih potom ukazao na problematičnost same istoričnosti objektivne slobode, te tvrdio da pojam slobode tek na nivou apsolutnog duha zadobija kvalitet istinske istoričnosti. To će mi omogućiti da u argumentaciji otvorim prostor za iznošenje teze o dijalektičkom jazu koji je prisutan u Hegelovom shvatanju dovršenosti slobode na dva različita nivoa njegovog sistema, u državi kao dostizanju konkretnosti slobode u području objektivnosti duha, kao i u prividno apolitičkom pojmu slobode u sferi apsolutnog duha, odnosno sferi konkretnog mišljenja, sferi same filozofije.

Ključne reči: Hegel, sloboda, država, istorija, apsolutni duh

Between Philosophy and State: Hegel's Dialectic of the Institutionalization of Freedom

Abstract

Hegel considers, in his system of philosophy, different specifications of freedom; he distinguishes between subjective, objective and absolute freedom. I am interested, in this paper, primarily in the dialectics of objective freedom, which Hegel introduces in his Philosophy of Law, in order to point out the problematics of the historicity of objective freedom, and to argue that the concept of freedom gains the quality of true historicity only at the level of the absolute spirit. This will allow me to open the space, within my argument, for presenting the thesis about the dialectical gap which is present in Hegel's understanding of the perfection of freedom at two different levels of his system: in the state as attaining the concreteness of freedom in the domain of the objectivity of the spirit, as well as in the apparently apolitical notion of freedom in the sphere of the absolute spirit, that is, in the sphere of concrete thinking, the sphere of philosophy itself.

Keywords: Hegel, Freedom, State, History, Absolute Spirit

III

BOOK SYMPOSIUM

SEMINAR O KNJIZI

BOOK SYMPOSIUM ON *KANT'S POLITICAL LEGACY: Human Rights, Peace, Progress* (Luigi Caranti) / Edited by Bojan Kovačević

SEMINAR O KNJIZI *KANTOVO POLITIČKO NASLEĐE: Ljudska prava, mir i progres* (Luiđi Karanti)

UDK: 141.7 Kant I.
https://doi.org/10.2298/FID1804567K
Original Scientific Article
Received: 21.06.2018. Accepted: 04.08.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Bojan Kovačević

SUBJECTIVE UNIVERSALITY OF GREAT NOVELISTS AS AN ARTISTIC MEASURE OF HISTORY'S ADVANCE TOWARDS ACTUALISING KANT'S VISION OF FREEDOM

ABSTRACT

The main idea behind this article is that in order to understand the meaning that Kant's political philosophy is rendered to by the given socio-historical context of a community we need to turn for help to artistic genius whose subjective "I" holds a general feeling of the world and life. It is in this sense that authors of great novels can help us in two ways. First, their works summarise for our imagination artistic truth about man's capacity for humanity, the very thing that Kant considers to be the scientifically improvable "fact of reason". Second, works of great writers offer for our insight destinies of individuals who decide to pursue moral dictate in a society, thus actualising the potential that lies hidden in all of us, making us worthy of respect. As we lack objective scientific standard of measurement, artist's universal feeling of the world is impressed upon us through a narrative about a man who, in a given society and in a given moment, decides to exercise his autonomy and seek the divine in himself. Contemporary social scientists' attempts to prove historical progress is characterised by the very lack of humbleness. Referring to the great novelists' works in this article is aimed to remind scientists of restraint and self-control demanded from them by the citizen of Konigsberg.

KEYWORDS

Kant, novel, freedom, peace, Tolstoy, genius, science, progress, history

History as Man's Moral Progress or an Eternal Search for Instruments of Passive Neutralisation of Social Conflict

Kant's hope for historical progress was inspired by his observation of the French Revolution.

> The revolution of a spirited people that we have witnessed in our times may succeed or fail. It may be so filled with misery and atrocities that any reasonable person, if he could hope, undertaking it a second time, to carry it out successfully, would nonetheless never decide to perform the experiment at such a cost.—Nevertheless, in the hearts of all its spectators (who themselves are not involved in the show), I assert, this revolution meets with a degree of *sympathy* in wish that borders on enthusiasm, a sympathy the expression of which is itself associated with danger. This sympathy can thus have no other cause than a moral capacity in the human race (Kant 2006: 155).

Bojan Kovačević: Associate Professor, Faculty of Political Sciences, University of Belgrade; bojan.kovacevic@fpn.bg.ac.rs.

One should bear in mind the distinction between being an actor and being a spectator, between engaging and judging, which Kant held important (Arendt 1992). Being an actor and being a spectator are based on two principles that differ in their essence. The principle of action tells the one engaged, and thus incapable of having a comprehensive insight into the consequences of one's actions, that no revolution, albeit one against the worst of tyrants, is justified. Spectator, on the other hand, must stay outside the game, impartial. Contrary to actor, spectator must remain autonomous in regard to the event. Kant observed affairs in France from the standpoint of a distanced spectator making a judgment about the event in a foreign country. It made him feel the true pleasure of man in face of a great work of art. "It is simply the spectators' mind-set, which reveals itself *publicly* in the face of this show of large-scale transformations and which makes known such a universal and yet unselfish sympathy with the players on the one side against those on the other, even at the risk that this partiality could become quite detrimental to them"(Kant 2006: 155). Convinced that his feelings of sympathy for revolutionaries aroused by this historic turmoil is shared among all spectators gifted with good taste, Kant concluded that the significance of the French Revolution lies in the fact that it serves as the historical indicator of the "moral tendency of the human race" (Kant 2006:155).

The world-historic events in France that inspired faith in historical progress in Kant were seen as an introduction into a continuous civil war by some other spectators. "Once virtue enters the arena of political action, then the moral dualism that, within the framework of the existing State, had guided the indirect assumption of power and made possible an overweening criticism, automatically justifies civil war. Civil war is an innocuous occurence. Although it does lead to violence and murder, it is none the less shaped by political criticism" (Koselleck 1988: 180). Instead of Catholics and Protestants, today's civil war actors are liberals, socialists, conservatives, nationalists, social-democrats who have been provided with an excuse to assert their views of the world, politics and society upon the rest of the community by the outcome of the Revolution. As opposed to the war led by those participating in the St. Bartholomew's Day massacre on the 1572 bloody Parisian August eve, in the civil war led under the guise of modern democratic state there is no blood and violence. Why? Intoxicated by Rousseau's appealing idea of *volonté générale* the French had decided to get rid of the monarch and destroy the Old Regime's institutions, but this fact alone did not make them up to the task of making an active political decision about their collective destiny (Burke 1910). The Revolution failed to lead the citizenship from the state of self-incurred immaturity, yet it paved the way for appearance of new mechanisms for depoliticisation and neutralisation of social conflicts as a substitute for an absolute monarch's sovereign decision. Passive neutralisation of social conflicts is focused on obscuring the issues of the source and origin of politics (Schmitt 2005, Schmitt 2007). Memories of the emergency, contingency, decision are to be sedated by liberalism, parliamentarism, reducing politics to administration, positive laws, ideology, material prosperity and civic security. Seen from this spiritual-historical point of view of decisionism, Kant's faith in historic progress is yet another instrument of depoliticisation and neutralisation of social conflicts.

These two opposed positions in regard to the French Revolution rest upon divergent views of man's nature.

Kant speaks of man in different terms (Arendt 1992: 22–27). In his ethics, man is a being with an inherent moral law and potentiality of acting in accordance with duty. "An action is said to be in accordance with duty only when every thought of advantage to be expected from it, every calculation of present or future pleasure likely to result from it, indeed every material aim of any other kind, is eliminated and only adherence to the universality of the law, which reins in all contingent and particular impulses, remains as the sole ground of determination" (Cassirer 1981: 244). However, when in his political writings he discusses man as part of humankind, Kant actually attempts to define conditions that could inspire people to act in accordance with the laws of practical reason. Man as part of humankind is seen by Kant as selfish, greedy, motivated by increasing material wealth, but at the same time possessing a capacity to learn from his own mistakes. "Those always who have their dear self before them as the sole focal point of their efforts and who attempt to make everything turn on the great axis of selfinterest are the most common, and nothing can be more advantageous than this, for these are the most industrious, orderly, and prudent people; they give demeanor and solidity to the whole, for even without aiming at it they serve the common good, supply the necessary requisites, and provide the foundations over which finer souls can spread beauty and harmony" (Kant 2011: 34). As a result of the "unsocial sociability" mechanism, this selfish man gifted with reason gradually perfects his institutions whose backlash teaches him to distance himself from his self-interest and refine his ambition.

The notion of man possessing a capability to learn from historical mistakes and perfect his social institutions is foreign to those spectators who see French revolutionaries' enlightened promises as nothing more than a dangerous illusion (De Maistre 2006). In the world where the church has lost its monopoly on interpreting the image of the world, there is no one who can tell man's real nature. On the ruins of medieval order grew a world marked by radical lack of meaning. In such a world, citizens will never be up to the task of deciding on common issues through public discourse. To convince them that the foundations of their communality can be determined by everyday plebiscite (Renan 1996) meant letting their ambitions run wild, giving them false hope, which always results in eventual violence and disorder. Peace and security that absolutistic monarchy once gave them together with the freedom to contemplate the issues of good, beautiful and just in the quietude of their private sphere is the most that citizens can expect from history.

Even Kant himself believed that citizens are not to be given the right of active re-examination of postulates underlying obedience to the state immediately. "For many affairs that serve the interests of the commonwealth a certain mechanism is required, by means of which some members of the commonwealth must play only a passive role, so that they can be led by the government in the pursuit of public ends by means of an artificial unanimity, or at least be kept from undermining these ends" (Kant 2006: 19). Should a tax payer publically question his obligation to pay the tax, a civil servant challenge the grounds of his superiors' orders, and priests refuse to adhere to church rules, that would certainly jeopardise the order, warned cautious Kant. Order must be stable, however, to provide scientists with an environment where they can contemplate in peace and present their insights to the literary public. This kind of freedom that Kant calls the "public use of reason"

must not be limited in any way. It is preconditioned by peace based on civil obedience to sovereign's orders (Ciaran 2003). And vice versa, the unlimited public use of reason is necessary so that the decisions of those who decide upon common good would gradually grow to result in such reform of social institutions that would be aimed at actualisation of republican ideals. Refinement of political institutions will be followed by a development of civic ability to decide upon common interest issues in the spirit of republicanism. Over time, man will learn to observe the problems of communal living independent of his confession-, class-, profession- or nation-specific interests. Man's state of not being engaged in the game will extend to envelop more domains of social life and spill over to wider and wider communities. Thus, the idea of humankind, present in each one of us, will gradually become the principle not only of our judgements but of our actions as well, and actor and spectator will become united (Ardent 1992: 75).

Today's authors who have given credence to Kant's vision of nearing eternal peace point to empirical indicators of historical progress. "Liberal democracies – the political systems we have are closest to Kant's republics – are wealthy and peaceful (at least towards one another). International organizations (global and regional) have been developed and play a role that was unthinkable in Kant's time. A culture of human rights is rapidly taking hold of the global community. These may not be good enough 'signs' of the direction in which we are moving for the sceptic, but it is quite likely more than Kant himself would have hoped for two centuries ago" (Caranti 2017: 230). Others, however, who have conjoined Kant in his belief that the principles of democracy and human rights are so attractive that ultimately no people will be able to resist the temptation, have reached completely different conclusions by observing legal and political tendencies of our time (Maus 2015). They warn of the changes that constitutional democracies and international organisations are going through in the globalised world, showing features of historical retrogression and deviation from the path delineated by Kant in his political writings.

Instead of deciding which of the two opposing positions to embrace, this article questions the very grounds of attempting to measure historical progress scientifically. It is erroneous to expect objective empirical proof of humankind's progress from social scientists as they are not up to this task. Powerless to take a scientifically unbiased position on this matter, social scientists who in our time seek evidence of historical progress often unawares charter their ethics to serving the purpose of preserving the existing order, taking on the role of "political moralists". Kant viewed the French Revolution as a great artwork of history. Willingness of people to sacrifice their safety, property and lives to translating the ideas of freedom from philosophical books into history arouse in him a feeling of sympathy and inspired his reveries on eternal peace. On the contrary to this, today's scientists see empirical evidence of Kant's dream coming true in proliferation of externally liberal features of national and international institutions.

The main idea behind this article is that in order to understand the meaning that Kant's political philosophy is rendered to by the given socio-historical context of a community we need to turn for help to artistic genius whose subjective "I" holds a general feeling of the world and life. "The work of art is something singular and apart, which is its own basis and has its goal purely within itself, and

yet at the same time in it we are presented with a new whole, and a new image of reality and of the mental cosmos itself" (Cassirer 1981: 307). It is in this sense that authors of great novels can help us in two ways. First, their works summarise for our imagination artistic truth about man's capacity for humanity, the very thing that Kant considers to be the scientifically improvable "fact of reason". Second, works of great writers offer for our insight destinies of individuals who decide to pursue moral dictate in a society, thus actualising the potential that lies hidden in all of us, making us worthy of respect. As we lack objective scientific standard of measurement, artist's universal feeling of the world is impressed upon us through a narrative about a man who, in a given society and in a given moment, decides to exercise his autonomy and seek the divine in himself thus becoming "qualified members, or, perhaps more modestly, qualified applicants to another kingdom" (Caranti 2017: 27). Can society accept such a member? Does history really change in this respect at all?

In the works of artists belonging to different times and societies we find conclusions that are contradictory to those arrived at by today's social scientists lost in their quest for objective indicators of historical progress. Writers of great novels tell us that society cannot tolerate an individual who decides to act in accordance with duty or questions with his reason the social norms, laws and customs built into the foundations of the existing order. To defend its rules, society banishes or condemns to a tragic end the one who dares to choose the freedom to seek meaning despite grave consequences of his decisions. Great novelists manage to recognise Kant's philosophical idea of human nature that contains the kernel of potential for a meaningful development towards laws of freedom in the concrete, individual and particular. Hence their recurring rebellion against society that denies to man of their era a possibility to achieve humanity, reducing him to the banal, empirical, earthly, immersed in calculating the costs and benefits of his acts.

Artistic genius reports with sorrow that history so far confirms the insights of philosophers belonging to the spiritual-historical sphere of decisionism. In the civil society whose foundations were set in the late 18th century by the French and American Revolutions, man has failed to polish the facets of his ambitions against historical mistakes. Stability in modern states survives as long as citizens, lulled by economic prosperity and semblance of constitutional democracy, cheerfully accept the condition of self-inflicted immaturity and witlessly adapt to the existing social norms. Yet, when loss of civil security and material wealth shake them awake from the state of passive apathy forcing them to pose the question of political power's source and origin, violence, chaos and non-order of the natural condition return to the stage once again.

On Curability of Human Contingency

"Is there a remedy for the contingent state of man? Is his life incurably accidental, as Lucretius thought and as existentialists maintain today, or has man, despite his duality, preserved some discoverable link with non-accidental and non-contingent Being, so that he may entertain a hope for self-identification? Or, in other terms, is he summoned or destined to return a state of completeness and non-contingency?"

(Kolakowski 1978:12–13) This question has been nagging at philosophers since time immemorial. Philosophical trouble is rooted in the assumption, hope, faith, that man is not what meets the eye, that there is more to him than merely his banal empirical existence. It is a faith that there is something in man that can elevate him to divine heights.

The transition from medieval to modern epoch was accompanied by a triumph of the idea that man is the sole creator of what he is to become in this world. The idea that man is "the molder and maker of thyself; thou mayest sculpt thyself into whatever shape thou dost prefer" (Mirandola 1998: 5) has exited humanistic thinker Pico della Mirandola. It is the fact that God left it to man to choose his own appearance and gifts for his adornment that makes man admirable. This Adam who can choose to crawl the earth like an animal or become a philosopher, even "a divinity clothed with human flesh" (Mirandola 1998: 6), for Mirandola is a miracle. The human capacity to determine his own unsteady, changeable and varied nature is the foundation of man's dignity. "The seeds that each man cultivates will grow and bear their fruit in him. If he cultivates vegetable seeds, he will become a plant. If the seeds of sensation, he will grow into brute. If rational, he will come out a heavenly animal. If intellectual, he will be an angel, and a son of God. And if he is not contented with the lot of any creature but takes himself up into the center of his own unity, then, made one spirit with God and settled in the solitary darkness of the Father, who is above all things, he will stand ahead of all things" (Mirandola 1998: 5).

Kant spent his philosophical life seeking the answer to the question of what man can know (Kant 1998). The quest led him to conclude that contingency of human existence is impossible to overcome. Cognition cannot overcome the inexplicable experientiality of facts, we can merely acknowledge it, it is a given. Any attempt at enveloping wholeness of the world with thinking unavoidably leads to antinomies of the mind. Nonetheless, at the same time, man cannot resist thinking about that which is in the focus of his interest, yet incomprehensible – about God, freedom and the immortal quality of soul. Hence, for Kant these ideas become the unattainable limit that must be the goal of our ambitions. They point to where the endless road of self-development that man treads in search of the divine in him leads to. "Kant opens a new chapter in the history of philosophy's attempt to overcome the contingency of human existence, setting up freedom as man's realization and establishing the independence of the autonomous reason and will as the ultimate goal of man's unending pilgrimage towards himself, a self that will then be divine" (Kolakowski 1978: 50).

Unwillingness to accept the banal contingency of human condition will determine the direction and result of the search for an idea, which would, following in the footsteps of Mirandola and Kant, provide a philosophical justification of attachment to human rights in today's world (Caranti 2017: 57–104). Endeavour to provide a firm base for human rights in the notion of human dignity have resulted in shifting the discussion from the field of geopolitics and legal positivism to the field of philosophy, which is its greatest merit. Human rights are protected by the existing national and international treaties because of the overwhelming conviction that humans have dignity. "We assume that one of the major tasks of philosophy,

applied to the tricky field of human rights, is the attempt to spell out what lies behind the intuition – taken for granted in all major human rights treaties – that humans have dignity" (Caranti 2017: 62). Although inspired by the work of the great humanistic thinker, for the new concept of dignity Mirandola's proposition that it is his chameleon-like nature that earns man his respectability is insufficient. Man deserves his rights to be protected not because he is left with an ability to choose his own life path, but because moral decision is always close at his hand regardless of the circumstances surrounding him (Caranti 2017: 61). Man is a being worthy of dignity due to his capacity to "silence all natural impulses, even the strongest instincts of survival, and act from our conception of duty" (Caranti 2017: 57).

Such foundation of human rights is inspired by Kant, yet it is not Kant's. Indeed, the novelty in comparison to Kant introduced by this attempt is that duty does not necessarily result from the too strict and rigid categorical imperative, but from a version of moral law. "For example, one may believe that morality's source is in Good and still adhere to divine commands not out of fear of divine punishment or similarly heteronomous motives, but because one endorses those commands and makes them truly one's own" (Caranti 2017: 64). What is important is man's cognition that we have a certain commission, absolutely independent from what this commission can do for our lives. Man's capacity to act following dictates of duty is the very thing that comprises dignity. It is the thing that makes human beings worthy of respect.

In Search of an Artistic Proof of Man's Capacity to Find the Divine in Himself

What links the philosophical attempt to provide foundations for human rights in today's world with Mirandolo and Kant is the belief that in man lies a hidden purpose, whose actualisation can lead him to the point where the individual and the universal, freedom and necessity, reconcile. "The true philosopher does not accept the conditions under which life has been given to man" (Arendt 1992: 22). Besides philosophers, this feeling of human purpose is also ingrained in artists, authors of great novels included.

The transition from the epoch of epic into the modern epoch of novel was made at the moment when the medieval worldview, which maintained that human evil nature was determined by the original sin, crumbled. "The novel is the epic of an age in which the extensive totality of life is no longer directly given, in which the immanence of meaning in life has become a problem, yet which still thinks in terms of totality." (Lukacs 1971: 56) Man who, hoping for redemption and earning his place in heaven, lives in compliance with external dictates of the church was replaced by Mirandola's chameleon with an inner potential for both divine and animalistic. It was only this kind of man that could become a hero of novel. Writers are not interested in the banal kinds who, out of habit, fear or indifferent dullness, submit to small-town rules of living, whose predictable lives can be subsumed under the laws of experiential perception. Novel as a work of art holds "a specific sharpness, a gravity all its own, in face of which nothing that is merely lifelike--which is to say nothing that is dramatically trivial--can survive" (Lukacs 1971: 57).

Artistic genius is inhabited by a spiritual archetype, a subjective feeling of purposefulness against which experiential facts of the world are formed. Kant's work demonstrates that it is impossible to envelop the wholeness of the world experience scientifically, with theoretical mind, by thinking. At the same time his insight reveals that artistic genius is gifted with the capacity to bring forth the principle of historical development upon which the entirety of nature rests in the singular and the particular (Kant 2000: 187–197). "Genius is the talent (natural gift) that gives the rule to art. Since the talent, as an inborn productive faculty of the artist, itself belongs to nature, this could also be expressed thus: Genius is the inborn predisposition of the mind (*ingenium*) through which nature gives the rule to art" (Kant 2000:186). Therefore, art is quite a special kind of connecting singular with the whole. Artist's talent enables him to find the way to present fruits of his imagination to all people who are given the gift of good taste. When man finds himself in front of a work of artistic genius he remains within his own self but feels at the same time relieved from all contingency, as an agent of a universal feeling. Thus, artistic genius manages to capture the very thing that eludes scientist. Artist succeeds in revealing the secret and the power of "universal communicability" thus conveying to spectators his intuition of the wholeness of the world, encapsulating it and feeding it to our imagination. "Genius and its act stand at the point where supreme individuality and supreme universality, freedom and necessity, pure creation and pure lawfulness indissolubly coalesce" (Cassirer 1981: 321).

This article inspired by the philosophical attempt to position the notion of dignity in the core of human rights could not feature any random writer, but the one who matches Kant's definition of genius, Leo Nikolayevich Tolstoy. This literary great wrote the novel "Resurrection" (Voskreséniye, 1899). It is a book about the nobleman Nekhlyudov, who lazily gets up from his bed, listlessly washes his face in his luxurious bathroom, puts on his elegant clothes with boredom, and reads, with resigned contempt, a letter from a wealthy countess who wants to lure him into a promise of marriage. His spirit, weary of meaninglessness of living, remains insensitive to joys of the beautiful spring day while leaving for the court, in bad humour, where he is called to serve as a juror in a hearing of a prostitute who is accused of murdering a brutish patron of hers.

However, once he sees the face of the wretched girl, the soul of the lazy nobleman disappointed with life is shaken from the very bottom. He recognises the girl who used to work on his aunt's estate and whom he, a handsome wealthy young lord, seduced and took before leaving to go to the army once upon the time. The unfortunate accused Maslova was left with a child after that night, losing the child to illness, while her life went astray, landing her in brothel where she was brutalised by miserable men of various ages and affinities, until the moment when she fell a victim of deceit and, unaware of what she was doing, poisoned a patron, consequently ending up in prison accused of a cruel murder. The revelation of the woman's sad destiny removes the layers of selfishness, self-absorbedness, laziness and apathy covering Nekhlyudov's divine sparkle. Suddenly he catches a glimpse of something beautiful in his own self. The decision to help poor Maslova leads him more and more to become aware of the light still simmering within. He gives a fortune on lawyers, does his best to keep her out of prison; the fact that she is

just disdainfully repulsed by him does not deter him from his intention to follow her to Siberia to help her, tend to her, protect her... In the inhumane conditions of prison life, far away from society and its banal rules, Nekhlydov manages to reach the heights of the divine in himself.

On the Republic yet to Be Created

So, Mirandola, Kant and Tolstoy believe that man has the capacity of making a moral decision. This conviction serves as the basis of philosophical attempt to found human rights on man's dignity. We are assured of the existence of that which is improvable, not the subject of scientific knowledge, by philosophers and Tolstoy in different ways. Awareness of the fact that every one of us possesses "an inescapable authority of the moral law" (Caranti 2017: 82) for Kant is a "fact of reason". Would we be capable of doing the right thing if monarch forced us by a death threat to give false testimony against an innocent man? The answer to this question is impossible to give beforehand, but what Kant finds certain is that each one of us would know what the right thing to do is and that it could be done (Caranti 2017: 82). This thought experiment is not a scientific proof, as there is no proof that moral decision is close at hand to each one of us in all situations. Yet, Tolstoy tells us in "Resurrection", by his ability to impress the universal view of man and society upon our imagination intuited by his subjective feeling of the world, that the basic proposition of Kant's ethics is no illusion after all.

At first sight, however, it can appear that Kant and Tolstoy arrive at different conclusions about the conditions necessary for man's pursuit of freedom. The society that Tolstoy knows, the society of banal conventions, hypocrisy and boredom, just buries ever deeper Good's light given to him at birth. "The aimlessness and insubstantiality of the life he describes expresses itself not only objectively, for the reader who recognises it, not only as the lived experience of gradual disappointment, but also as an a-prioristic, established, agitated emptiness, a restless ennui" (Lukacs 1978:149). Nekhlyudov gains his dignity in the train to Siberia, travelling together with prisoners convicted for the most serious crimes. Can people who have buried their light deep inside, who, unlike children, birds and insects, are incapable of recognising the beauty of a spring day, judge, punish and imprison each other according to their own laws at all without causing an even greater evil? In War and Peace (Voyná i mir, 1869), Bolkonsky discovers the truth about the beauty of the world mortally wounded on the battlefield, while hungry and barefoot Pierre Bezukhov finds it in the image of the simple peasant Platon Karataev in prison. Tolstoy's man succeeds to find humanity only by escaping the world of conventions, society and culture.

As opposed to Tolstoy, Kant and his interpreters today believe that peace and stability of civil society offer encouragement for a gradual moral perfection of people. However, one must exercise caution here. Kant does not claim that man can act in accordance with duty only in a certain civil society. Quite on the contrary, Kant and his philosophical disciples maintain that each man in every situation can find humanity in himself by following the categorical imperative. Even in inhumane circumstances, man still has the freedom to use his reason and make

a decision independent of the direct or indirect interests. "An individual can be autonomous even if she is deprived of external freedom. Think of the case of a slave" (Caranti 2017:30). Thus, the difference between the great philosopher and genius writer lies not in their divergent understanding of the society and man of the time they lived in. On the one hand Tolstoy says: the society I know does not encourage man to perfect his nisus. Moreover, this society eradicates humanity in man. The decision following the dictates of duty unavoidably means an escape from society, rejection of culture and return to nature, harmony, consonance and simplicity. On the other hand, Kant argues in his political writings that society is capable of changing, developing, of perfecting its institutions. With history's advancement towards actualisation of the ideals of the republic, society will generate more and more impetuses for actualisation of human potentials. In other words, writers speak of a specific society in a specific moment. Reflecting on specific in them spawns subjective feeling of the universal. Therefore, it would be wrong to wonder which one is right – Kant or Tolstoy.[1] Kant writes about a republic that is yet to appear. The philosopher writes about a process, a journey. From novelists writing in different epochs we can learn how far we have gone on this journey. Kant tells us how far we could go.

The hope of the moral progress of humankind has informed the political writings of Köngisberg-based philosopher. At the beginning there are just republican institutions that keep in check the destructive passions of individuals. In order to establish such institutions no moral transformation of people is necessary. Institutions that set selfish desires, interests and attitudes of individuals against one another, preventing them to take over the place of public good, would be suited equally to the devil's people. A bad man establishes a state and becomes a good citizen after that. Moreover, perhaps human selfishness and pursuit of self-interest is necessary for creating a base one day "over which finer souls can spread beauty and harmony" (Kant 2011: 34).

Nevertheless, as opposed to the American Founding Fathers, who held that man's selfishness and egocentrism are ineradicable, Kant proposes a possibility of moral perfection of men. In civil society individuals gradually perfect their ambitions. "The justice of institutions gradually permeates individuals' souls, and they in turn adhere more authentically and steadily to the principles on which their government is based, thereby generating further institutional progress" (Caranti 2017: 126). Citizens gradually come to learn to distance themselves from their own self-interest and view problems through the eyes of others. When war disasters and suffering finally teach entire nations to see problems through the eyes of others, they enter into agreements and gradually transition into the state of eternal peace, preserving their republican institutions.

Like freedom, Kant's republic is indeed an ideal to be constantly sought but never completely attained. "No existing republic can be satisfied with the level of normative and institutional development achieved" (Caranti 2017: 201). The same goes for the idea of citizen. To be a citizen in Kant's republic means to be able to distance ourselves from our own interests and implies a high level of freedom from indirect and direct pressures. Kant's citizen is the one in whom actor engaged in the

1 I am thankful to Caranti for this important suggestion.

game and unengaged spectator coalesce, the one who feels at home in the world. They are still non-existent in the empirical reality of Kant's time. Man's lifespan is too short to actualise all of his potentials. A full development of the seeds planted in man by nature can be achieved only at the end of the historical road delineated by Kant in his essay "Idea for a Universal History from a Cosmopolitan Perspective" (Kant 2006: 3-16). In it, Kant attempts to reveal natural purpose in the senseless current of human reality.

Hannah Arendt notices a tone of irony in Kant's political writings, drawing the conclusion that Kant himself did not take them too seriously (Arendt 1992: 7). She reminds us of the fact that Kant called the most important among them, Perpetual Peace, "reveries" (Arendt 1992: 7). The German philosopher discovered significance of the political as distinguished from the social rather late in life, when he was woken from his snooze by the French and American Revolutions. Therefore, his political writings should be understood more like "play with ideas" or "mere pleasure trip" (Arendt 1992: 7) than as serious theses that the social science of the future is to prove.

Today's philosophers dedicated to study of Kant's political thought disagree with that (Caranti 2017: 207-256). What our time requires is to clean Kant's plan of the human history's development from dogmatic deposits. Once this is done, revealed before us will be the reasons why "non-linear progress towards the cosmopolitan constitution, rather than regress or stagnation, is the most likely development of human affairs" (Caranti 2017: 210). This thesis rests upon the faith that people are relatively benevolent and capable of learning from experience over time to understand their true interests. Owing to the mechanism of "unsocial sociability" people will, in time, learn lessons from their social conflicts, which will force them to reform their national and supranational institutions towards actualisation of the republic and eternal peace. Does the expansion of institutions of representative democracy and strengthening of the international institutions' power truly mean that humankind has progressed and that citizens have learned their lessons from wars, misery and conflicts? Is Carl Schmitt right when he argues that institutions of liberal democracy are instruments of passive neutralisation of social conflicts that can only postpone, more or less, the return of chaos, violence and wars on the stage of history? Do peace and stability in civil society, contrary to what Kant and those who continue in his footpath today think, require an individual in the state of self-inflicted immaturity, blind adherence to the decisions of authorities, national narrow-mindedness and lack of interest in the world problems? Do we still live in the same civil society whose artistic truth was offered to us by Tolstoy, a society that does not encourage but discourage people's actualistion of their human potentials? Who should we believe? Where to look for conffirmation of Kant's theses?

A Search for Objective Evidence of History's Progress – Science in the Service of Preserving Order

By positioning freedom as a fundamental principle, while eternal peace as a duty derived from this principle, Kant presented all future interpreters of historical progress with a conundrum. For Kant republic and peace go hand in hand. Internally ordered form of government and peace are equally important. Until both

legally needed elements, peace and freedom, are united, the solution is not found (Maus 2015).

The main problem with the authors engaged in theory of democratic peace is the fact that they have tried to simplify Kant's conundrum in such a way to enable empirical measurement of speed on the road of historical progress. Those contemporary authors who can be credited with rerailing the debate back to the field of philosophy, show where even the best of democratic peace theorists have failed in their ambitious attempt to empirically confirm the thesis that democracies never engage in war between each other (Caranti 2017: 177–181). They always miss something in their attempts to measure the degree of human history's progress. So, instead of objective scientific insights, their works too often feature an arbitrary subordination to the ruling ideology and servicing interests of the most powerful world's state.

It must, however, be admitted that our time poses a very difficult task before those researchers who dare make the claim that history is on the track of finding a solution to Kant's conundrum. This includes references to strengthening of international organisations and declarative advocacy of democracy and human rights all over the world as empirical evidence of Kant's "reveries" turning into historical reality. We live in the time of illusions. Observed from the outside, liberal democracies are truly scattered all over the world. But liberal democracy's institutions are almost everywhere devoid of their fundamental philosophical and historical sense.

Kant's hope for actualisation of republican ideal was sparked under a great influence of the revolutions in France and the USA that marked the historical moment in which he was writing his political texts. American and French constitutions yielded by the two revolutions rely on two essentially different principles of freedom (Arendt 1973). Kant criticised the American Constitution, which reflects the Founding Fathers' distrust of the people and democracy. Convinced that the existing American States provide a better framework for the appearance of the future republic than a large federal state where the will of the sovereign people is structurally limited, the German philosopher sided with the antifederalists (Maus 2015: 77). History would validate the claim showing that the American *demos* has only on rare and exceptional "constitutional moments" – such as the 1861-65 civil war and the 1930s New Deal reforms – managed to leave the Madison's federal prison to reach an autonomous decision on their collective destiny. Moreover, American historical experience speaks of a constant progress in finding original constitutional and political mechanisms that will leave the population in a permanent state of dull passivity in regard to public affairs (Wolin 2008). To spectator not engaged in the game today's world where the American model of democracy and human rights is externally imposed upon other countries by military interventions, can hardly seem to be coming closer to Kant's ideal of putting eternal peace in service of freedom.

Kant pinned more hopes on the promise of freedom given on the ruins of the French Revolution than on the federal Constitution of the USA. Experts on his thought claim that for Kant there was only one republic, and it was the French Republic (Maus 2015: 78). How far has Europe gone in actualising republican principles announced by the Revolution? It is important here to bear in mind that for

Kant, states were the only true guardians of the place where the future republic was to appear. State sovereignty is to be maintained at all costs for its very capacity to one day become people's sovereignty. On the one hand, European states built into their constitutions the principles of the French Revolution, while engaging in wars for territories, supremacy, resources, on the other. For many of today's spectators, the peace guaranteed by the European Union is a proof of Europeans' capacity to learn from their historical mistakes and gradually come closer to actualising Kant's ideal. Their mistakes lie in the fact that they have left out from Kant's eternal peace formula the key element of freedom. The fundamental belief of the French revolutionaries that each authority must have its source and origin in the people is almost completely forgotten in Europe today. Peace in Europe today results from the process of putting democracy under control, stripping it of meaning and neutralising it.

The existential crises that hit the EU after the soaring of the member states' debts in 2009 offered to spectators an opportunity to grasp the idea underpinning the European integration (Kovacevic 2017). As it turned out, integration represents a "hidden technocratic revolution" depleting the states' democracies of more and more decisions without transferring the framework for actualising the republican principle to the European level. Despite peace, human rights and institutions of representative democracy at national and European levels, dull apathy has nested in the souls and minds of European countries' citizenry instead of republican ideals. Lulled in material prosperity and safety the member states' citizens have indifferently accepted cancellation of the possibility of social learning and self-enlightenment. It has been done by politicians, governments, courts, the European Commission, Court of Justice, social scientists, whom Monnet's artful method of integration put into service of actualising revolutionary goals of Europe's political unity and creation of a governing system that will structurally separate the will of the people from the decision-making processes. Thus, over time, the European integration process disarmed all the potential guardians of state sovereignty, bringing crumbling down the only framework for the appearance of the future republic available so far. The peace in Europe where the republican principle has been sacrificed for the sake of prosperity and security today resonates with the silence of a graveyard.

Great Novels in The Service of Measuring the Speed of Progress on the Historical Road to Moral Perfection of Men

Revealing contradictions and misconceptions in the works of today's social scientists seeking empirical evidence of the historical progress are insufficient in solving the dilemma whether the hope of moral perfection of humankind and improvement of social institutions is justified. Are social scientists up to the task of finding the answer to this question at all?

A negative answer to the latter question was suggested by Kant himself. The one who would attempt to offer a scientific proof that social conditions for the development of human potentials have been improved would have to envelop the entirety of historical experience. That would unavoidably lead into to the labyrinth of commeasuring causes and consequences from which it is impossible to find a

way out by applying cognitive processes. The notion of the entirety of experience is scientifically unattainable, showed Kant with conviction. For this reason, we suggest, following in the footsteps of what the German philosopher wrote in The Critique of Judgment about the problem of reconciling the beautiful and the teleological, to seek the measure of historical progress in works of artists. More accurately, in the novels of great writers who wrote what the time and society they lived in whispered into their ears about the relationship between an individual and order.

A writer seeks neither evidence nor unbiased objectivity. Artistic feeling is an "I"-feeling. But it is this very "I" that holds the universal feeling of the world and life. It seems that this artistic feeling always tells the same: society cannot tolerate an individual who defies its banal conventions and sets off on an autonomous search of meaning and dignity. For this, he is doomed to exile, excommunication, misery, tragic death. Such a destiny, which befell Socrates, Hegel found justified. "The spirit of this people in itself, its constitution, its whole life, rested, however, on a moral ground, on religion, and could not exist without this absolutely secure basis. Thus because Socrates makes the truth rest on the judgment of inward consciousness, he enters upon a struggle with the Athenian people as to what is right and true. His accusation was therefore just, and we have to consider this accusation as also the end of his career" (Hegel 1892: 426). By dooming to failure those who question its customs, beliefs, value system, society has been trying to defend its boring order since antiquity.

The artistic truth about the Athenian society that condemned Socrates to his tragic end was conveyed to us by Euripides in his tragedies. This society cannot stand Hippolytus who, staying true to this oath of preserving his chastity, refuses to accept manly and sovereign duties imposed on him by the society (Hippolytos, 428 BC). Having acted in accordance with the feeling of his inner duty, heedless of consequences, he makes a moral decision to keep quiet and preserve honour of Phaedra, his father's wife who was inspired by Aphrodite to fall in love with her stepson in order to exact vengeance on Hippolytus who denies her worship refusing to transition from the world of boyish innocence into the world of adult men. Making his horrible death under the hooves of frightened horses inevitable in the sequence of events, Euripides' artistic genius offers us the truth of his times about the relationship between polis and an individual who decides to defy its rules.

Centuries after Euripides' tragedy had been written, the malign excommunication would be the destiny of Moliere's *Misanthrope*, too (*Le Misanthrope ou l'Atrabilaire amoureux*, 1666). Refusing to accept hypocrisy, banality and superficiality of Parisian society, Alceste engages into a conflict with the monarchist society, refusing to succumb to its "*honnête homme"* norm. The society, which finds Alceste's ridiculing and subversion of banal norms of behaviour and human relationships dangerous for its survival, exiles and dooms to miserable loneliness this rigid moral puritan, who refuses to "silently adopt the spirit of the time". Shakespeare's Hamlet (Hamlet, 1602) will share Hippolytus and Alceste's destiny. The bitter rebellion of the Danish prince against the society founded on hypocrisy, lies and betrayal will end once Shakespeare's hero, unlike Kant, concludes that the world is impossible to change and put back on the right track. Hamlet will disdainfully reject the maturity heralded by such a conclusion and, thus resigned to his fate, go to meet his death.

One philosophical interpretation of Dostoyevsky suggests that he foretold a historical quake, a creation of a new man, society and world. "He belongs to the new world. Only formal analysis of his works can show whether he is already the Homer or the Dante of that world or whether he merely supplies the songs which, together with the songs of other forerunners, later artists will one day weave into a great unity: whether he is merely a beginning or already a completion" (Lukacz 1971: 152-153). If the Russian literary genius is observed from this point of view, similarity between the protagonist of his The Idiot (Idiot, 1869), Prince Myshkin, and Socrates comes as no surprise. Socrates' enquiry of the beautiful, the virtuous and the good, coupled with encouraging an unbound curiosity in the youth, jeopardised the foundations of his Athens' order. Myshkin's truthfulness of an innocent child will wrack chaos with the souls and minds of the members of the society whose artistic truth is delivered by Dostoyevsky. Myshkin's presence alone, his child-like questions and goodness help people to recognise, if only for a moment, the divine, the light, the spark lying hidden under the layers of hypocritical and rotten society, the very spark that Tolstoy's Nekhlyudov finds on the train to Siberia. Myshkin's boyish spirit, however, at the same time stirs the selfish, banal, passion-blinded and the vengeful that the existing social customs and rules inspire in man. Perhaps Dostoyevsky was truly a man of the coming age, but his answer to the question of the relationship between an individual and order was the same as Euripides', Moliere's and Shakespeare's. Defending the order, the narrow-minded society eventually dooms Myshkin to lunacy and sends him back to the same Swiss sanatorium for mental patients from which a train brought him to Russia at the beginning of the novel.

German experience holds special importance for testing Kant's vision of the moral progress of humankind that learns from horrible war experiences.

So, in his novel *The Tin Drum* (*Die Blechtrommel*, 1959), Gunter Grass writes about a boy who refuses to grow up and starts to speak in order to avoid engaging in the world of adults and sinking into brutality of World War II.

Heinrich Boll has left to the world a book about a lonely clown who roams the trains and towns of post-war Germany (*Ansichten eines Clowns*, 1963). This twenty-seven-year-old boy from Bonn is haunted by a clear memory of his mother's voice passionately telling her children at the dinner table: "everyone must do his bit to drive the Jewish Yankees from our sacred German soil!" Haunted by this voice, he tries to escape the society in which this same woman has come to preside over the Executive Committee of the Societies for the Reconciliation of Racial Differences and regularly goes to the Anne Frank House in Amsterdam. Boll's clown remains a pure and innocent boy who refuses to engage in the hypocritical world of adults showing its face from the war ruins of Germany. Perhaps this hypocrisy is necessary for modern society to function at all, for people to enjoy security and material prosperity. Perhaps hypocrisy is a necessary element of the passive neutralisation of social conflicts written about by Carl Schmitt. In such a society, there is no place for Boll's hero. He remains a lost clown who performs his sad act about coming and going at railway stations.

In his books, Peter Handke roams alone the gloomy global society of today searching in people's eyes for a spark that testifies to the existence of human

potential for the beautiful and the sublime that Kant and Mirandola pinned their hopes on. In his novel *The Moravian Night* (*Die morawische Nacht*, 2008), he finds it by mere chance in the eyes of a girl reading a book on a bus in Denmark or a boy who, holding a ball, stands in the midst of a destroyed cemetery in Kosovo.

***

The foregoing is not to lead one to conclude that Kant was wrong. It is important to understand that Kant's freedom, republic and citizen are ideals the world constantly comes closer to, yet never reaches. However, these are not ideals detached from reality or above it. Kant recognised the grain of purposeful historical progress in the reality he knew. Therefore, the question is not whether or not Kant was right. The question is how to measure how far society has come in cultivating the grain of humanity captured by Kant's philosophical imagination in the world of his time. In contrast with today's social scientists, writers of great novels say from their respective epochs: not an inch.

Social Scientist as a "Political Moralist"

Kant's philosophy has always attracted artists whose vivid imagination found in it all kinds of inspirations and giving it a plethora of interpretations. Indeed, the fate of philosophical ideas is determined by mirrors these ideas will reflect in, regardless of the protests of philosophical schools or experts for certain thinkers. One of the most wonderful mirrors to reflect Kant's thought was Goethe's poetic soul (Cassirer 1970). The views of the philosophical and poetic geniuses differed in many ways. While for Kant the beautiful and the good, genius and scientist, art and nature must stay apart, Goethe does not accept a sharp division between science and art (Cassirer 1970: 85). Therefore, Kant would perhaps frown upon the humble intellectual experiment conducted in this article, which looks for a measure of historical progress in the works of the great novelists. Goethe might approve of it. Or he might not. We cannot answer this question with any degree of certainty, however disdainful Kant experts might find it. This article, indeed, was not aimed at offering yet another in the long series of "correct" interpretations of Kant's political philosophy. The humble aim of the paper was to create yet another in the endless series of small mirrors for reflecting the great German philosopher's thought, using history of ideas, political theory and history of literature.

Kant's gift that Goethe wholeheartedly embraced is setting limits to that which can be grasped by mind. Goethe was grateful to Kant for calling scientists on humility, warning them not to probe into that which cannot be known. The greatest German poet saw Kant's philosophy as something liberating, something that by setting the clear limits on the experiential knowledge clears the field where artistic imagination will freely seek the truth. Everything that science is unable to prove is left to artistic genius (Cassirer 1970: 78).

Contemporary social scientists' attempts to prove historical progress is characterised by the very lack of humbleness. Today's interpreters of Kant's thought have a task to show them where they are wrong, to remind them of the limits of what can be known. Refereeing to the great novelists' works in this article about

Kant was actually aimed to remind scientists of restraint and self-control demanded from them by the citizen of Konigsberg.

To disagree with the authors who see tendencies of historical progress in today's world does not mean to reject their understanding of "moral politician" (Caranti 2017: 235–255). Even if coming closer to the idea of republic and establishing the state of eternal peace is accepted, moral politician can contribute to facilitating the process leading towards attainment of this goal. His task is to carefully manipulate facts of the empirical world into compliance with the universal law of righteousness. But, to achieve this, it is impossible to follow a scheme, applying rules given beforehand to reality. Politics cannot be reduced to science. A politician's skill cannot be learned. The world where politics is possible at all cannot be a closed and predictable world of given facts combined and recombined following the rules derived from scientific observation. "She must be endowed by nature with a talent to know how and when scientific and moral norms of various kinds are to be applied" (Caranti 2017: 255). It is for this reason that talent, which only few are gifted with, is needed for both politics and art. What is needed is imagination, creativity. Therefore, it comes as no surprise that it is "more likely to meet a good man than a good politician" (Caranti 2017: 255), the interpreters of Kant's political philosophy remind us.

However, they claim that besides talent, a politician needs knowledge, the kind of knowledge that the best social science today can give him. "She must have acquired solid and wide knowledge of the empirical laws relevant to her decisions, that is potentially all those of the social sciences plus history" (Caranti 2017: 255). The question we pose at the end of this article is: Can a social science that sees the legal and political tendencies of today's world as proofs of historical progress bring a politician to the right path?

Amid the European public debt crisis of the early 2010s, two important thinkers wrote books that attracted great attention. One comes from the world of social sciences, while the other is a novelist. They are Jurgen Habermas and Michel Houllebecq. In Habermas' book, the systemic crisis that has led to a suspension of democracy and collapse of the previous way of life in the countries of Southern Europe is just a necessary step on the path towards the development of a republican Constitution of Europe, which is merely an introduction into actualisation of Kant's ideal of a world civil society (Habermas 2012). On the other hand, Houllebecq in his *Submission* (*Soumission*, 2015) writes about the lost French literature professor left without references points that used to be provided to European man by profession, religion and family. Terrified by such a freedom of a child left alone in the cradle, professor eventually submits to the new worldview offered by Islam for the sake of comfort.

So, the question is who Kant's "moral politician" is to believe? Habermas, who provides the gloomy European reality with a philosophical justification uttering the words of comfort: everything is going as planned, exactly as it is supposed to, reforms are supposed to be modest and limited in their reach? To Houllebecq, who wakes him up from his stupor with a scream and warns: historical regression is at work, even the little autonomy and dignity whose respect a European had won since the French Revolution in his national state has now vanished; unless you

want Europe to slip into dark despotism, something major must be done quickly and society radically changed?

The answer can only be given by a true "moral politician". A political theorist must satisfy himself with an insight that the conclusions of the most influential European philosopher and artistic truth have found themselves on the warring sides in today's European society. The theorist can claim with certainty that Habermas will not share the destiny of Socrates, who brought the constitutional foundations of his polis under question with his unbridled and truth-seeking curiosity. The theoretician can also conclude that Habermas has ignored Kant's requirement for scientific humbleness in his attempt to find in the development of a supranational governing system of the EU scientific indicators of historical progress. The French Revolution kindled the spark of the sublime in Kant as an unengaged spectator and encouraged him to dream of an eternal peace. The European integration, as a hidden revolution, has derailed European nations from the path of gradual self-enlightenment and learning through historical experience. It is hard to believe that this wearisome process led by technocratic logic can inspire an unengaged spectator to dream of making the idea of freedom true. Sympathies of the unengaged spectators might have been awakened for a short while by the rebellion of the citizens from southern states in the height of the public debt crisis (Douzinas 2013). If such sympathies really existed, they must have turned into disappointment quickly when the Greeks, French, Spanish, Italians and Portuguese were discouraged by fears of losing material security from rebellious demands for freedom. Yet, even though the European Union does not encourage dreaming of freedom, it has so far managed to secure peace. That explains why philosophers who see before their eyes the horrors of war suffering from the previous European epochs place their work into the service of preserving the present supranational order. The legitimate fear of new conflicts on European soil overwhelms their faith that autonomous learning and self-enlightenment of Europeans is possible, leaving them with nothing else to do but to take upon themselves the role of "political moralists".

Refernces:

Arendt, Hannah (1973), *On Revolution.* London: Penguin books.
Arendt Hannah (1992), *Lectures on Kant's Political Philosophy*, Chicago: The University of Chicago Press.
Burke, Edmund (1910), *Reflections on the French Revolution.* London: J.M. Dent & Sons Ltd.
Caranti, Luigi (2017), *Kant's Political Legacy. Human Rights, Peace, Progress*. Cardiff: University of Wales Press.
Cassirer, Ernst (1970), *Rousseau, Kant Goethe*. Two Essays. Princeton: Princeton University Press.
Cassirer, Ernst (1981), *Kant's Life and Thought.* New Haven: Yale University Press.
Ciaran, Cronin (2003),"Kant's Politics of Enlightement", *Journal of the History of Philosophy*, 41: 51–80.
De Maistre, Joseph (2006), *Considérations sur la France : Suivi de Essai sur le principe générateur des constitutions politiques*. Bruxelles: Editions Complexe.
Douzinas, Costas (2013), *Philosophy and Resistance in the Crisis.* Cambridge: Polity.
Habermas, Jurgen (2012), *The Crisis of the European Union: A Response*. Cambridge: Polity Press.

Hegel, Georg Friedrich, (1892), *Lectures on the History of Philosophy*, Volume One. London: Routledge and Kegan Paul.
Kant, Immanuel (1998), *Critique of Pure Reason.* Cambridge: Cambridge University Press.
Kant, Immanuel (2000), *Critique of the Power of Judgement.* Cambridge: Cambridge University Press.
Kant, Immanuel (2006), *Toward Perpetual Peace* and Other Writings on Politics, Peace, and History. New Haven and London: Yale University Press.
Kant, Immanuel (2011), *Observations on the Feeling of the Beautiful and Sublime and Other Writings*. Cambridge: Cambridge University Press.
Kolakowski, Leszek (1978). *Main Currents of Marxism. Its Rise Growth and Dissolution*. Volume one. Oxford: Clarendon Press.
Koselleck Reinhart (1988), *Critique and Crisis. Enlightenment and the Pathogenesis of Modern Society*. Cambridge: The MIT Press.
Kovačević, Bojan (2017), *Europe's Hidden Federalism: Federal Experiences of European Integration*. London and New York: Routledge.
Lukacs, Georg (1971), *The Theory of the Novel*. Cambridge Massachusetts: MIT Press.
Maus, Ingeborg (2015), *Menschenrechte, Demokratie und Frieden, Perspektiven globaler Organisation*. Frankfurt am Main: Suhrkamp.
Mirandola, Pico (1998), *On the Dignity of Man*. Indianapolis/ Cambridge: Hackett Publishing Company.
Renan, Ernest (1996), *Qu'est-ce qu'une nation et outres* écrits *politiques*. Paris: Imprimerie nationale.
Schmitt, Carl (2007), *The Concept of the Political*. Chicago and London: The University of Chicago Press.
Schmitt, Carl (2005), *Political Theology. Four Chapters on the Concept of Sovereignty*. Chicago and London: The University of Chicago Press.
Wolin, S. Sheldon, (2008), *Democracy Incorporated. Managed Democracy and the Specter of Inverted Totalitarianism.* Princeton and Oxford: Princeton University Press.

Bojan Kovačević

Subjektivna univerzalnost pisaca velikih romana kao umetnička mera napredovanja istorije ka ostvarenju Kantove vizije slobode

Apstrakt

Osnovna teza članka jeste da bi za razumevanje smisla koji Kantova politička filozofija zadobija u konkretnom društveno-istorijskom kontekstu određene zajednice trebalo iskoristiti pomoć genijalnih umetnika u čijem se subjektivnom „ja" smestilo opšte osećanje sveta i života U tom smislu, pisci velikih romana mogu nam pomoći na dva načina. Prvo, njihova dela sažimaju za našu uobrazilju umetničku istinu o čovekovom kapacitetu za ljudskost, onome što za Kanta predstavlja naučno nedokazivu „činjenicu razuma". Drugo, u delima velikih pisaca možemo posmatrati sudbinu pojedinaca koji u određenom društvu odluče da deluju u skladu sa moralnim zakonom, ostvarivši tako mogućnost koja leži u svakome od nas i čini nas vrednim poštovanja. U odsustvu objektivnih naučnih merila, umetničko opšte osećanje sveta prenosi nam se kroz priču o čoveku koji u konkretnoj zajednici u određenom vremenu odluči da ostvari svoju autonomiju, potraži božansko u sebi. Odsustvo skromnosti ono je što odlikuje pokušaje današnjih društvenih naučnika da dokažu istorijski progres. Pozivanje na dela pisaca velikih romana u ovom članku o Kantu ima za cilj da opomene naučnike na uzdržanost i samograničavanje koje je od njih zahtevao građanin Kenigsberga.

Ključne reči: Kant, roman, mir, Tolstoj, genije, nauka, progress, istorija

UDK: 342.7:172.4 Kant I.
https://doi.org/10.2298/FID1804586F
Original Scientific Article
Received: 23.06.2018. Accepted: 05.08.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Luigi Filieri

AUTONOMY, DIGNITY AND HISTORY IN CARANTI'S *KANT'S POLITICAL LEGACY*

ABSTRACT
In this paper I discuss some relevant theses of Caranti's *Kant's Political Legacy*, whose aim is to provide a consistent account of how we could develop Kant's political thought and see to what extent Kant's insights can help us to critically understand the 21st century's political world. First, I will focus on autonomy as the ground of dignity and discuss Caranti's arguments against the exclusiveness of the Categorical Imperative as the sole principle of true moral agency. Second, I will take into account Caranti's views on history and consider whether human rational nature can be regarded as containing teleological – though non-biological – elements, thereby questioning Caranti's Separability Thesis.

KEYWORDS
Autonomy, dignity, moral agency, teleology, history, human nature.

Prologue

Our world, in our time, is marked by the always increasing urgency to find good solutions to very relevant political problems. In 2018, mankind has at its disposal enough technology to plan a voyage to Mars, but still no satisfactory remedies for poverty, terrorist attacks and plenty of social and juridical issues which it is impossible to mention exhaustively: from the thousands of deaths in the Mediterranean sea to rape cases in India; from labor exploitation in Asia, Africa and South America to the recent challenges to the free market, without forgetting Brexit, the Middle-East crisis and so on.

Everybody will agree that the answer to these political problems has to be political as well, either in terms of single states or by involving supranational institutions. In turn, this leads politicians to search for strategies and intervention-hypotheses. But again, this raises further questions. Plenty of political scientists, sociologists and political philosophers thus investigate the sense and meaning of individual laws or jurisdictional controversies, as well as of socio-economical themes. Consider, as an example, the range of the discussion on human rights, transnational justice and, more recently, environment protection. In order to discuss these issues, one should have – in the first place – an idea of what would be the essential feature of every single political move (extremely broadly conceived), namely what makes these issues *political*. This sort of theoretical regression would reach its final step by asking: *what is* politics, after all? How should we understand the fact that

Luigi Filieri: PhD candidate in Philosophy at the University of Pisa and Research Fellow at the Italian Institute for Philosophical Studies (IISF-Neaples); l.filieri7@gmail.com.

even the most tribal communities still have some kind of political organization? If one wants to pursue this line of thought beyond this step, he or she will reach the point where the issue is simply: *what is* man? Or, who are we? To answer this question means to understand how politics is essential to philosophical investigations. After all, Plato already taught us that the city and the soul are tied to each other (see Fussi 2012).

We do not want here to move beyond the specific question of politics. Instead, we think it is relevant to understand why we cannot avoid going back to the political philosopher Kant when searching for answers concerning the contemporary political world.

For an example, as Paul Guyer stated,

> there has been much more extensive discussion of Kant's views about perpetual peace, world government, and cosmopolitanism since the 1990s, especially 1995 (200 years after the publication of *Perpetual Peace*), and up to a brand-new book called *Kant's Embedded Cosmopolitanism* by Georg Cavallar. Then its topics of world peace, world government, cosmopolitanism, international justice, etc. were very extensively discussed, much more than they were discussed before. Now, is that just because of the contingent fact that it happened to be the two-hundredth anniversary of *Perpetual Peace* 20 years ago? I do not think so. It is because of globalization, it is because of the issues about international justice, global justice and so on, that are generally more pressing issues now than they were earlier; and that impacts what seems important in scholarly studies. Similarly, there has been an explosion of non-Kantian literature on global justice in the Anglo-American philosophy in the last twenty years. So, that has become a major topic in Kant scholarship as well. I can imagine that, as democratic forms of government come under increasing pressure, even in the long-established democracies, as they have been in recent years from issues of immigration, from inequality, from resentment, all the ways in which governments and long established democracies in Europe and even in the United States are under pressure, that then people may turn back to Kant's political philosophy, which sometimes seems to people to be old fashioned and boring, because it is a profound defense of the primacy of freedom and of the need for limits to paternalistic government or worse. Kant's views about what he calls republican government, the equal standing of the individual before the law, the separation of powers etc. etc., become urgent and important, just as the discussion of the international issues became in the last twenty years (Guyer 2017: 22–23).

Two important points emerge: 1) Kant has left a huge legacy of political concepts and reflections; 2) this legacy is no dead letter. In a few words, these are the basic premises of Caranti's *Kant's Political Legacy*, which belongs to a very valuable series of researches into Kant's political philosophy[1]. A third premise would be our initial consideration, namely the very fact that we feel the increasing urgency to find good solutions to our political problems. Our political world can find in Kant's thought a very valuable instrument to understand what is happening, why it is happening, and how to manage a change towards what should happen. Obviously, one could ask: *why* Kant? The philosopher of Königsberg was not the sole thinker who dedicated his time to political themes. For a good answer, Caranti's reader can refer to

1 Let me just quote – besides the already mentioned Cavallar 2015 – Kleingeld 2011 and Krasnoff, Sánchez Madrid, Satne 2018.

the first lines of page 2 of *Kant's Political Legacy*. Kant cannot offer any ready-for-use political solutions, but only a direction for our efforts. Let me just clarify that this direction is of the greatest importance. As I said before, our political world is extremely fragmented. Its problems thus require a systematic approach.

Though it could seem superfluous to say that a globalized environment needs worldwide perspectives, it is precisely with Kant's political philosophy that we see how the direction we are seeking has to be of the widest range. For a simple fact, namely because the question *what is man?* is a Kantian question. And again, the extent of Kant's thought shares with Plato what we could call the unavoidability of politics. It is impossible to avoid developing a political reflection when philosophizing on man. As I said, our political world is fragmented; not only because we are facing very heterogeneous problems, but also because the very status of each political issue is not clearly determined. Human rights, peace, progress have a single meaning only within the pages of a dictionary. The concrete forms these ideas assume do not display any uniformity. Rather, these forms represent the reason why it is necessary to include political reflection in every philosophical inquiry of man. In these terms, Kant's critical philosophy is a valuable resource.

Now, it is worth noticing that Caranti's book is not simply a book on Kant's political philosophy. The title of Caranti's book has to be taken very seriously. Caranti is not simply interested in discussing what Kant has left us. Instead, this book represents a clear example of how a great philosophical legacy can contribute not only to the comprehension of the contemporary world, but also to the definition of new thought-paradigms and, why not, even intervention strategies at a more general level (guidelines, analysis, etc.). Caranti's book tells us what to do with Kant's thought, given that this thought, as I said, provides us with a direction for our efforts. In this sense, one can easily understand why Caranti has chosen to focus on human rights, peace and progress. The political challenges of our world precisely require us to say what has to be understood as a human right, and for which reasons. In addition, peace and progress cannot just represent the essential features of a benevolent person. Again, a political philosopher has at least the duty to understand what peace and progress mean, why they should be pursued and, eventually, how they could be reached. Obviously, this is beyond a critical discussion of Kant's texts. Our world is not Kant's. Nonetheless, Kant seems to have at least tried to discuss themes whose urgency in his time was surely lesser than in ours. This is what we should call the greatness of a philosopher. The fact that Caranti has not only recognized it, but also tried to make valuable political perspectives out of it, represents one of the major merits of *Kant's Political Legacy*. Though much can be said of *Kant's Political Legacy*, in what follows I will focus on two specific topics in order to both discuss Caranti's interpretation of Kant and make some critical reflections on Caranti's theoretical proposals.

1. Autonomy and Dignity

My first concern is that of autonomy as the ground of dignity. In the third chapter of his book, after having discussed Kant's ground for the normativity of human rights, Caranti develops an alternative view:

> We start from the premise that we are autonomous beings, a view that is not only at the centre of Kant's philosophy but also – as we will show – of (a) common sense and (b) all major cultural traditions and revealed religions, at least when reasonably interpreted. Autonomy is not to be understood merely as the ability to choose one's path in life, or as the ability to be rational in the sense of purposive agents. With Kant, we refer to a capacity distinct from and 'higher' than practical freedom. We have in mind the ability to act under self-imposed moral constraints. This capacity – it will be argued – shows us as worthy creatures, and reveals the deepest and most stable layer of human value. Reflecting on our autonomy, we turn out to be beings with this fascinating feature: being able to silence all natural impulses, even the strongest instincts of survival, and act from our conception of duty (Caranti 2017: 57).

The core of the proposed interpretation lies in the following arguments. As a first point, Caranti wants to detach the concept of autonomy from the exclusiveness of the Categorical Imperative. This hypothesis can be considered as Kantian in form, though the content radically diverges from Kant's account. According to Caranti, autonomous agency does not need to exclusively follow the moral law, since following other available rules (the Golden Rule for example) would still mean to act autonomously. In a few words, differently from Kant, true autonomy does not require a morality grounded on the Categorical Imperative. Recalling Pico della Mirandola's *Oration on the Dignity of Man*, Caranti focuses on the concept of self-determination. According to Pico, human dignity lies in the capacity to became whatever one decides to be. However, Caranti's account moves one step further, for he wants to distinguish between autonomy intended as mere self-determination and autonomy intended as the capacity for moral agency[2]. If the question is *why do human beings possess dignity?*, Caranti's answer would be: *because they are autonomous, namely free to act under self-imposed moral constraints.* As I said, these constraints are plural, and the moral law loses its exclusivity.

We see again why Caranti's move is Kantian only according to the form. Caranti indeed holds that the moral constraints have to be categorical in kind, though this does not require one of the formulas of the Categorical Imperative. The difference from Kant on this point could be summed up in this way: while Kant stated that an agent is truly autonomous when consciously self-constrained according to the moral law, Caranti instead claims that, given the capacity to act under self-imposed constraints, an agent is autonomous by the very fact of acting in this way, no matter what the constraints properly consist in, nor how successful the action is. Thus, duty and Categorical Imperative no longer go together. Another point Caranti insists upon is the independence from contingent sensuous motives. Acting according to duties involves a rejection of whatever could conflict with the same duties, given that such obstacles, however, influence the will. Since they are attractive and powerful, resisting them thus requires a self-constriction.

The reasons why Caranti differentiates his account from Kant's are mainly two. On the one hand, an account of autonomy which does not depend on the Categorical

2 See Caranti 2017: 61: "The distinction between the common-sense notion of autonomy (self-determination) and the one at work here (capacity for moral agency) is important in understanding the way in which our approach links the possession of a faculty to the intrinsic worth that entitles humans to the protections of human rights. In fact, why should the sheer possession of a capacity ground any worth?"

Imperative is transcultural. Several different cultural traditions, and even religions, share something similar to the Golden Rule or, at least, a principle for autonomous agency which always requires self-constriction and resistance to sensuous motives. Through this, Caranti thinks it is possible to outline a Kantian perspective which could be accepted outside the orthodoxy of Kantian interpreters. On the other hand, by renouncing the Categorical Imperative, Caranti's account of autonomy both disregards any metaphysical assumptions and demands to be compatible with common sense. To sum up, once we sever autonomy from the exclusiveness of the Categorical Imperative, we obtain a wider, transcultural, but still universally valid account of autonomy whose recognition grounds the dignity of human beings.

The thesis is agreeable and thorough. Caranti's arguments demand that we 1) develop Kant's political legacy without emulating it; 2) provide a transcultural account of autonomy; and 3) allow the latter to ground the concept of human dignity. This said, given that a book is interesting when it raises questions, let me now ask mine. The first: is this account of autonomy strong enough to satisfy Caranti's demands? For sure, Caranti's notion of autonomy seems more promising than Kant's, but one could wonder whether Kant's notion of the Categorical Imperative as the *ratio cognoscendi* of freedom has still something to say. Caranti carefully distinguishes between practical freedom and autonomy (pp. 24–33) in this way: while the former concerns the capacity of setting ends for oneself without being influenced by sensuous inclinations, with the latter an agent sets his or her own end independently of any empirical motive. In the first case the inclinations are necessary, but not exhaustive, conditions for the will's choice; in the second, they play no role. Though this distinction is clear, it could be reframed by focusing on the ground of an action. For example, I am practically free to recycle in order to protect the environment or not to be fined. What grounds my choice here is an end that is external to the object of the action.

Very far from here is the choice of jumping into the water, at my own risk, to save the life of someone endangered. The object of my action and the ground of the choice are in this case one and the same. It must be noticed here that the causality of freedom does not deal with cognition, but with actuality. Indeed, for Kant,

> [...] the moral law is not concerned with cognition of the constitution of objects that may be given to reason from elsewhere but rather with a cognition insofar as it can itself become the ground of the existence of objects and insofar as reason, by this cognition, has causality in a rational being, that is, pure reason, which can be regarded as a faculty immediately determining the will (*KpV*, 5: 46 (177)).[3]

According to this picture, a truly autonomous agent ought to take the grounds of existence of his or her moral action as the sole end of his or her will. Obviously, this does not mean to know with certainty that the action is truly from duty, for according to Kant there can be no experience of freedom[4]. However, if the grounds

3 *KpV*, 5: 46 (177). See Ferrarin 2015: 128–136.

4 As already mentioned, consider someone who jumps into the water to rescue someone endangered. He or she recognizes the duty to help a castaway, for his or her action comes from the self-imposed constraint to risk his or her own life for the action to be successful. The command of the moral law makes the subjective maxim of our rescuer a universal law. Therefore, the moral action of the rescuer is one of those objects whose ground of existence

of existence of the object of the action were not the end of my action, this would certainly compromise the morality of the same action, for my agency would be driven by heteronomous motives and towards external ends.

This brings us back to the question whether Caranti's account of autonomy is enough. For sure, Kant's account seems to satisfy at least demands 2)[5] and 3), but then we should ask Caranti why he thinks he can do without the Categorical Imperative when Kant's notion of autonomy already did the job (precisely resting on the Categorical Imperative). For sure, a first answer could be that demand 1)[6] would otherwise remain unsatisfied. On this point I think Caranti's approach is extremely valuable, for it tries to make the best possible use of a huge legacy, moving many steps beyond mere interpretative work. For what concerns the other two issues, I think it could be useful to specify that the Categorical Imperative is not properly a law, but rather the mere form of a law (see Longuenesse 2005: 246–264). I would not say that, according to Kant, one must subjugate his or her maxim to one of the formulations (see Caranti 2017: 63–64) of the Categorical Imperative. Rather, I think that the law according to which I determine my maxim has to be the (only) one through which my maxim itself becomes a law. For this to happen, the moral law legislates the mere form of an action, whatever its content may be.

If I have understood Caranti correctly, I think he holds that autonomous agency does still need a principle, though he refuses Kant's claim that only the Categorical Imperative is entitled to be the principle at stake. What I suggest here is to consider the Categorical Imperative as the mere form of a law, and not as a principle which commands what to do. This would be, in my humble opinion, the reason why Kant thought that freedom in the proper sense (namely autonomy, not mere practical freedom intended as rational agency) is achievable only through the Categorical Imperative. Compared to the Golden Rule (which, as Caranti correctly claims, belongs to several cultural and religious traditions), the Categorical Imperative displays two main distinctive features. First, it does not say what to do, but only how to act. The moral law is no precept. Rather, since it merely provides us with the form of a law, it makes us responsible of what ought to be done[7]. Second, the key word of the moral law is *act*, not *will*. The action at stake, to be a mark of true autonomy, has to be legislative. In turn, to this end, the action must conform to nothing but the Categorical Imperative.

It is true that the Golden Rule can boast a wider recognition across different cultures. However, I think that an account of autonomy built on this principle,

---

is the moral law. Thus, while in the speculative synthesis the matter is always given in intuition, in the practical case we regard the moral law as a kind of performative cognition, for acting according to the categorical imperative makes objects real in the first place.

5 I take here *universality* as wider than *transculturality*. The universal validity of the moral law does not follow from its content, but from the formal element of the imperative. Thus, the moral law radically diverges from other formulas, for I think the imperative cannot be regarded as a formula. Put succinctly, the moral law does not provide us with universally valid maxims or precepts, but only with a formal principle according to which – exclusively – our subjective maxims can be regarded as universal laws. For one of the most clear and precise readings of Kant's account of the moral law see Kleingeld 2017.

6 That is to develop Kant's political legacy without emulating it.

7 On Kant on responsibility see Blöser 2015, Willaschek 2003.

though legitimate, is weaker than the one Kant tried to provide. Put succinctly, if it is correct to claim that human dignity rests on autonomy, I think that true autonomy requires more than acting according to self-imposed duties. As Kant claimed, autonomy requires us to act as if our subjective maxims could be regarded as universal laws. Autonomy does not merely mean to give a law to oneself, for it also requires me to give laws by myself. This means that autonomous agency does not merely require us to act under self-imposed constraints, that is according to a law, for it also makes of my subjective maxims the type (see Ferrarin 2004: 148–156) of a universal law. In a few words, in this case the will is not merely self-legislated but, more properly, it is legislating by itself (see Kleingeld 2018: 71).

Here is where the Golden Rule account misses a crucial point the Categorical Imperative can provide. However, this is not meant to disregard the distinction between developing and emulating a legacy. On the contrary, my point concerns what to do *now* with the Categorical Imperative, namely how should we understand the relationships between the supreme principle of morality and the ground of human dignity.

When I ask Caranti whether autonomy is enough, I do not pose a rhetorical question. On the contrary, I think the question is really open and in need of satisfactory answers. Credit is due to Caranti for having contended with one of the most significant elements of Kant's political legacy.

2. Teleology and History

The second topic I will discuss here concerns the relationship between teleology, nature and history. In chapters 7-9, Caranti develops his Kantian view of progress. As he already stated in the case of autonomy, this interpretive perspective is also Kantian in terms of method and form, though it does not entail any mere acritical description of what Kant said. Thus, again, it is Kantian, but not Kant's. Caranti takes into account three texts: the *Idea for a Universal History with a Cosmopolitan Purpose* (1784), *The Contest of the Faculties* (1798) and *To Perpetual Peace* (1795). By focusing on *Idea*, Caranti introduces one of the main theses of Part III, namely the *Separability Thesis*:

> the theory of 'natural dispositions' and the theory that spells out the consequences of social unsociability are separable and independent. One can believe in social unsociability (proposition 4) as well as accept the account that spells out the predictable institutional repercussions of such a mechanism (propositions 5–9) without endorsing the 'natural dispositions talk' in which they are embedded. Not accidentally, Kant himself will introduce the concept of unsocial sociability in the First Supplement of *To Perpetual Peace* after an account of nature completely different from the one introduced by the first three propositions (Caranti 2017: 214).

Let us clarify Caranti's aim. The first three propositions of *Idea* hold a problematic view that is incompatible with Darwinian evolutionary theory. According to Kant, the natural dispositions of all creatures are destined to develop both fully and completely. Moreover, human's natural disposition to reason, best develops only within the space of a society. Caranti thinks that Kant's natural dispositions correspond to what contemporary science calls *genetic materials*, whose main feature is

to be subjected to mutation and natural selection. In other words, there is no stable natural disposition to count on. Finally, Caranti also notes that Darwinian biology overcomes Kant's idea that mere mechanical laws (without any teleological orientation) cannot account for the production of any natural feature. The contemporary scientific paradigm indeed excludes any intentionality in nature, resting instead exclusively on mechanical explanations (which are in turn consistent with mutation and natural selection). In order to solve the problem, Caranti then focuses on the remaining six propositions. Though Kant here connects (fourth proposition) the mechanism of social unsociability to natural dispositions – for nature employs antagonism as a means to fully develop human nature – nonetheless Caranti claims that the mechanism of social unsociability occurs independently of any natural disposition. Therefore, one can abandon the teleology of natural dispositions and, at the same time, retain the mechanism of social unsociability.

Now, this thesis is part of a wider project, that is grounding a progressive view of history. Once this step has been completed, Caranti addresses three further explanatory issues. First, the teleological development of history has to be understood without any natural purpose; second, the validation of such teleology rests exclusively on a natural mechanism (there is no need to unify an alleged lawless aggregate of events); third, our teleology does not play a mere regulative role, since it has to be shown that progress (though non-linear) will occur more likely than regress or stagnation. To be fair, Caranti's project is even wider. In chapter 8 he discusses the Guarantee Thesis, according to which nature has a plan towards the achievement of perpetual peace. The link between Caranti's Separability Thesis and Kant's Guarantee Thesis lies in the fact that neither play a mere regulative role. In Caranti's words,

> there is nothing a priori wrong or dogmatic in focusing on certain fairly stable human inclinations (such as the pursuit of happiness or self-love), as well as on certain very general and uncontroversial empirical facts about the world, to infer from them a thesis about the likelihood of a certain evolution of human affairs (Caranti 2017: 233–234).

Caranti would here borrow the scientific paradigm of *unintended consequences*: our free choices, through consequences we are not always able to master, outline the shape of our political world. Thus, Caranti thinks that through these steps it is possible to argue for an actual development of human political agency towards a condition of perpetual peace, though the latter does not represent a pre-fixed end of nature.

Two things are very remarkable. First, by insisting on the fact that neither the Separability Thesis nor the Guarantee Thesis play a mere regulative role, Caranti proves how seriously he takes the necessity to develop Kant's political thought. Caranti's moves can be easily considered as attempts towards new political paradigms which, in turn, would improve our understanding of the political environment surrounding us. Moreover, though he does not explicitly say that the two theses are *constitutive*, I think they at least outline a coherent political view through which we are not merely able to explain the past but, most importantly, through which we can intervene in our political present. After all, in chapter 9, Caranti masterfully synthesizes the steps of his philosophical research by outlining the figure of

a 21st century moral politician, which is first of all required to understand how his or her own political reality is developing in history and, consequently, to find new political concepts through which another reality can be made actual. Therefore, the reader is strongly recommended to read pp. 246–255.

This said, there is still a question we can ask Caranti. What does he precisely mean when he speaks, more or less explicitly, of humans' natural dispositions? If these dispositions are one and the same with the *genetic materials* Caranti refers to, there is obviously no room for an objection to the Darwinian view. However, one could wonder whether Kant thinks that reason properly belongs to human nature or if instead it represents what is beyond nature. The dichotomy is of the greatest importance. On the one hand, as Caranti correctly puts it, assuming a hidden plan of nature in order to ground a progressive view of history is not a legitimate move. On the other, in his *Idea* (first proposition) Kant also says that

> Reason in a creature is a faculty of extending the rules and aims of the use of all its powers far beyond natural instinct, and it knows no boundaries to its projects. But reason itself does not operate instinctively, but rather needs attempts, practice and instruction in order gradually to progress from one stage of insight to another (*IaG*, 8: 18–19 (109)).

Had Kant meant that reason is a natural disposition among others, we would then face the problem of having a natural disposition which disregards and even conflicts with other natural dispositions. The other available option is to think of reason as something (in nature) which – at the same time – also transcends nature. In both cases, the problem is that of making nature and reason consistent with each other. I think it could be misleading not to insist on the fact that the natural disposition called reason is only partially natural. This is not meant to be a criticism of Caranti's objections to the first three propositions of *Idea* but, more modestly, a further deepening of Caranti's use of the distinction between *homo phaenomenon* and *homo noumenon* (pp. 27–28). The question is: *how many natures does a human being possess?* Insofar as it belongs to both the kingdom of nature and to that of reason, a human being seems to have two natures, not merely one.

Now, on the one hand, these two natures can be regarded as separated. In this regard, human beings are either part of natural causal chains (which also include instinct) or rational agents which set ends for themselves. This is only part of the story, however. I would suggest that the duplicity mentioned above does not entail any strong dualism. The very fact that human beings are subjected to instinct *and* are capable of rational agency is not enough to argue for a dualism of human nature. What is instead possible is to think of autonomy (or freedom in the proper sense[8], not merely practical freedom) as another nature or, in other words, a purposive nature. It is true that autonomy means to act under self-imposed moral constraints and independently of any heteronomous sensuous (in these terms, natural) motives; however, the action of jumping into the water, at my own risk, to save someone endangered is still something which cannot disregard the (natural) laws of fluid dynamics or the principles of human anatomy.

8 See *KpV*, 5: 47 (178), the moral law is "a law of causality through freedom and hence a law of the possibility of a supersensible nature".

Autonomous agency still disregards heteronomous sensuous motives, though it does not disregard nature *in toto*, since all autonomous actions take place in the realm of the natural world. As it is very well known, Kant's point was to find what unifies the two causal chains of nature and freedom. Beyond instinct and rational agency, which could be regarded as mere natural dispositions (though it is not easy to think of rational agency as a *genetic material*), autonomous agency thus seems to display another nature. When Kant says that reason knows no boundaries to its projects, he is first of all stating that reason does have projects.

What if one of these projects is that of providing reality, i.e. actuality, in the natural world, to the faculty of *homo noumenon*, that is autonomy? If that were the case, there is no need to abandon the first three propositions of *Idea*, though it is clear they are not enough. Nor is the progressive view of history[9] endangered. Kant is not saying that human nature is biologically destined to perpetual peace but, rather, that, given that nature could always represent an obstacle to reason's projects, reason needs something more than practical freedom to achieve its end. What is then required? Three things, mainly. First, autonomy as the faculty of acting according to self-imposed moral constraints and independently of any sensuous motives. Second, a civil and rightful constitution as the guarantee of external freedom. Third, moral politicians as agents within the gap between the first two requirements, namely agents whose main task is to plan and develop all the necessary means to the end of perpetual peace. If reason is 1) part of our nature and 2) the faculty to determine our (second) nature as autonomous agents, there is no need to exclude the development of our natural dispositions (propositions 1-3) from the steps through which we achieve perpetual peace. To sum up, the development of our natural dispositions is not merely natural, for reason brings our nature far beyond instinct and sensuous inclinations, either by resisting or completely disregarding them.

Let me conclude by saying some words on the last page of *Kant's Political Legacy*, for it raises a significant issue. Here we read that the distinctive marks of a moral politician are not easy to find. As Caranti puts it,

> she must have assigned primacy to the moral law over self-love. This secures a firm adherence to the principles of right, an adherence 'for the right reasons', as Rawls would put it. She must have acquired solid and wide knowledge of the empirical laws relevant to her decisions, that is potentially all those of the social sciences plus history. She must be endowed by nature with a talent to know how and when scientific and moral norms of various kinds are to be applied. And she must have strengthened and refined this talent through practice. Not a little thing. Actually, a thing for very few (Caranti 2017: 255).

Unfortunately, this picture is tragically true. In fact, the political problems we mentioned also result from the inefficiency of those – the few – who were expected to provide a solution. However, it is also true that political agency does not coincide with political leadership. What I recognize as one of the major legacies of Kant's political thought is precisely this shift from politics intended as the activity of a few specialists, to politics as a co-responsibility. This would also highlight the link Kant saw between morality and politics (see Fonnesu 2017), that is, between

9 On teleology and history see Wilkins 1966, Allison 2012, Cicatello 2016. See also Yovel 1980: 125-198.

internal and external law. Our autonomy, as the ground of our dignity, is one of the most relevant human capacities. What I think Kant teaches us, is that this capacity leads us to political responsibilities. By this I mean that being – *in esse*, not merely *in potentia* – truly autonomous agents, requires us to shape safe and rightful political environments. Thus, our autonomy is displayed not only when we decide to save someone's life, but also when we pay our taxes, excel at our job, make some sacrifices to provide the best education to our sons and daughters, and even when we make of our moral consciousness a bulwark against unfair laws. This is also why Kant thought that history is concerned with the narration of human actions intended as the appearances of freedom (*IaG*, 8:17 (108)). To one of these appearances – the French Revolution – Kant once seems to have dedicated the following words: "God, let your servant die in peace, for I have already lived this memorable day!" (Ypi 2014: 265). After all, that was a thing for more than just a few.

References:

Allison, Henry E. (2012), "Teleology and History in Kant: The Critical Foundations of Kant's Philosophy of History", in *Essays on Kant*, Oxford: Oxford University Press, pp. 236–253.

Blöser, Claudia (2015), "Degrees of Responsibility in Kant's Practical Philosophy", *Kantian Review*, 20(2): 183–209.

Caranti, Luigi (2017), *Kant's Political Legacy*. *Human Rights, Peace, Progress*. Cardiff: University of Wales Press.

Cavallar, Georg (2015), *Kant's Embedded Cosmopolitanism. History, Philosophy, and Edication for World Citizens*. Berlin and Boston: De Gruyter.

Cicatello, Angelo (2016), "Kant's Idea of History", *Epekeina*, 7(1-2): 1–15.

Ferrarin, Alfredo (2015), *The Powers of Pure Reason. Kant and the Idea of Cosmic Philosophy*. Chicago-London: The University of Chicago Press.

Ferrarin, Alfredo (2004), *Saggezza, immaginazione e giudizio pratico. Studio su Aristotele e Kant*. Pisa: ETS.

Fonnesu, Luca (2017), "Politik als Moral", in Mori, Massimo (hrsg.), *Vom Naturzustand zur kosmopolitischen Gesellschaft. Souveränität und Staat bei Kant*, Wiesbaden: Springer, pp. 169–189.

Fussi, Alessandra (2012), *La città nell'anima. Leo Strauss lettore di Platone e Senofonte*. Pisa: ETS.

Guyer, Paul (2017), "Kant and the Autonomy of Reason. An Interview by Luigi Filieri", *Studi Kantiani*, XXX: 9–26.

Kant, Immanuel (1788), *Kritik der Praktischen Vernunft*. Translated by Gregor, Mary J. In Gregor, Mary J. (ed.) (1996), *Immanuel Kant. Practical Philosophy*, Cambridge: Cambridge University Press, pp. 133–271. In the text: *KpV*, original pagination of the *Akademie-Ausgabe* (page of the English translation).

Kant, Immanuel (1784), *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*. Translated by Wood, Allen W. In Louden, Robert B., Zöller, Günter (eds.) (2007), *Immanuel Kant. Anthropology, History, and Education*, Cambridge: Cambridge University Press, pp. 107–120. In the text: *IaG*, original pagination of the *Akademie-Ausgabe* (page of the English translation).

Kleingeld, Pauline (2018), "The Principle of Autonomy in Kant's Moral Theory: Its Rise and Fall", in Watkins, Eric (ed.), *Kant on Persons and Agency*, Cambridge: Cambridge University Press, pp. 61–79.

Kleingeld, Pauline (2017), "Contradiction and Kant's Formula of Universal Law", *Kant-Studien*, 108(1): 89–115.

Kleingeld, Pauline (2011), *Kant and Cosmopolitism. The Philosophical Ideal of World Citizenship*. Cambridge: Cambridge University Press.
Krasnoff, Larry, Sánchez Madrid, Nuria, Satne, Paula (eds.) (2018), *Kant's Doctrine of Right in the Twenty-first Century*. Cardiff: University of Wales Press.
Longuenesse, Béatrice (2005), *Kant on the Human Standpoint*. Cambridge: Cambridge University Press.
Wilkins, Burleigh T. (1966), "Teleology in Kant's Philosophy of History", *History and Theory*, 5(2): 172–185.
Willaschek, Marcus (2003), "Freiheit als Bedingung für Verantwortung. Ein kurzes Argument für den Kompatibilismus", in Mischer, Sibille, Quante, Michael, Suhm, Christian (hrsg.), *Auf Freigang. Metaphysische und ethische Annäherungen an die menschliche Freiheit*, Münster: Lit Verlag, pp. 199–205.
Ypi, Lea (2014), "On Revolution in Kant and Marx", *Political Theory*, 42(3): 262–287.
Yovel, Yirmiahu (1980), *Kant and the Philosophy of History*. Princeton: Princeton University Press.

Luiđi Filijeri

Autonomija, dostojanstvo i istorija u Karantijevoj knjizi *Kantovo političko nasleđe*

Apstrakt

U radu razmatram neke važne teze Karantijeve knjige *Kantovo političko nasleđe*, čiji je cilj da ponudi konzistentni uvid u to kako možemo razvijati Kantovu političku misao i da odgovori do koje nam mere Kantovi zaključci mogu pomoći da kritički razumemo politički svet 21. veka. Prvo se usredsređujem na autonomiju kao osnovu dostojanstva i raspravljam o Karantijevim argumentima koji u pitanje dovode ekskluzivnost kategoričkog imperativa kao jedinog načela istinskog moralnog delovanja. Zatim, razmatram Karantijeve poglede na istoriju i postavljam pitanje da li se može prihvatiti da racionalna priroda ljudi sadrži teleološke – iako ne-biološke – elemente, dovodeći na na taj način u pitanje Karantijevu tezu o razdvajanju.

Ključne reči: autonomija, dostojanstvo, moralno delovanje, svrhovitost, istorija, ljudska priroda

UDK: 172.4 Kant I.
https://doi.org/10.2298/FID1804598C
Original Scientific Article
Received: 21.07.2018. Accepted: 12.08.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Luigi Caranti

REPLIES: AUTONOMY AND HUMAN DIGNITY. A REASSESSMENT OF *KANT'S POLITICAL LEGACY. HUMAN RIGHTS, PEACE, PROGRESS*

ABSTRACT

The paper centers on some problematic theses of my book *Kant's Political Legacy. Human Rights, Peace, Progress* (UWP 2017). This reconsideration is occasioned partly by comments I received and partly by my own process of self-criticism. I focus on the point that commentators have mainly criticized, that is, the link I suggest between human dignity and our capacity for moral behavior, or autonomy. The first part recalls the basic features of my Kant-inspired and yet in many regards anti-Kantian account of the relation between dignity and autonomy and replies to some criticisms received from orthodox Kantians. The second part is strictly connected to the first because it deals with the reasons we have to believe that we are autonomous. While in the book I sketched Kant's own reasons for the 'reality of freedom,' as he puts it, I focus now on Bojan Kovačević's suggestion to look at characters in novels written by artistic geniuses (in particular Leo Tolstoy) to find indirect evidence in favor of autonomy. This allows me to reflect on the kind of evidence one can legitimately expect in the proof at issue. Thirdly, I reply to a classical objection, ignored in the book, that impacts with equal force Kant's ethics and my own position. The problem concerns people with temporary or permanent impairment of rational capacities. If I let human dignity depend on our capacity for autonomous behavior, am I committed to the counterintuitive (and rather devastating) conclusion that children or people suffering from momentary or irreversible loss of rational capacity, and *a fortiori* of autonomy, do not have dignity and therefore do not deserve to be protected by human rights?

KEYWORDS

morality, autonomy, dignity, human rights, Kant, mental impairment, speciesism

In my book *Kant's Political Legacy. Human Rights, Peace, Progress*, as the title indicates, I attempt to show how Kant's ideas can be illuminating for three themes crucial both for contemporary politics and for political theory. In particular, I sketch a foundation of human rights considerably different from any foundational argument available today in the literature. Further, I provide an interpretation of Kant's model for peace, that I show to be profoundly different from the one assumed by the majority of scholars and by the research program – the democratic (or liberal) peace – that made Kant's pacifism known and studied well beyond the circle of interpreters. Finally, in the most speculative part of the book I offer a defense of Kant's teleology, which I take quite seriously as meant to provide objective

Luigi Caranti: Professor of Political Philosophy, DSPS, Università di Catania; lcaranti@unict.it.

reasons to believe that progress towards a future of peace is more likely than any other outcome. Almost two years have passed since the book was published and I had the opportunity to rethink some of my theses, either through a process of endogenous reconsideration, or stimulated by the criticisms attracted.

Since commentators have mainly focused on the thesis, central for my foundation of human rights, that humans have dignity because they possess autonomy, I devote the paper to a reconsideration of this tenet. The first part deals with the legitimacy of my Kant-inspired and yet in many regards anti-Kantian reading of the relation between dignity and autonomy. I was aware that my idea that agency not inspired by the Categorical Imperative could count as authentically autonomous would attract criticisms by orthodox Kantians. I take here the opportunity to defend myself to the extent that this is possible. The second part is related to the first in that it deals with the reasons we have to believe that we are autonomous. While in the book I sketched Kant's own reasons for the 'reality of freedom,' as he puts it, especially in the second critique, I start now from Bojan Kovačević's suggestion to look at characters in novels written by artistic geniuses (in particular Leo Tolstoy) to find indirect evidence in favor of autonomy. This allows me to reflect on the kind of evidence one can legitimately expect in the proof at issue. Thirdly, I reply to a classical objection to Kant's ethics, that impacts just as deeply on my own reading, in a way that I had not done in the book. The problem concerns people with what I call temporarily or permanently impaired autonomy. If, with Kant, we let human dignity depend on our capacity for autonomous behavior, it seems that we are forced to the counterintuitive (and rather devastating) conclusion that children or people suffering from momentary or irreversible loss of rational capacity (and *a fortiori* of autonomy) have neither dignity nor human rights.

1. Autonomy and Human Dignity

The working hypothesis of *Kant's Political Legacy*, part I, is that all current foundational accounts of human rights lack a crucial ingredient. This is the simple intuition, latent in all major documents of human rights, that human beings are *worthy* creatures, despite the atrocities for which they have been responsible. Within 'humanity', by virtue of which, the documents say, we have human rights, there is supposed to be something extraordinarily valuable and awe-inspiring that makes it obligatory never to treat humans below a certain threshold of respect. This kernel of value, which a foundation should spell out, serves as protection not only against the violations of our dignity that others cause to us, but also against the degradation we can bring upon ourselves.

In starting from this working hypothesis, my approach goes in a direction opposed to the practical compromise dear to founding fathers of the culture of human rights. Jacques Maritain and Eleanor Roosevelt famously stated that we all agree on HR and human dignity "on the condition that no one asks us why." If a foundation is supposed to convince a skeptic (for what other reasons would you want a foundation?), I think we must say something as to why human beings have dignity, why they are to be esteemed independently of their religion, citizenship, gender, race. Only if we do so will we be in the position to counter the arguments

of those who believe that some, but not exactly all persons deserve the equal respect promised by HR. In choosing this path, I knew that my position would appear as a form of ultra-orthodoxy, or perhaps of 'foundationalism', to use Tasioulas' derogatory label (Tasioulas 2015: 46–47). To make things worse, in construing the required account of human dignity I took my lead from Immanuel Kant and even if I quickly depart from him on a number of crucial points, this choice is bound to be perceived as somewhat bizarre. Among other things, Kant's account of human dignity seems to rest on his controversial doctrine of our belonging not only to the phenomenal, but also to the noumenal world, a very metaphysical tenet indeed.

Still, a Kantian foundation should not be discarded before one critically explores its potential. Kant offers one of the most compelling accounts of our dignity in the history of thought, and this cannot be fully irrelevant for HR. Moreover, silence on the reasons we have to believe that human beings have dignity will not convince those who think that it is permissible to treat some of us below certain standards. If we do not know *why* we have dignity, it is very likely that we will not know what this entails, and we will have a difficult time clarifying why and when human rights trump normative considerations in the form of maximization of general utility, the furthering of certain ideologies, or the protection of traditional practices. In addition, failing to unpack the notion of human dignity gives no direction regarding related questions crucial for the theory of human rights. It is unlikely that we will know what rights should count as *human* rights unless we know why human rights should exist in the first place. It may not be impossible to come up with an answer to this and other similar questions before we clarify what dignity is and entails. But few would deny that some clarity on human dignity will help greatly in dealing with those questions.

If a Kant-inspired foundation of human rights is worth exploring, the following might be a reasonable way to proceed. The central thesis, in itself far from original, is that humans have dignity because they are capable of a unique form of freedom, namely autonomy, which allows them to perform actions motivated by what they take as morally obligatory. Autonomy is not to be understood merely as the ability to choose one's path in life, or as the ability to be rational in the sense of purposive agents. With Kant, we refer to a capacity distinct from and 'higher' than practical freedom. We have in mind the ability to act under self-imposed moral constraints.

The distinction between the common-sense notion of autonomy (self-determination) and the one at work here (capacity for moral agency) is important to understanding the way in which our approach links the possession of a faculty to the intrinsic worth that entitles humans to the protections of human rights. In fact, why should the sheer possession of a capacity ground any worth? The fact that humans have a peculiar capacity hardly grounds any special right or entitlement. Ours is probably the only species that kills for sheer amusement (cats may be another example) or for cold, long-term calculation of interest. It may also be the only one capable of lying with a high level of sophistication. Now, these peculiarities obviously do not ground any merit and therefore entitle us to no special protection.

Things, however, are different with autonomy. This feature is not only peculiar to, or most developed in, the human species. It also has an intrinsic value, as it shows humans as capable of behavior that exacts respect. We are not merely free;

we are free to act on a principle that we perceive as morally obligatory. And precisely because we have this capacity, precisely because morality is within our reach, we are entitled to an amount of respect unfettered by contingent circumstances.

The argument is largely inspired by Kant and yet it shares with Kant only the intuition of a link between a capacity for moral agency and dignity. Down the argumentative path, whose details I cannot present here, I take a couple of turns that make my approach significantly different from any Kant would be ready to endorse. To begin with, I hold that autonomous agency need not be restricted to agency under the auspices of the Categorical Imperative. I argue that authentic, duty-based agency occurs even when people act under different moral imperatives, such as the Golden Rule or other moral formulas, including the maximization principle dear to utilitarians.[1] Also, and perhaps even more against Kant, I deny that autonomous agency is a peculiarity of human animals, defending the weaker thesis that humans are merely capable of this form of agency *to the highest degree of development* in the animal world.[2]

As it was easy to predict, it is on these two major departures from Kantian orthodoxy that readers have turned their critical eye. In the context of this paper I reply to Luigi Filieri and Natalia Lerussi who both react to my severing autonomy from the Categorical Imperative. Filieri reminds me that the Categorical Imperative is not a law with its own normative content. Rather it is merely the form a law (better, maxim) must have if it must be permissible. So the Categorical Imperative merely prescribes that the maxim you are adopting could be universalized without logical or practical contradiction (formula of universalization) or could be willed to hold as a natural law (formula of the law of nature). In and of itself, however, it prescribes no content or matter. It follows that we do not need to alter Kant's idea that autonomous agency rests on the Categorical Imperative to have all the latitude in the moral law I want to avoid parochialism. If I understand him correctly, Filieri thus wants to question the necessity of liberalizing the boundaries of the moral law's content as to include the Golden rule or other allegedly less parochial formulae for the sake of transcultural validity.[3]

---

1 As it will be shown soon, I have radicalized my position since the publication of the book because I now argue that *any* moral formula can serve as inspiration for authentic autonomous behavior, on the sole condition that it is a general principle that does not merely serve my selfish interests.

2 Obviously for Kant human beings are not the sole entities capable of moral agency. He thinks that purely rational agents or partly rational entities whose existence we may speculate about, like respectively angels or intelligent extraterrestrials, are equally autonomous. Since what confers autonomy is ultimately our rational nature, it follows quite naturally that all rational entities are autonomous. In fact, as Kant puts it: "unless one wants to refuse the concept of morality all truth and reference to some possible object, one cannot deny that its law is so extensive in its significance that it must hold not merely for human beings but for *all rational beings in general* [*alle vernünftige Wesen* überhaupt]" (GMS 4: 408; see also 410n., 412, 426, 431, 442). So when we say that humans are the sole autonomous species, we implicitly mean 'of which we have experience", "part of the sensible world," or the like. This point that humans are autonomous simply because they embody a property (rationality) that other species may share with us will be crucial to show, in the third part, that Kant's position and ours have nothing to do with specisism.

3 For a criticism similar to Filieri's see Klein 2018.

Natalia Lerussi cuts deeper and presses me to determine whether my severing autonomy from the Categorical Imperative does not imply that I am making the law underpinning moral behavior fully arbitrary. In fact, it becomes incumbent on me to define with some precision which moral principles, in addition to the Categorical Imperative, are to be considered as legitimate bases of autonomous agency. I need to define new boundaries for the content of moral principles that, if adopted, are supposed to generate autonomous agency. Most fundamentally, the problem is whether only reasonably well-specified moral principles can inspire autonomous agency or any principle can do that on the sole condition that agents adopt it independently of any possible selfish interest. Is it the case that agents are autonomous even when they adopt heinous 'moral' principles? In the book I gave the example of the sacrifice of Nazi officials that decided not to surrender to the Allies. Is their sacrifice an act as autonomous as to the one imagined by Kant in his famous example of someone who refuses to give false testimony against an innocent man, knowing that this will lead him or her to death (KpV 5: 30)?

Let me reply to my two critics in the order I presented their comments. In response to Filieri, the main point is the following. In construing the Categorical Imperative basically as a formal requirement that any acceptable maxim must satisfy, the impression is that he is understanding it as indifferent to any specific normative content. As Filieri puts it, "the moral law legislates the mere form of an action, whatever its content may be" (Filieri 2018: 591). Obviously this cannot mean that the Categorical Imperative does not discriminate between permissible and impermissible maxims, checking only that they are formally correct, which I trust to mean amenable to universalization or transformation into a law of nature. This would mean that all the CI does is to check that I may be able to live or want to live in a world in which my maxim becomes a universal law. But this cannot be all the CI does, at least not without abundant qualifications.

One way to see this is to focus on what Henry Allison calls 'false positives' (Allison 2011 191–196), that is, examples of clearly and unequivocally impermissible maxims that could be easily either universalized or turned into a universal law. The maxim imagined by Paul Dietrichson and discussed by Barbara Herman (Herman 1993, 113–131) is that of killing babies who prevent our sleep by crying at night. If this were to become a law of nature, it is difficult to detect any logical or a practical contradiction. Let me give another example. Imagine that I want to live in a world in which people deceive one another at every opportunity. A sort of world in which it is universally accepted that if one is smart and clever enough to be able to fool others, he should be allowed to do it. Again, there are defensive strategies a supporter of the law of nature formulation could adopt against these counter examples, but it does appear that no practical or logical contradiction arises the moment we universalize this maxim or transform it into a law of nature.

Although a discussion of this would be way too long, and it would probably intersect the much discussed issue of the equivalence between the three formulas of the Categorical Imperative – an issue that has still excellent interpreters holding opposing views – the point relevant for replying to Filieri is that the CI does contain some 'matter' and not only the sheer form of universalizability. And this matter is best expressed by the formula of humanity, where the 'matter' famously is

the prescription to use "humanity", whether in your own person or in the person of any other, always at the same time as an end, never merely as a means" (GMS 4:429).[4] Once the matter is not forgotten, the above mentioned false positives immediately vanish. To focus on my own example: it does even if I, or all human beings, want to live in a world in which we cheat and deceive each other any time we have the opportunity to do so. Such a maxim, no matter whether endorsed universally or turned into a law of nature, violates the humanity in ourselves and others. As such it would be disqualified as impermissible. This shows is that the Categorical Imperative does have its specific normative content and cannot be reduced to a merely formal requirement, on pain of making Kant's ethics not only vulnerable to Hegel's famous objection of 'empty formalism,' but more importantly close to a parody. And precisely because CI has a specific content, the question opens up whether we want to make that specific content (respect for humanity as an end in itself) as a necessary condition for autonomous agency.

Natalia Lerussi shares with Filieri an uneasiness related to my liberalized version of autonomy. Yet her worries are more profound.[5] To use her words:

> if autonomy is guaranteed, as Caranti wants it to be, through the determination to act in conformity with any general principle, we are confronted with an undesirable alternative: if the principle in question is not given to the agent by reason, but rather they must decide it voluntarily, the question arises of the *lack of a criterion* for deciding between general moral principles. The consequence is that the decision to subordinate oneself to one general principle or another appears arbitrary. If, on the other hand, the principle in question is given to the agent, not by means of their reason, but rather, for example, through culture, the state or religion, it becomes difficult to maintain an acceptable concept of "autonomy". This second alternative also brings the difficulty that it does not permit us to establish a sharp division between moral determination mediated by, for example, the Golden Rule and by a morally perverse principle, such as the determination to act in accordance with the will of the Führer in Nazi Germany. (Lerussi 2018: 630)

Although Lerussi phrases her criticism in the form of a dilemma, it seems that her basic point is that if autonomy is construed as the capacity to act on any general unselfish principle freely endorsed by the agent, then two consequences (I do not think they are two horns of a dilemma) follow. To begin with, it seems that I am allowing for the possibility that the law come from a source other than our pure reason: tradition, religion, culture, accepted morals and so on. How could one call this autonomy? On the other hand, if agents are autonomous merely because they act on a principle, with the sole condition that it must not respond to selfish interests, then I am opening to all sorts of perverse principles. For example, I am committed to say that a principle of unfettered loyalty to Nazism is just as good Kant's respect for humanity to generate autonomous agency. Both points are taken as self-evidently devastating for my conception of autonomy. The consequence of the first is that my conception of autonomy is stained, to say the least, with heteronomy. The consequence of the second is that I go very close to relativism.

4 References to Kant's works are indicated with standard abbreviations and follow the *Akademie* pagination.

5 For a criticism similar to Lerussi's see dos Santos 2018 and my reply Caranti 2018.

Replying to Lerussi allows me to confess that my position has become much more radical than it was in *Kant's Political Legacy*. While I was well aware of both points in the book, I think I am ready now, in a way that I was not before, to bite both bullets. What interests me is the human capacity to ignore one's selfish interests for the sake of some greater cause, even if this greater cause does not coincide with any of the most popular moral imperatives and the motivation for action is not 'pure' in the sense of originating from the recognition of the sheer obligatoriness of the course of action in question. Just to give an example, I think that the Christian martyrs who refused to worship the emperor to avoid the persecutions carried out by the pre-Constantine Roman authority were sacrificing their lives not because they had the Golden Rule or the Categorical Imperative in mind, but out of a free endorsement to *their* 'greater cause', namely obedience to their God. And I want to say that this conduct is as autonomous as the conduct, to use Kant's example, of the person who refuses to give false testimony against an innocent man just because he thinks this is the right thing to do. Quite in line with common sense, I think that any sacrifice of one's interests for the sake of a greater cause is a potential act of autonomy.[6]

In addition to readiness to sacrifice one's immediate interests, in the book I introduce a second condition for autonomy, namely the ability for this greater cause to pass minimal moral standards. To be frank, I am no longer certain that my theory needs this second condition, which, among other things, brings with it the necessity to have a fairly detailed account of what these 'minimal moral standards' are. In other words, I am now inclined to think that even the refusal by Nazi officials to surrender in Berlin and their sacrifice for loyalty to the Führer should be construed as an autonomous act. I think now that we can legitimately admire this act without compromising in the least our strict and profound reprobation of the intrinsic merits of the cause that inspired it. I also think that this position is closer to our moral intuitions as one may think at first sight. In fact, we do admire acts of 'integrity' and self-denial, independently of the merits of the cause for which people sacrifice themselves. We do admire the 'integrity' of Nazi officials who sacrificed their lives for what they believed in, no matter how much we despise the ideology for which they sacrificed.

Finally, and perhaps less problematically, let me reinforce a point about which I did not change my mind and that is relevant to both Filieri and Lerussi. I think that we should be more 'flexible' than Kant about the standards of moral behavior to hope that our autonomy/dignity based conception of human rights be sufficiently consonant with the diverse moral sensitivities of the world. If we liberalize our conception of autonomy making it equivalent to the individuals' ability to act on principle they recognize as obligatory over and above their immediate or selfish interests, we come interestingly close to a conception of human value that has a chance to enjoy transcultural validity. As I try to show in a dense section of chapter 3 of my book, all major religious traditions, at least if reasonably interpreted, consider individual autonomy in my liberalized understanding as a central component of human worth (Caranti 2017: 95–104). And this ubiquitous presence

6 And I say 'potential' because there should not be, hidden behind the greater cause, some sort of unconscious selfish interest. This possibility was well known to Kant and I concede, as he does, that we can be certain about this.

is no accident. There is, so to speak, an *a priori* reason why no religion can fail to acknowledge the central, if not prevalent, value of individual autonomy. Since adherence to a received set of rules, including divine ones, is worth nothing unless it is free and uncoerced, religions cannot help but pay homage to individual autonomy, whether it does so explicitly or not.

2. Can Fictional Characters 'Prove' Individual Autonomy?

As already seen, I claim that the most promising basis of human dignity is our capacity for moral behavior, that I equate, in agreement with Kant, with our autonomy.[7] Once such a strong emphasis is placed on autonomy, one should be certain that autonomy is within our reach. Unless some compelling reasons are given to believe that we do have this capacity, it will appear as bizarre to let human rights rest on such uncertain possession. In fact, without a solid proof the impression will be that my foundation, in relying on some sort of 'invisible basis', produces more, not less skepticism.

Famously Kant himself was challenged to find a proof of the reality of freedom/autonomy. For some time, as shown by Dieter Henrich (Henrich 1973: 107-110), he thought that he could give a theoretical proof of this property of the will. Then he abandoned this plan as evidently incompatible with the limits he himself had come to set for our cognition through the publication of the first critique. He thus moved in the *Groundwork of the Metaphysics of Morals* (1784) to the second best strategy available, that of arguing that autonomy must be necessarily presupposed as a property of the will of entities who take themselves as agents: "now I assert that every being who cannot act except *under the idea of freedom* is by this alone – from the practical point of view – really free." (Gr 4: 448). In other words, I am free because if I weren't, I should stop considering myself an agent. Later, unsatisfied also by this 1784 strategy, Kant attempted to argue in the *Critique of Practical Reason* (1788), through what most commentators see as a 'great reversal' in the argumentative strategy, that the reality of freedom is to be grounded through an appeal to the 'fact of reason', namely the immediate consciousness we have of the binding force the moral law has on us.[8] Remember here the example of the refusal to give in to the threat by the immoral prince who asks from us a false testimony against an innocent man. On my reading, the fact of reason is precisely the immediate and inescapable consciousness that a) refusing is what would be right to do and b) that we *could* refuse. The immediate consciousness of duty reveals that we *are* free, or better that we *take* ourselves as such.

In short, all Kant was able to do, actually all he could do given the general picture of his philosophy, was to provide an argument 'from the practical point of view' for the reality of autonomy. Far from being a scientific or logical proof, his argument

7 This is what Henry Allison calls the Reciprocity thesis (Allison 1990: 201-213).

8 Although the question is still debated among Kant scholars, the vast majority of them do see the difference between the argument of part 3 of the *Groundwork* and the argument in the second critique, turning on the 'fact of reason,' as a reversal. Of this opinion Karl Ameriks, *Kant's Theory of Mind*, Oxford: Clarendon Press 1982, p.226 and Henry Allison (Kant's Theory of Freedom, CUP 1990, pp. 227-90).

only showed (and was meant to show) that autonomy cannot be taken away from the description of our agency without taking away also the very notion of an agent. Even in the most sophisticated version of the argument, which I take to be that of the second critique, freedom or autonomy is revealed through a fact that cannot be displayed, but at most replicated any time in our consciousness, when we encounter duty and discover that we could follow it in any circumstances. This is not insignificant and yet falls short of something we can literally 'point to' a sceptic.

Bojan Kovačević's extended comments on my book contain a lot of interesting suggestions but I want to take them mainly as speaking to this central problem of the proof of our autonomy (or lack thereof). Kovačević's central idea – I reckon – is that we need to live with the fact that we won't ever be able to obtain a proof of our autonomy. Aiming to that would be an act of intellectual hubris, oblivious of the Aristotelian recommendation to search in a discipline for a degree of precision fit for the discipline itself. In the same way in which it would be absurd to be content with a low degree of certainty and precision in a mathematical proof, it would be equally absurd to ask for mathematical precision in social science, let alone philosophy. Nonetheless, Kovačević continues, this does not mean that we have to take our autonomy as an article of faith. We can collect some hints about the reality of this capacity, that would add up to Kant's proofs 'from the practical point of view', from the work of artistic geniuses such as Leo Tolstoy. Literature at times presents us characters that, despite their being a product of human imagination, can bear witness to the human capacity to make the moral law the supreme principle of their actions.

The Tolstoyian character preferred by Kovačević is the nobleman Nekhlyudov in the novel *Resurrection*. Nekhlyudov leads a life full of pleasures to the extent of becoming bored at every spark of beauty in life and insensitive to human suffering. But when he is called as juror in a case against a prostitute who is accused of murder, and he recognizes the girl he once seduced and abandoned, thereby causing the beginning of all her misfortunes, Nekhlyudov hears distinctly the call of duty and devotes his entire life to mitigate the girl's suffering, up to the point of following her to Siberia and sharing with her the condition of inhumane imprisonment. The obvious moral of the story is that, in giving up all his pleasures and choosing the hardships involved in helping the girl, Nekhlyudov is our autonomous hero.

Of course Kovačević does not take into consideration the possibility that our hero is moved to help the girl by a feeling of empathy or compassion that for Kant would make his helping no less heteronomous than his past pursuing of (boring) pleasures. But this is not a major difficulty. At least on our liberalized account of autonomy, what matters is the capacity to act on a principle to the detriment of our immediate selfish interests. The fact that the adoption of that principle can be somehow linked to some higher and more distant inclination (compassion in this case) does not remove the fact that the agent had to prioritize the feeling of compassion over the inclination to avoid all the pain involved in helping the girl. A moment of deliberation about what should be done, as well as of distance from his immediate impulses was therefore presupposed, something which suffices for construing Nekhlyudov's conduct as autonomous under our liberalized account.

The central point to be discussed about Kovačević's reading is a different one and concerns the extent to which art can provide evidence of our autonomy. Novels

are 'proofs', obviously, only to the extent in which readers can connect to the characters presented and identify with them in a manner not dissimilar to the one Kant expects from us when he presents the above cited 'story' of the noble man who resists the prince's threat. After all, Kant's thought experiment is no less a product of the imagination than characters in a novel. It all depends on whether we find it easy to identify with them and therefore find 'evidence' of the human capacity revealed by their actions.

This opens up the general question of the relevance of examples of moral heroism for the sake of proving our autonomy. Not only literature but also history contains stories of moral heroism. But here comes the problem. Literature and history also have numerous stories of moral abjection. The point is not so much whether the atrocities humans have carried out cancel out their good deeds. This was the way I phrased the problem in the book and solved by insisting that no cruel thing actually done can erase a capacity. Since we recognize worth in human beings for what they are capable of, not for what they have done, this worth cannot be touched by actual human evil. The truly difficult question is rather about a comparative assessment of *capacities*. If we focus on a special *capacity* humans have, shouldn't we also consider humans' *capacity* – as special as autonomy – to reach a level of baseness animals are not capable of?

The answer to this objection must likely take one of the two following forms: either one denies that the bad capacity in question is really special of human beings, at least in the same way or to the same extent in which the good one is. For example, one could question whether cruelty towards members of our species or of others is truly a peculiarity of humans. It would not be difficult to find examples in the animal world of 'cruel' behavior. After all, also cats kill not only for the sake of feeding themselves and many mammals kill or eat their offspring or that of other members of the same species.[9] Alternatively, and more promisingly, one could point out that the goodness of autonomy, its comparative advantage if you want, is not exhausted either by its peculiarity. In addition to that, unlike the capacity for extreme cruelty that always seems to serve some selfish need, autonomy presupposes a distance from our impulses, and a deliberating freedom that may even lead to our sacrifice, that makes it distinct and superior than the other 'special' capacity. Only our ability to follow a moral law severed from our selfish impulses reveals us as demi-gods, finite creators, absolute rulers of our own lives. To use the expression by Pico that Kovačević much appreciates, autonomy and only autonomy reveals a spark of 'divine' in us.

3. Impaired Autonomy

If human rights are grounded on a capacity (autonomy), does that mean that humans who are temporarily or permanently impaired in that capacity do not enjoy the protection of human rights? If that were the case, my foundation would run into a fatal difficulty. Human rights are commonly understood as tools in the hands of the weak to defend themselves from abuses of all sort by the strong. If we end up

9 Here is an incomplete list: hamsters, rats, lions, meerkats, and about forty different species of primates.

denying human rights to people who are particularly weak, like those with mental impairments, then something essential has gone wrong.[10]

Fortunately, neither Kant's moral thought nor my foundation are bound to such a counter-intuitive conclusion. First let's distinguish the cases we are dealing with. In the case of children, one can hardly quarrel with the fact that they are given less rights than adults and that this happens precisely for the fact that their rational capacities are not fully developed. For example, children do not have the right to vote because it is assumed that their ability to think autonomously is not developed sufficiently. At the same time, their potentiality for reaching full autonomous status is part of the reason why they have all other rights (human or not) we usually attribute to people. Actually, sometimes they have certain rights – like access to certain state benefits designed to help their development – adults do not have. Hence we attribute more or less rights to children precisely by using their autonomy (or potential development thereof) as a moral compass.

Analogously, elderly people who have lost in part or full their ability to think are denied certain rights (think of all the restrictions that come with a declaration of *non compos mentis*) and yet keep other rights because we still respect them for what they were once capable to do (think and act autonomously). Respecting a rational creature when its capacity for fully autonomous behavior is in place seems to entail respecting her even when she happens to lose – in part or fully – that capacity. Marc cannot be said to be truly respecting Charles *now*, when Charles is a fully autonomous agent, if it is understood that Marc can do whatever he wants with Charles the moment the latter loses his ability to think. If that is the case, then Marc was not respecting Charles even when Charles was healthy.

The same point can be seen from another angle: Imagine how odd it would sound if I were to tell you: "I respect you because you have this wonderful capacity for moral agency. Hence, I make sure that you enjoy all the rights that come with that status. But also rest assured that the moment that capacity will vanish I will stop considering you a subject of rights up to the point that you are degraded to the level of animals or the like". One can certainly restrict, like in the previous case, the number of rights enjoyed by the impaired individual (we deny the right to vote or to use property to someone after *non compos mentis* is declared). And yet the subject does not lose all its rights. While there is latitude for discussion about precisely which rights (human or not) the person should retain, what matters here is the principle. We cannot ignore her (intact) capacity for suffering, for having interests and needs without affecting negatively, in retrospect, the way in which we were treating her. *Mutatis mutandis*, the same can be said for people who have lost – temporarily or permanently – their thinking ability because of illnesses or accidents of different sorts.

In my mind, a whole different case is that of people who were born with severely diminished rational capacities that we know will not improve in the future or with no

10 Obviously this objection is nothing but the reformulation, in the language of human rights, of a classical complaint against Kant's moral thought. In its general form, the complaint is that Kant links so closely morality and dignity to human reason that he risks denying children, some elderly, individuals with temporal or permanent mental impairment the protection that moral status entails.

rational capacity at all. In these cases, I submit, we are permitted to treat these individuals in the same way in which we treat other sentient animals. And that our respect should be made dependent (and perhaps proportional) to their capacity for suffering.

This line of thought is slightly different than the one advanced by Allen Wood (1998) and Onora O'Neill (1998) to deal with this classical impasse for Kant's moral thought. Wood, for example, argues – and O'Neill agrees – that we should abandon the idea that rational nature is to be respected always as embodied in persons. We should also respect rational nature in general, which entails respecting "fragments of it or necessary conditions of it, even where these are not found in fully rational beings or persons" (Wood 1998, 198). This is what happens with small children and people who have severe mental impairments or diseases which deprive them, either temporarily or permanently, of the capacity to set ends according to reason. My argument is not that we respect the fragments of the rational nature these particular subjects embody. Rather I am arguing that we respect the *full* rational capacity these subjects have the *potentiality* to reach or used to embody. It remains an open question (to me at least) whether Wood and O'Neill would agree with my suggestion that these cases are radically different from the ones posed by human subjects who never were and never will be rational.

Recently, in the context of a reply to Peter Singer's famous attack on speciesism (Singer 1975), Shelly Kagan (Kagan 2016) has made a compelling case to show that one can both identify the source of our worth in rational nature (in my language, autonomy) and be able to attribute to people with diminished rational capacity the same protection we attribute to 'normal' humans. While the argument in my opinion ultimately fails, it contains a crucial intuition that helps us to understand better my defense against the objection under consideration. Kagan starts from the premise that the reason why we attribute to ourselves a higher status than animals is not generically what makes us human, but specifically the part of our genetic setting that constitutes us as *persons*, which Kagan rather loosely defines as a rational capacity and self-awareness. We can see this when we realize that we attribute the same degree of moral consideration to entities that we identify as persons, but that are not human. We do not think, to use Kagan's example, that an evil act against superman or E.T. is less of a problem because these two individuals are not human. The offence is serious because superman and E.T. are *persons* in the specified sense. This intuition, that I fully share, shows why Kantians are not speciesists. Going back to a point we already made, for Kant what confers dignity to us is not our belonging to a species, but the fact that this species, at least in the vast majority of its members, *embodies* a property that non-human entities like angels or intelligent extraterrestrials embody or could embody as well. More importantly, Kagan's emphasis on personhood as the true seat of value helps us to see why impaired autonomy is not an insurmountable obstacle for our (and Kant's) account. The point we made above can now be reformulated as follows: those who have the potentiality to be or used to be persons can never be said to be truly respected unless one extends the privileges that come with that status to the time in which respectively personhood has not fully actualized or is lost. [11]

11 Kagan, however, does not use this intuition to arrive to the natural conclusion regarding humans who were born with no rational capacities. Instead of saying, like us, that these

Conclusion

The relation between human dignity and autonomy, combined with my – we could say – deflationary view of the conditions that make human agency autonomous, is undoubtedly the nerve of the foundation of human rights I offer in *Kant's Political Legacy* (first part). Among the many difficulties my foundation encounters two stand out as particularly acute. One is the problem of what I called elsewhere the possibility of deontological barbarians, that is, people who satisfy my condition of autonomy (they act on principles that do not serve their selfish interests) and yet are clearly perverse from a moral point of view. Nazi officials are unfortunately only one easy example of this human category far more populated than one may think. While in the book I tried to avoid the thesis Nazi officials display autonomy when they sacrifice for the Führer, I am now inclined to accept this conclusion and to see it, perhaps too benevolently, not as a vice, but as a virtue of my account. After all, on any reasonable interpretation of human rights, they have to apply and defend from inhumane and degrading treatment not only good people, but also the worst of us. And focusing on the capacity to sacrifice for a 'higher' cause, no matter how perverse this cause is, may not be the wrong move in order to offer protection even to these individuals.

Secondly, the problem of proving that we are autonomous, obviously not a secondary concern for a position that puts so much emphasis on this faculty, shares all the difficulties originally encountered by Kant. In fact, my liberalized version of autonomy is no less a 'breach' in the natural course of events than Kant's original notion. What I tried to argue, taking myself as an orthodox Kantian here, is that we are obliged to look for a proof 'from the internal viewpoint,' that is, from the way we are not merely accustomed but obliged to look at ourselves. In other words, the proof of autonomy must turn on the self-image of agents that is forced upon us by the immediate consciousness of our capacity to act, and sacrifice if necessary, for a selfless cause. In this vein, fictional characters do contribute to the proof of our autonomy – I think Kovačević's basic intuition is correct – on the sheer condition that our identification with them is strong.

Finally, I replied to a classical objection to Kant's ethics that applies with equal force to my account, although I did not address it in the book. This gave me the opportunity to show not only that the objection is considerably less devastating than what critics think, but also that neither Kant's nor my position are stained by speciesism, a very welcome and unexpected result in the context of a foundation of rights that seem to risk that stain in their very concept.

---

humans are not persons and therefore are not entitled to human rights, Kagan argues that, while these individuals are not 'persons', they are nonetheless members of a 'persons-species', that is, of a species that *most of the times* display in its members the required feature of personhood. In virtue of this belonging to a person-species these members deserve a better treatment than primates or animals in general. Ironically, Kagan seems here to fall in the position he was attacking. If being a member of the species brings to individuals a special value *merely* in virtue of what the species *normally* displays, and that, by definition, these particular individuals born with no rational capacity do not, why should we grant them the same protection we reserve to individuals endowed with the relevant feature? It seems that one could do so only through the endorsement of that speciesism from which Kagan's modal personism was supposed to be sharply distinguished.

Bibliography

Allison, Henry (1990), *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
Allison, Henry (2011), *Kant's Groundwork of the Metaphysics of Morals. A Commentary*. Oxford: Oxford University Press.
Caranti, Luigi (2017), *Kant's Political Legacy. Human Rights, Peace, Progress*. Cardiff: University of Wales Press.
Caranti, Luigi (2018), "Reply to my critics" *Estudios Kantianos* 6 (1) 2018: 117–129.
Dos Santos, Robinson (2018), "Sobre autonomia e dignidade como base para justificação dos direitos humanos" *Estudos Kantianos* (6) 1: 39-48.
Filieri, Luigi (2018), "Autonomy, Dignity and History in Caranti's *Kant's Political Legacy*", *Philosophy and Society*, 29 (4): 586-597.
Henrich, Dieter (1973), "Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft" in G. Prauss (Ed.) *Kant: zur Deutung seiner Theorie von Erkennen und Handeln*, Köln: Kiepenheuser & Witsch: pp.77-115.
Herman, Barbara (1993), *The Practice of Moral Judgment*. Cambridge MA: Harvard University Press.
Kagan, Shelly (2016), "What's Wrong with Speciesism" *Journal of Applied Philosophy* 33 (1): 1–21.
Klein, Joel Thiago (2018), "Uma Fundamentação Kantiana dos Direitos Humanos, da Paz e do Progresso: Um Debate com Luigi Caranti" *Estudos Kantianos* (6) 1: 97–106.
Kovačević, Bojan (2018), "Subjective Universality of Great Novelists as an Artistic Measure of History's Advance towards Actualising Kant's Vision of Freedom", *Philosophy and Society*, 29 (4): 567-585.
Lerussi, Natalia (2018), "Luigi Caranti, *Kant's Political Legacy. Human Rights, Peace, Progress*", *Philosophy and Society*, 29 (4): 629-633.
O'Neill, Onora (1998), "Kant on Duties Regarding Nonrational Nature: II.", *Proceedings of the Aristotelian Society* Supplementary Volumes, 72: 210–228.
Tasioulas, John. 2015. 'On the Foundations of Human Rights', in R. Cruft, M. S. Liao and M. Renzo (eds), *Philosophical Foundations of Human Rights*, Oxford: Oxford University Press, pp. 45–70.
Singer, Peter (1975), *Animal Liberation*. London: Jonathan Cape.
Wood, Allen (1998), "Kant on Duties Regarding Nonrational Nature: I." *Proceedings of the Aristotelian Society* Supplementary Volumes, 72: 189–210.

Luiđi Karanti

Kant's Political Legacy. Human Rights, Peace, Progress
Odgovori: autonomija i ljudsko dostojanstvo. Preispitivanje knjige Kantovo političko nasleđe: ljudska prava, mir i progres

Apstrakt

Članak se usredsređuje na neke problematične teze moje knjige *Kantovo političko nasleđe. Ljudska prava, mir, progres* (UWP 2017). Preispitivanje je jednim delom podstaknuto komentarima koje sam dobio a drugim delom procesom samo-kritike. Osnovni problem koji su komentatori uglavnom kritikovali jeste moj pokušaj da dovedem u vezu ljudsko dostojanstvo i našu sposobnost za moralno ponašanje, ili autonomiju. Prvi deo članka podseća na osnovna obeležja mog Kantom inspirisanog a ipak u mnogo čemu anti-Kantovog shvatanja veze između dostojanstva i autonomije i odgovara na neke od kritika koje upućuju ortodoksni Kantovci. Drugi deo je striktno povezan sa prvim budući da razmatra razloge zbog kojih mi verujemo da smo autonomni. Dok sam u knjizi skicirao dokaze samog Kanta za "stvarnost slobode" kako on to kaže, sada se usredsređujem na predlog Bojana Kovačevića da se indirektni dokazi u korist autonomije potraže u likovima iz romana kao delima umetničkih genija

(posebno Lava Tolstoja). To mi omogućava da razmišljam od tome kakva se uopšte vrsta dokaza može legitimno očekivati u ovoj stvari. Treće, odgovaram na klasičnu primedbu, koja se u knjizi ignoriše, a koja podjednako pogađa Kantovu etiku i moju sopstvenu poziciju. Problem se odnosi na ljude sa trenutnim ili trajnim poremećajem racionalnih sposobnosti. Ako dozvolim da ljudsko dostojanstvo zavisi od naše sposobnosti za autonomno delovanje, da li se onda suočavam sa kontraintuitivnim (i prilično razornim) zaključkom da deca ili ljudi koji pate od trenutnog ili nepovratnog gubitka racionalnog kapaciteta, i *a fortiori* autonomije, nemaju dostojanstvo i stoga ne zaslužuju da budu zaštićeni ljudskim pravima?

Ključne reči: moralnost, autonomija, dostojanstvo, ljudska prava, Kant, mentalno oštećenje, vrstovnost

IV

REVIEW ESSAY

KRITIČKI OSVRT

UDK: 316.42(4-11)
https://doi.org/10.2298/FID1804615V
Review Essay
Received: 02.08.2018. Accepted: 07.10.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Jelena Vasiljević

CONTENTIOUS POLITICS IN THE EUROPEAN (POST-SOCIALIST) (SEMI-)PERIPHERY: MAPPING REBELLION AND SOCIAL PROTESTS IN SOUTHEAST AND EASTERN EUROPE

Florian Bieber and Dario Brentin (eds.), Social Movements in the Balkans. Rebellion and Protest from Maribor to Taksim, Routledge London and New York, 2018. / Veronika Stoyanova, Ideology and Social Protests in Eastern Europe. Beyond the Transition's Liberal Consensus, Routledge London and New York, 2018.

ABSTRACT
This essay takes a critical and reflective look at two recently published books on contentious politics in the Balkans and Eastern Europe: *Social Movements in the Balkans* (ed. by F. Bieber and D. Brentin, Routledge 2018) and *Ideology and Social Protests in Eastern Europe* (V. Stoyanova, Routledge 2018). Focusing on regions somewhat neglected in scholarly analyses of the recent global upsurge of protests, these books aim to fill the gap by highlighting some contextual and regional specificities: a position of economic and geo-political (semi)periphery, weak or unconsolidated democratic institutions, post-socialist and transitional environments, societal (ethnic) divisions, etc. By critically assessing both contributions, in a manner that looks for their complementarity, this essay: examines the characteristics of popular mobilizations and grievances in Southeast and Eastern Europe; questions dominant narratives of political and economic transition and EU integration; re-evaluates socialist heritage and post-socialist political trajectories; discusses the (im) possibilities of articulating political alternatives to representative democracy and free market economy; and addresses the burden of conflicting memories and attitudes towards the region's socialist past (and, in case of post-Yugoslav states, ethnic conflicts from the 1990s).

KEYWORDS
social protests, crisis of representative democracy, radical democracy, Balkans, Southeast Europe, Eastern Europe.

Introduction

The global economic crisis of 2008, subsequent worldwide austerity policies, and the general neoliberal preference for technocratic governance have all shaken the trust of citizens in democratic institutions and representational models, giving rise

Jelena Vasiljević: Research Fellow, Institute for Philosophy and Social Theory, University of Belgrade; jvasiljevic@instifdt.bg.ac.rs.
This paper is a result of the project 43007 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

to a global cycle of protests movements, from *Indignados*, to Occupy movements, to experimentations with ad hoc citizens' popular assemblies and direct democratic decision-making (della Porta 2013, Kaldor and Selchow 2013, Mercea 2016, della Porta et al. 2017). Within Europe and the EU, these conditions have been coupled with a questioning of the transparency and accountability of EU governing structures as well as a critique of the stance of EU institutions towards indebted countries (on the EU periphery) and those aspiring to become member states (and already part of the EU accession process). However, scholarly interest in new forms of contentious politics and citizens' protests has mostly ignored the regions of Eastern and Southeast Europe, typically 'reserved' for the study of ethno-national conflict, transition successes and (mostly) failures, and contested memories of socialist and post-socialist experiences. Two notable exceptions, both of which aimed to offer a systematic overview of citizens' dynamic engagement in contesting the political-economic "order of the day" in the region, were a 2015 volume on radical politics in the Balkans by Horvat and Štiks and a 2017 special issue of *Europe-Asia Studies* on activist citizenship in Southeast Europe, edited by Fagan and Sircar.

This essay takes a critical and reflective look at two recently published books that are reversing this trend – *Social Movements in the Balkans: Rebellion and Protest from Maribor to Taksim*, edited by Florian Bieber and Dario Brentin (2018, Routledge) and a monograph by Veronika Stoyanova, *Ideology and Social Protest in Eastern Europe: Beyond the Transition's Liberal Consensus* (2018, Routledge) – effectively putting Southeast and Eastern Europe on the map for the study of contemporary social and protest movements. First, I will outline some of the themes and topoi that these contributions share and that, in fact, emerge as distinctive issues regarding popular mobilizations and protests in this part of Europe. Then, each book is discussed separately but is critically assessed with an eye for their complementarity.

Emerging Topics and Points of Contestation

Some problems and topics are saliently present and/or marked as distinctive in both volumes and appear across different country/protest case studies; but also seem to appear generally in contemporary social and protest movements, thus confirming their relevance beyond the region. First, there is a noticeable "return" of class and capitalism – following the end of belief in "the end of history" – as once again useful and legitimate tools of analysis, but also as matters of political contestation. Linked to this is another theme, an articulated critique of the general paradigm of post-1989 transitions to liberal democratic governance and free-market societies; and tightly connected to this is an emerging critique of EU integration as a process that has failed to deliver its promises. However, this cluster of criticisms and contestations are predominantly articulated as negation, or dissatisfaction, without accompanying visions about possible or desirable alternatives.

Also notable is that this general critique of dominant forms of politics – rather than of, say, the concrete political programs of parties in power – and consequent widespread, deep mistrust of all parties, has manifested itself in two different articulations among citizens. Most often, the protests and indignations of citizens fall

short of transforming into any kind of clear political program or goal; but sometimes, they inspire citizens to experiment, *in situ*, with direct forms of participation, in the form of *prefigurative* politics (as was the case in Greek public squares, or in Bosnian towns during so-called plenum-assemblies; see more in Milan 2017 and Zaimakis 2018).

Finally, both volumes reviewed here share insight about two competing major frames of interpretation among researchers of protests, and among protest protagonists themselves. One of these perspectives, which could be termed neo-Marxist or belonging to the academic Left, holds that the general conditions of neoliberalism and the dogma of free-market liberal democracy have led to non-democratic practices, the exclusion of citizens from decision making processes, and ultimately, to the reaction of citizens in forms of mass protest. The other perspective, shared predominantly by democratization scholars, claims that it is weak institutions and a lack of rule of law or accountability of political leaders that are to be blamed for a lack of democratic culture and practices.

In *Ideology and Social Protests in Eastern Europe*, Stoyanova firmly embraces the first position, arguing: "[T]he dominance of the transitological approach to the region should be seen as part of the wider post-Cold War Washington consensus which gave rise to an era of 'technocratic governance' that eschewed any popular grounding ... Such an elite consensus engendered a depoliticized technocratic culture of 'governance' which lacked accountability – politicians could not be held accountable for policies which were predetermined, even inevitable" (2018: 19, 20).

In *Social Movements in the Balkans,* on the other hand, Bieber and Brentin argue in their Introduction that these two frames of interpretation should not be seen as "irreconcilable alternatives," asserting that: "The critique of neoliberal policies helps explain how economic and political transformation has failed to deliver states that respond to citizens' needs and protect from predatory elites. Yet the emphasis on democratization and rule of law highlight why some liberal democratic market economies have been able to mitigate the economic crisis and respond to citizen demands ... [m]arket economy and representative democracy as such were not a problem ... Instead, the core problem remains weak institutions that are easy prey to the dominance of strong parties driven by narrow interests" (2018: 5). However, not every chapter in this edited volume expresses the same view.

Social Movements in the Balkans

Protests "from Slovenia to Turkey" are covered in Bieber and Brentin's volume. In other words, the entire Balkan region is examined, encompassing countries with both comparable and very different recent histories – some are already part of the EU, some are "potential candidate countries"; some are post-socialist, others are not. Still, despite any differences, common among these countries is the predicament of existing on the semi-periphery of Europe. Stretching their focus thusly, the editors subtitled the book: *from Maribor to Taksim*. These points of reference not only identify the geographic boundaries of the region under discussion, but also infer how versatile yet similar – in demands and sites of occurrence – all these protests were. Maribor is a city and Taksim a city square. Maribor is in the

small country of Slovenia, once an industrial heartland in socialist Yugoslavia, then drastically de-industrialized; while Taksim is at the heart of a true cosmopolis (Istanbul). But both are sites of tremendous change, swift developments, and political processes that have radically challenged their residents to conform, adapt, or resist imposed transformations.

The chapters of *Social Movements in the Balkans* address the diverse issues that incited the protests in question: austerity measures, the privatization and commodification of common goods, corruption, authoritarian political actions, etc. Introducing the volume, Bieber and Brentin invoke Garret Hardin's seminal essay, "The tragedy of the commons," implying that this is the one thing all recent protests in Southeast Europe do share. They describe "a sense of grievance with the way the authorities administer the common good, public spaces and the state" (p. 1).

In the opening chapter, Heiko Wimmen considers informal political movements in Bosnia-Herzegovina from 2005 to 2013, backed by international organizations, to strengthen civil society and encourage bottom-up citizen participation. Wimmen asks to what extent these efforts have compensated for the shortcomings of a political system that sustains and nurtures ethno-national politics and divides. Analyzing several attempts by civil society organizations and other informal actors (who were distanced from and dissatisfied with formal NGOs) to influence both political parties and citizens (to exercise pressure on parties and overcome political abstinence), the paper concludes that these "occurred *parallel to* rather than *cutting across* the dominant ethnic cleavage lines" (p. 13, original emphasis). The conclusion is both pessimistic and optimistic: on the one hand, civic resistance has not brought about major changes, nor has it succeeded in bridging ethnic divisions by promoting inclusive Bosnian civic citizenship; on the other hand, citizens involved in resistance efforts have built capacities, broadened networks of support, and accumulated experiences. It is also important to measure the failures (and successes) of civil society against the wider political system in which they operate, and in the case of Bosnia-Herzegovina, this system is designed to maintain division and exploit fears of instability.

In the subsequent chapter, Gal Kirn discusses the example of Maribor – once one of the most important industrial cities in Yugoslavia and today almost completely de-industrialized – as typical of the periphery. There, wintertime protests in 2012-2013 raised questions about the "myth of Slovenia as a 'success story' of the Balkans." The protests erupted in reaction to corrupt deals made by Maribor's then-Mayor, who was behind a shady public-private partnership responsible for a new road radar system that resulted in more than 2,000 speeding tickets in just two weeks (in a town of 100,000). The chapter argues that even though protests resulted in a positive outcome, as the Mayor was ousted, they oscillated between a moral and almost apolitical outcry (demanding accountability) and a more complex and rigorous critique of the very foundations of liberal democracy – thus sharing this "in-betweenness" with other recent protest movements on the periphery. Kirn adopts a Rancièrian framework of understanding "the people," the dispossessed, as a political figure of dissensus, but he argues further that the context of de-industrialization and the dismantling of welfare services and infrastructure makes the *surplus population* equally significant. Kirn is interested in the possibilities for

this population to become politicized; this is why, he concludes, the most important outcome of the Maribor protests was not the resignation of the Mayor, but the emergence of new political actors and their coming together in a political coalition (the United Left), which entered Parliament in 2014.

Valentina Guerguieva authored the next chapter, on a series of protests in Bulgaria, by presenting personal accounts of the activists themselves, collected through in-depth interviews and participant observation. A special focus is given to the environmental movement that rose in 2012 to oppose the use of fracking to extract shale gas. But Guerguieva also examined two protests in 2013 – in the winter, against a sudden rise in electricity costs; and in the summer, over the election of a media mogul, suspected of having criminal ties, to the position of Director of the State Agency for National Security. She concludes that these protests mostly took on the form and logic of negation, as they were "reactions of indignation, expressions of disapproval" that mostly dissolved after protest actions ended. Nonetheless, Guerguieva argues, these "reactive mobilizations" can in fact have a positive impact and could be seen as a *counter-democracy* (she cites the work of Pierre Rosanvallon), and that "the positive work of mistrust" is a tool for maintaining pressure on the government. This argument resembles the traditional understanding of the role of civil society, and interestingly enough, it is precisely this notion of civil society and its imaginaries in Bulgarian protests that form the object of the study of Stoyanova's book. In that volume, however, as will be shown shortly, the potential for citizens' protests to act as checks and balances is viewed with more nuance and less optimism.

In the next chapter, on protests in Greece, Kostis Plevris argues that physical and geographical *spaces* of protests are much more than mere containers and backgrounds to social activities. This chapter differs somewhat from the others by taking a more theoretical approach, and is a bit detached from the actors, actions, and demands that dominate other accounts in the volume; but even so, Plevris points to some shared problems identified in the Introduction, noting that in contemporary struggles for the commons, spatial claims are important and "only a particular space enables social struggle to flourish." This seems very true in times when important means of struggle include occupations of workplaces, factories, universities, or cultural venues (like theaters). Equally, social *spatial* organization, like communal work in neighborhoods, is becoming increasingly important as traditional institutional networks of support and welfare infrastructures rapidly wither away.

Ksenija Berk then presents another look at the Slovene protests, but in this chapter the focus is on "the visual communications" and the creative, artistic interventions of protestors – a much written about topic in studies on recent waves of protests. She skillfully juxtaposes simple, effective, yet smart and mobilizing visuals (often using irony and parody), such as posters and stencils, to more sophisticated artistic interventions that fail to contribute to the general cause of the protests and serve instead as "vehicles for self-promotion" and "example of aesthetic populism in protest graphics" (p. 89, 90).

The chapter by Željka Lekić-Subašić confronts the use and influence of social media, another prominent object of study in relation to contemporary protest movements. This chapter focuses on online activism and the role of social media in Southeast Europe specifically, with a case study on Bosnia-Herzegovina. It aims to

address questions frequently posed about the prominent role of social media today: What is the transformative political potential of social media? Can social media help mobilize and organize? And, does social media positively affect democracy or actually prevent "real" activism by offering a comfortable substitute and the "feel-good" effect of "doing something" by clicking? Her conclusion, which moderates both an overly optimistic embrace of new technologies and widespread fears related to their use in fighting oppressive regimes, is that social media is primarily a very helpful tool for disseminating information and organizational details but cannot substitute for "traditional" forms of organization. An important caveat, however, is that this research is only illustrative, as it was conducted using very limited material, based on a survey of active students in Sarajevo and secondary resources related to protests in Slovenia.

Chiara Milan and Leonidas Oikonomakis offer a very insightful comparative analysis in the next chapter, examining the characteristics of protests in Greece, Turkey, and Bosnia-Herzegovina. In all three countries, protests that were analyzed started in order to challenge a specific single issue – austerity measures in Greece, the demolition of Gezi Park in Istanbul, and the inability of the Bosnian government to issue ID numbers for newborns. However, in Greece and Turkey, these protests evolved into a more sophisticated critique of general governing structures and even of representational democracy, establishing grounds for experimenting with autonomous forms of self-organization and direct democracy and thus becoming part of the global "movement family" of protest cycles since 2011. This was a testimony to the realization of citizens in Greece and Turkey that those singular issues which sparked initial outrage were merely the symptoms of a more generalized problem: the exclusion of citizens from affecting political decision making that directly impacts them. But this was not the case in Bosnia-Herzegovina and, using concepts from social movements studies, like *frame resonance*, *network*, and *resources*, Milan and Oikonomakis argue that the movement there was hindered from further development and politicization because it lacked two important elements: "the resonance of the overarching frame ... and strong local and transnational movement networks" (p. 115). The inability of the movement in Bosnia-Herzegovina to question the structural context generating contested issues was decisive; and this is intricately related to the post-Dayton political system in which the questioning of ethno-political elites has easily translated into a questioning of post-war stability. Still, the analysis of Milan and Oikonomakis does not fully explain the outburst of protests in Bosnia-Herzegovina that followed only a year later, culminating in a series of self-organized participatory assemblies, or plenums. This fact is only very briefly mentioned in the chapter, noting that it is "at odds" with the analysis of the 2013 protest (p. 125).

In the chapter that follows, Marius I. Tatar puts forth a very useful overview and mapping of the statistical data on protest politics in Southeast Europe. Using the European Values Survey and World Values Survey datasets, and then taking a closer look at the cases of Slovenia and Romania, the chapter takes a longitudinal approach to presenting key characteristics of both protesters and the contexts of their protest, from a dynamic and comparative perspective. The figures and tables it includes may prove useful for other research and may inspire other

cross-comparative analysis of available data. All in all, in mapping some specificities of the region, Tatar reveals that protesters who have sparked recent waves of protest in Southeast Europe tend to have similar socio-demographical and attitudinal profiles to those in Western Europe – meaning, those most likely to protest have higher levels of education and higher incomes, and nurture post-materialist values – but also that they "have stronger ideological identifications on the Left-Right axis" (p. 147).

Mark Kramer authors the closing chapter, exploring the somewhat controversial but significant question of how "anti-regime protesters can be influenced by external actors" (identified as foreign governments, foreign media outlets, international organizations, transnational movements, diaspora, spillover effects from neighboring countries, etc.). His analysis is particularly focused on protests in Bulgaria, Romania, and Turkey in 2013, but also deals more generally with the issue of influence by foreign actors. Especially important and nuanced is his discussion of the role of transnational advocacy networks (TANs) and "external 'norm entrepreneurs' (individuals, groups, organizations or governments that promote normative concepts such as human rights, democracy and political freedom)" (p. 163), in correlation with domestic political opportunity structures, where Kramer carefully weighs their mutual influence and co-dependence. Also very useful is Kramer's systematic elaboration of the specific ways in which protests and political upheavals from other countries can exercise influence: through diffusion, demonstration effects, or spillovers. However, the empirical research supporting important insights in the chapter is lacking, and therefore, it serves mostly as a useful theoretical and methodological consideration for future research.

Ideology and Social Protests in Eastern Europe

While *Social Movements in the Balkans* brings an interesting variety of methodological and theoretical approaches and provides an overview of a very wide region, it falls short of providing a comprehensive, systemic framework for interpreting and understanding new protest cultures in the region. *Ideology and Social Protests in Eastern Europe* endeavors to offer precisely that, although by focusing on a much narrower object of study.

The book is a very ambitious work in which Veronika Stoyanova introduces an important and fresh perspective in understanding recent protest movements. Firstly, it focuses on Eastern Europe, a neglected area when it comes to the subject matter. Secondly, it brings an original approach by studying antagonisms, not between those who protest and those who are protested, but *within* popular protest mobilizations and among different factions of protesters themselves. By focusing on two waves of Bulgarian protests in 2013, viewed as different in many ways and having mobilized distinct social groups, Stoyanova distills the strife between them into a question that goes straight to the heart of all of today's popular mobilizations for social change: "whose vision for change constitute[s] the right way forward?" (p. 4).

Stoyanova takes a very complex and sophisticated approach that is at times slightly confusing due to the great (and sometimes unnecessary) multiplicity of perspectives she presents in relation to the main object of the study, which is:

understanding post-socialist power relations through discourses on civil society, as they were produced and utilized during the two waves of protests in Bulgaria in 2013 – known as the Winter and Summer protests. This is a very original approach to recent protest movements, harnessing the notion of *civil society* and the peculiarities of its meanings and usages in post-socialist contexts.

The theoretical grounding of Stoyanova's volume does not rely so much on social movement literature (in contrast to most chapters in the other volume reviewed here), as she finds it unsuitable for explaining post-socialist dynamics because it links social movements to reactions to strong states and markets. She firmly embraces the position of the academic Left, but distances herself from the usual post-structuralist and post-Marxist approaches of the New Left, as she wishes to reanimate the notions of *human agency* and the power of the powerless. Thus, she borrows theoretical framing mostly from Gramsci and Ernst Bloch. It is Gramsci's understanding of civil society she uses in asking how different factions of protestors claim to represent the authentic voice of civil society ("how the idea of civil society itself gets mobilized within these struggles," p. 7). Not surprisingly, tightly connected with this is another Gramscian theme of (organic) *intellectuals* and their role in (de)legitimizing political actions, in developing hegemonic discourses, and in presenting specific ideologies in universal terms. She further combines Gramsci with Bloch's ideas about the functionality of utopian thinking: "The same way Gramsci believed that in the mystifying fog of the subalterns' 'common sense' there resided a kernel of 'good sense', Ernst Bloch argued that there always existed a kernel of *utopian surplus* at the core of every ideology" (p. 38, original emphasis).

Arguably, the strongest points of the book reside in Stoyanova's critical and theoretically insightful scrutiny of how notions of civil society and transition have been used, in both academia and politics, to describe the post-socialist condition. She argues that the concept of civil society has been underpinned with ideas of Westernization and "catching-up," and was thus imbued with ideological and utopian content. From 1988 onwards, she further argues that for East European societies, this was "propelled by idealized visions of the future and of notions such as democracy, Europeanization and morality" (p. 88).

> "[T]his discourse conceived civil society as comprising the 'energetic minority' of citizen-protesters, as opposed to the 'passive majority' of 'docile subjects' ... Secondly, civil society was opposed to the 'state' or to 'political society' (in a Hegelian manner), which served to depoliticize the concept and firmly ground it in an anti-statist (and anti-communist) ideology. These two currents ... corresponded to the two main tasks of the early transformational utopia ... first, that of a project of modernization ... and second, the project of 'decommunization' or that of the dismantling of the (repressive) state ... Key to not just the manufacture of consent, but to the mobilization of appeal to the neoliberal ideology ... was the emotional appeal to autonomy and self-determination inherent to the idea of civil society" (pp. 88, 89).

This critique is then connected with critique of the dominant, liberal, anti-populist interpretations of popular mobilizations in Eastern Europe and the fear of overcoming the post-1989 liberal consensus. She rightly notes that democratization scholars, who embrace a strong critique of populist tendencies – and who worry more about the decline of trust in liberal institutions than about the decline, or rather exclusion, of citizens in political participation – conflate *elite* forms of

populism (represented in populist political parties) and *popular* forms of populist mobilization, imputing to the latter the illiberal tendencies of the former (pp. 20, 21).

What has come to be known in Bulgaria as the Winter Protest took place in February 2013 and started in reaction to an unexpected and unprecedented rise in electricity costs. Citizens in many Bulgarian cities protested, and even expressed demands for the nationalization of energy distribution companies, the privatizations of which were seen as the core of the problem. As a result, the government resigned, and after elections that May, a new coalition government took office. But just over a month into its mandate, the new government made the controversial decision to appoint a media mogul, known for his ties to the mafia, to head the State Agency for National Security. Once again, tens of thousands of Bulgarian citizens poured into the streets to protest, and this time the corrupt government and its links to the mafia were their focus – resulting in the Summer Protest.

What Stoyanova set out to study were the antagonistic discourses of these two protest waves. The crux of her argument is:

> "[T]he Winter protest wave made a counter-hegemonic attempt to challenge the liberal-capitalist discourse of the transition's organic intellectual elite ... by articulating a popular-national political identity. The latter's organic expert-intellectuals then responded first with intense criticism and mockery of the Winter protest and a refusal to 'join in' the latter's protest, despite the common 'enemy without' (corrupt power-holders); and later by attempting to re-assert the hegemony of their liberal-capitalist vision during the Summer protest wave, which they sought to dominate" (p. 99).

Stoyanova rightly detects two prevailing, mutually antagonistic discourses – both claiming they are about civil society, as they articulate different demands regarding the power relations between state, market, and citizens – but to my mind, she distinguishes them in perhaps a too 'ideal-type' manner. She contends that the Winter Protest attempted to de-monopolize liberal understandings of key notions, primarily of civil society, and in fact developed a counter-hegemonic discourse of what she termed *civil society of the people* (a Gramscian understating). Even though this discourse was fragmented, there allegedly "lurks a *radical conception of democracy* as the horizon of social critique, epitomized in the utopian desire to *democratize civil society*" (126, original emphasis). In contrast, Stoyanova found the Summer Protest discourse was much more homogenous and rounded, and that it defended a "civil society of the middle class" (132–172). This was framed as "the protest of the moral, productive (and creative), tax- and bills-paying, and even beautiful middle class, which has finally 'risen' to do away with the communist remainders and finish the 'incomplete transition' to European 'normality'" (p. 132).

Some Concluding Thoughts

Though it is convincingly shown by Stoyanova how differing discursive strategies aimed at producing cultural-political distinctions, I believe the question is still open as to whether this testifies about a true inter-class clash and whether we can really talk about different classes representing the two Bulgarian protests in 2013 (the Winter Protest allegedly belonging to the subaltern class, and the Summer to the middle class). It seems to me that Stoyanova is perhaps conflating discursive

tropes with socio-economic strata. After all, what is analyzed in the book are *discourses*; which represent the only empirical material and are not accompanied, for example, by other socio-demographic data about protesters. So, it might as well be the case that these two waves of protests articulated two conflicting discourses – subaltern or middle-class interests – but this fact does not mean that the classes themselves were originators and distributers of these distinct discourses. While Stoyanova quite insightfully and rightfully problematizes the presumed coherence of claims and political stances of those fighting *against* the regime, she ironically assumes a simplified and reductionist approach to what she sees as factions within protesting blocs. Her argument thus tends to homogenize the groups fighting to dominate the discourses of dissent. What I want to underline is that distinct discourses do not necessarily overlap with distinct classes; in fact, this is precisely where the power of discourses resides – in their ability to assert autonomy from concrete class or group interests.

Given the title of the book as well as its Introduction, both of which promise to offer perspective on Eastern Europe, another weak point of Stoyanova's book is the betrayal of this promise, beyond the case study of Bulgaria. Despite announcing that she will take a comparative look at Hungary, Romania, and Macedonia, this is almost completely lacking. Instead, just several paragraphs in the closing chapter serve to legitimize the title of the book but offer no meaningful comparative insight. Consequently, brief descriptions of protests in those countries are not only superficial, but necessarily distorted as well, as they are presented in such a way as to support the prevalent analysis of the Bulgarian case (with an exclusive emphasis on discursive class- and culture-based derogation of 'subalterns' by 'middle class' proponents of the protests).

Still, Stoyanova is absolutely right to claim that "the critical studies of political struggle can benefit tremendously from a stronger East European (post-socialist) perspective, which offers empirical data on the workings of ideological struggle in the aftermath of large-scale social change" (p. 6) and I found it thought provoking to consider how this framework of analysis could be used in post-Yugoslav countries, as they also share the "predicament" of belonging to the post-socialist world. Bieber and Brentin also rightly note that the elaboration of the post-Yugoslav space, as a specific region where grievances and protests can easily have 'spill-over effect', is often omitted from analyses of emerging protests movements. Yet, even in their volume, a comparative or a systematic approach to this region is missing.

Some thoughts that spring to mind when imagining research on theories of ideological discourses in protest movements in the post-Yugoslav space include: First, it would probably be difficult to detect distinct protests articulating distinct discourses (liberal or more radical), as seems to be the case in Bulgaria. Rather, it would be more likely to find different discourses overlapping within the same events or protests. Somewhat connected to this is an assumption that the generational gap could play an important role, as it could be argued that, for instance, the discourse about liberal values is more likely to be found among older generation, whereas younger people are prone to more radical ideas. Perhaps most importantly, though, anti-communist discourse – which Stoyanova sees as central in liberal 'agenda setting' – is very different in post-Yugoslav states, due to the radically

different nature of their socialist periods and, consequently, their radically different cultures of socialist memories. It has already been noted, for instance, that phenomena like *Yugonostalgia* or *Titostalgia* (Petrović 2016, Velikonja 2010) are not homologues to nostalgia for socialism in other post-socialist countries. Finally, the fact that post-Yugoslav states are also post-partition and post-conflict societies adds another layer of meaning to notions of civil society (overlapping with the sphere of anti-war activism during the 1990s) and transition (not only from socialism to capitalism, but also from war and conflict to peace and stability).

References:

Bieber, F. and D. Brentin (eds.) (2018), *Social Movements in the Balkans. Rebellion and Protest from Maribor to Taksim*. London and New York: Routledge.

della Porta, D. (2013), *Can Democracy be Saved? Participation, Deliberation and Social Movements*. Cambridge: Polity Press.

della Porta, D., M. Andretta, T. Fernandes, F. O'Connor, E. Romanos and M. Vogiatzoglou (eds.) (2017), *Late Neoliberalism and its Discontents in the Economic Crisis. Comparing Social Movements in the European Periphery*. Palgrave Macmillan.

Fagan, A. and I. Sircar (2017), "Activist Citizenship in Southeast Europe", *Europe-Asia Studies* 69 (9): 1337–1345.

Horvat, S. and I. Štiks (eds.) (2015), *Welcome to the Desert of Post-socialism. Radical Politics after Yugoslavia*. London and New York: Verso.

Kaldor, M. and S. Selchow (2013), "The 'Bubbling Up' of Subterranean Politics in Europe", *Journal of Civil Society* 9 (1): 78–99.

Mercea, D. (2016), *Civic Participation in Contentious Politics: The Digital Foreshadowing of Protest*. Basingstoke: Palgrave.

Milan, C. (2017), "Reshaping Citizenship through Collective Action: Performative and Prefigurative Practices in the 2013–2014 Cycle of Contention in Bosnia & Hercegovina", *Europe-Asia Studies* 69 (9): 1346–1361.

Petrović, T. (2016), "Towards an Affective History of Yugoslavia", *Filozofija i društvo* 27 (3): 504–520.

Stoyanova, V. (2018), *Ideology and Social Protest in Eastern Europe. Beyond the Transition's Liberal Consensus*. London and New York: Routledge.

Velikonja, M. (2010), *Titostalgija*. Beograd: Biblioteka XX vek.

Zaimakis Y. (2018), "Autonomy, Degrowth and Prefigurative Politics. Voices of Solidarity Economy Activists amid Economic Crisis in Greece", *Partecipazione e conflitto* 11 (1): 95–120.

Jelena Vasiljević

Politike osporavanja na evropskoj (postsocijalističkoj) (polu)periferiji: mapiranje pobune i društvenih protesta u Jugoistočnoj i Istočnoj Evropi

Apstrakt

Ovaj rad se kritički i polemički osvrće na dve nedavno objavljene knjige o politikama osporavanja na Balkanu i u Istočnoj Evropi: *Social Movements in the Balkans* (prir. F. Bieber i D. Brentint, Routledge 2018) i *Ideology and Social Protests in Eastern Europe* (V. Stoyanova, Routledge 2018). Fokusom na regione koji su donekle zapostavljeni u akademskim analizama nedavnih globalnih talasa protesta, ove dve knjige nameravaju da popune tu prazninu, te osvetle neke od kontekstualnih i regionalnih specifičnosti: položaj na ekonomskoj i geopolitičkoj (polu)periferiji, slabe i nekonsolidovane demokratske institucije, postsocijalistički i tranzicioni uslovi, društvene (etničke) podele itd. Kritički procenjujući obe studije i sagledavajući ih u njihovoj komplementarnosti ovaj rad se dotiče sledećih problema: karakterisike masovnih mobilizacija i protesta u Jugoistočnoj i Istočnoj Evropi; propitivanje vladajućih narativa o političkoj i ekonomskoj tranziciji i evropskim integracijama; ponovno vrednovanje socijalističkog nasleđa i postsocijalističkih političkih putanja; (ne)mogućnost artikulisanja političkih alternativa predstavničkoj demokratiji i tržišnoj ekonomiji; teret sukobljenih sećanja i stavova prema socijalističkoj prošlosti.

Ključne reči: društveni protesti, kriza predstavničke demokratije, radikalna demokratija, Balkan, Jugoistočna Evropa, Istočna Evropa.

V

REVIEWS

PRIKAZI

LUIGI CARANTI, *KANT'S POLITICAL LEGACY: HUMAN RIGHTS, PEACE, PROGRESS*, CARDIFF, UNIVERSITY OF WALES PRESS, 2017.

Natalia Lerussi

Although it is true, as is stated in the introduction to the book, that Kant's political philosophy has served as the inspiration for the foundation of important political devices in the course of the previous century, essentially after the Second World War (human rights, the United Nations, the continental blocs, etc., the development of a humanitarian language in general), it is also true that the reception of Kant was the product of two clearly isolated agencies. These are, on the one hand, the academic world, concerned centrally with hermeneutic questions, often innocuous from the practical point of view, and, on the other, the public in general, which, in view of the application of Kant's thought, has frequently used his ideas with little theoretical exactitude and, for this very reason, failed to take advantage of the conceptual riches which would be offered by an attentive and careful reading of its sources. From this point of view, the book is presented as a point of communication between these two hitherto separate spheres of reception. Thus, its objective is not only to offer a coherent vision of Kant's political philosophy on the basis of a detailed analysis of the sources, but, in addition, to elucidate the theoretical tools which are to be found in his thought with which we can confront at the global level the great political challenges of our century. The book is a resource for the specialists alone but also for political agents who aim to take intelligent action in the search for peace and the defence of human rights. It is a hermeneutically careful interpretation of Kant's political philosophy which is, at the same time, sharp and committed to the problems of our times.

The book is divided into three Parts (I. Human Rights, II. Peace and III. Progress), which are divided, in turn, into three Chapters each. The different theoretical frameworks used, the discussions into which the book intervenes, and the hypothesis developed in each Part are stated clearly in the Introduction. This thus offers a rapid and useful overview of the most important contents of the book.

In Part I, "Human Rights", the problematic of this type of "right" is analysed, the most important questions are determined, and the "justification problem" is described, together with the different responses to this question which we have at present. In this context, Caranti makes Kant intervene in the discussion through the reconstruction of the elements which are fruitful for a "strong" justification of human rights, without for that reason falling into dogmatisms or a "simple-minded humanism". In relation to the problems

associated with a certain "occidentalism" into which the defence of human rights could fall, the author shows that Kant's political philosophy provides us with a sane and pertinent combination – which enables us to confront those problems – of the universalism of the principles with a particularism in the application of those principles. Towards the end of this first Part, the author develops an argument about human rights which, although inspired by Kant, distances itself from him on several important points.

According to the author's argument, although Kant did not and could not know the technical term "human rights", it is possible to offer a reconstruction of a justification of this kind of right on the basis of an analysis of the grounding which the philosopher gives to the inborn "right to external freedom". Kant would have defended a conception of human rights (that is, the inborn "right to external freedom") through a substantive conception of "human dignity", based, not merely on practical liberty (as a faculty of choosing between alternative actions), but rather on the concept of autonomy (as "the ability to act under self-imposed moral constraints", p. 57). According to Caranti, autonomy appears as a privileged basis of the concept of human dignity, as a foundation which serves both common sense and "all mayor cultural traditions and revealed religions" (p. 57). In the face of other solutions to the "justification problem" which the author presents and discusses in detail, the Kantian proposal is offered as the most solid alternative. However, with the end of making the Kantian justification of the concept of "human rights" (that is of the "right to external freedom") even more inclusive and modern, the author proposes two important modifications. In the first place, he argues that we must accept (against Kant) that not only human beings, but also the animals, act morally, with a certain degree of autonomy and are, because of this, worthy of respect and the subjects of rights. In the second place, distancing himself again from the literal meaning of Kant, who, as is known, bases the concept of autonomy on the "categorical imperative", the author proposes underpinning the concept of autonomy with any general principle which could function as an original grounding of moral decision. That is, it could be supported not only by the categorial imperative, but also by some principle inspired, for example, by the Aristotelian concept of the ethical life or even the "Golden Rule" (one of whose formulations would read, "Do unto others as you would wish them to do unto you", p. 64), etc.

With respect to Caranti's second proposal, I would like to make one observation. Kant bases his concept of autonomy on the respect for the categorical imperative as the consequence of strong internal reasons. In particular, through the strictly formal character of this principle Kant guarantees that it belongs exclusively to reason and thus guarantees that no external content or end interferes with moral determination. Rational beings are autonomous when they act out of respect for the categorical imperative because they do so in relation to a principle which has its source in pure practical reason. Now, if autonomy is guaranteed, as Caranti wants it to be, through the determination to act in conformity with any general principle, we are confronted with an undesirable alternative: if the principle in question is not given to the agent by reason, but rather they must decide it voluntarily, the question arises of the *lack of a criterion* for deciding between general moral principles. The consequence is that the decision to subordinate oneself to one general principle or another appears arbitrary. If, on the other hand, the principle in question is given to the agent, not by means of their reason, but rather, for example, through culture, the state or religion, it becomes difficult to maintain an acceptable concept of "autonomy". This second alternative also brings the difficulty that it does not permit us to establish a sharp division between moral determination mediated by, for example, the Golden Rule and by a morally perverse principle, such as the determination to act in accordance with the will of the Führer in Nazi Germany. Although the author's proposal

makes Kantian morals more flexible and more suitable to our contemporary vision of the moral life – this is unquestionably a major effort – this point should be qualified. It is not a question of satisfying specifically Kantian requirements, but rather those of anyone who seeks to defend a consistent concept of "moral autonomy".

Part II, "Peace", analyses the "Democratic Peace Theory" (DPT). Initially developed by the political philosopher Michael Doyle in "Kant, Liberal Legacies and Public Affairs" during the nineteen eighties, this has had a far-reaching influence in the present. One of the central theses of this theory which, according to Doyle, would have been announced by Kant long before it could be proved empirically, is the directly proportion relationship between democracy and peace. Although there are different variants of the DPT, which Caranti describes and discusses, in general, the central point defended within this framework is that, for 200 years, that is to say, since their very birth, the liberal democracies (the "republics", in Kantian terminology) have not entered into nor have had to enter into war with each other. The objective of this section of the book is to weigh up the problems and defects of the DPT through an analysis of the tract, *To Perpetual Peace* (1795), and of the different critiques which have been made of this theory from a conceptual and historical point of view. (In fact, as the author recognizes, the DPT is challenged by multiple and important counterexamples from the nineteenth century to the present.) The three permanent conditions of peace (republicanism, federalism and cosmopolitan right) outlined by Kant in the tract of 1795, read carefully, offer, according to the author, a valuable resource for a more solid DPT. In Caranti's reading, it is not just a question of the fact that the three conditions must function *jointly* so as to guarantee peace, as Doyle argues, but also of *understanding them correctly*. Thus, the author offers an interpretation of those conditions which he contrasts with that offered by the defenders of the DPT. Unlike the reading which the latter defend, Caranti offers one which contains the following points. First, the republic must be distinguished from "liberal democracy": the latter watches over the particular interests of the individuals, while in the former the equal status of the citizens is protected. Second, the federation of the nations must be considered a league not only of republican but also non-republican states. Finally, foreigners' right to visit must be interpreted not merely as a "right to trade", but also, in general terms, as a right to community or to society, or, inversely, as the obligation on the part of the states to guarantee the human rights of foreigners, a guarantee which would lead to the development of a global moral consciousness. Moreover, the author argues that the "right to trade" which would promote peace does not imply, according to Kant, the right to all forms of trade, but only to those forms of exchange which are fair for all the parties involved.

In this section the author shows how, understood properly, Kant's political philosophy presents itself as a valuable resource for correcting a theory which has had a major impact in our times and, at the same time, for defining those fundamental guidelines which politicians ought to follow in order to promote peace. Particularly worthy of attention is the charge that Caranti makes at the end of this Part, according to which the DPT, even in the version refined by Doyle, would oscillate in a way which is "politically dangerous". For Kant, as the author correctly argues, republicanism, federalism and the right to visit are *norms of action* which guarantee or supposedly lead to peace, but the DPT *moves* between a descriptive or explanatory theory about how relations are and have been between democratic (liberal) states, on the one hand, and a theory which *serves to justify certain wars*, that is, those which involve a supposedly undemocratic state, on the other. (The DPT functioned in this way in George W. Bush's speeches during the war against Iraq but is also used in this way by Doyle himself [p.198], although with a different meaning from that of the former.)

Finally, Part III, "Progress", analyses the Kantian conception of history presented both in the essay *Idea for a Universal History with a Cosmopolitan Purpose* (1784) and in *To Perpetual Peace* (1795). It aims to establish and make credible the theoretical reasons for which, according to Kant, both greater respect for human rights and the establishment of peace can be expected in a nearer or more distant future.

On the basis of the tract of 1795, Caranti argues that the Kantian conception of historical progress could be vindicated if two conditions are observed. In the first place, the mechanism of the "unsocial sociability" has to be separated from what is for contemporary standards, according to the author, the highly problematic notion of "natural dispositions" (this is, Caranti's "Separability Thesis") and, in the second place, the conception of historical progress has to be studied in conjunction with the general premises of what, for Kant, human nature is. The prediction of human development towards progress is possible taking uniquely into consideration "certain constant feature[s] of human beings" (p. 210) such as, for example, "unsocial sociability" (p. 215), the "limited benevolence and ability to learn" (p. 216), or the "pursuit of happiness or self-love" (p. 233), and certain objective circumstances of the external world. The result is to offer a reconstruction of the argument for which, according to Kant, there are reasons – highly plausible according to present-day canons - to believe that a non-linear progress towards a cosmopolitan constitution (and therefore not its alternatives: regression or stagnation) is the most probable course of development in human affairs. Because of this, the author distances himself from the usual interpretations, according to which historical "progress" must be understood in a merely regulatory or even "practical" form. In line with this, the author analyses the First Supplement to the text of *To Perpetual Peace* (1795), in which Kant presents his much discussed "Guarantee Thesis" according to which "nature ensures that humans will one day achieve a condition of perpetual peace" (p. 218). Caranti offers an interpretation of this affirmation which, by contrast with other readings, makes it not only compatible with Kant's critical philosophy, but also epistemologically significant. In that Supplement Kant would not defend a teleology based on "nature or providence ends", but rather a theory about the existing mechanisms by means of which progress towards the cosmopolitan condition can be forecast. In this way, Caranti reads the texts of 1784 and 1795 as being in harmony with each other.

Here, I would like to make an observation about the strategy of eliminating teleology from Kant's conception of history. According to the point of view of the model of the development of nature which appears to have won the battle (Darwinism or some version of Darwinism), all reference to natural "ends" or "dispositions" (or "ends of providence") must remain outside the realm of scientific discourse. In this context Caranti attempts to cleanse Kant's history of philosophy of its teleological assumptions. He thus shows that his conception of historical progress is perfectly sustainable without those assumptions. Thus, the natural mechanism would be a concept sufficient for explaining human development even from a point of view which would have to be admitted by Kant himself. But we must note that, as the author demonstrates, according to Kant, the horizon of the natural mechanism would coincide with the *highest ends of reason*, the institution of a perfect civil constitution (a republic) in both its internal and external relations, perpetual peace, etc. Is this coincidence not just a little too happy? Should not this coincidence itself be the object of explanation? Is not this explanation, precisely, something which Kant's teleological comprehension of history could offer us? However it may be, I think that underlying the author's theoretical perspective is a tendency to consider all forms of teleological thought to be "dogmatic", when in fact one of the great achievements of Kant's critical philosophy is that of offering a "critical grounding" for that kind of reasoning. Perhaps not only some kind of Darwinism, but also social

and political theory could take advantage of a critically grounded teleological comprehension.

In the final Chapter of the last Part, history is approached from the point of view of the political agent and their interest in articulating their efforts towards progress, that is to say, in this context, towards perpetual peace. In particular, that section studies the nature of the prudent subordination of the political to the moral which is expected of the *moral politician,* a figure which, in Caranti's analysis condenses two fundamentals. On the one hand, it contains the thesis of the subordination of the political to the moral, and on the other, it accommodates the thesis of the primacy of politics, for the art of prudence which the *moral politician* exercises exceeds the moral and defines the sphere of autonomy proper to politics. Unlike the First Supplement of *To Perpetual Peace* in which, according to what we saw above in Caranti's interpretation, it is the "mechanism of nature" which leads to and guarantees peace, in the Appendix to the same text this role is exercised by the "the moral politician's good will and his free decision inspired by duty" (p. 242). This does not, however, make us fall into a contradiction. To the contrary, in Caranti's reading, we are confronted by two complementary theses: the moral politician does not replace the labour of the natural mechanism, but rather accelerates or promotes that which nature does independently of the will of human beings. The *moral politician* (unlike the *political moralist* and, above all, the *moralizing politicians*, figures Caranti also analyses carefully), therefore, defines the responsibility which human beings, in general, and politics in particular, have in the face of the urgency of providing responses to the problems and challenges of our times.

Kant's Political Legacy: Human Rights, Peace, Progress is a book which is accessible to the general public, although it is profound and enriching in all its sections. It is structured well, contains intelligent argumentative strategies, includes an up to date literature and participates in a series of fundamental current debates. It offers important theses both for the field of Kant studies and for politicians, functionaries, and activists concerned with giving a solid theoretical grounding to the defence of human rights, the search for peace and the hope for progress. It opens a road to the renovation of the reception of "Kant's political legacy" which I hope will increase still more and diversify both within the academic world and outside of it.

ALESSANDRO FERRARA, *ROUSSEAU AND CRITICAL THEORY*, LEIDEN, BRILL, 2017.

Marjan Ivković

Alessandro Ferrara's *Rousseau and Critical Theory* can be considered a form of reconstruction with a systematic intent. Ferrara succinctly reconstructs the most important tenets of Rousseau's philosophy (Part 1), interrogates his reception within Critical Theory's three generations – the original Frankfurt School, Habermas and Honneth, followed by authors outside the "core" of the tradition such as Charles Taylor and Frederick Neuhouser (Part 2), demonstrating in the final part the enduring relevance of some of Rousseau's most valuable insights, not just for present-day Critical Theory, but for political philosophy and theories of the self in general. Engaging Ferrara's study, however, one realizes that the aims and contents overflow the boundaries of the above genre as the contours of an original "Rousseauian" Critical Theory gradually begin to take shape – let us first take a glance at the reconstructive edifice that supports Ferrara's critical-theoretic perspective.

Ferrara's reconstruction of Rousseau in this work draws heavily on his *Modernity and Authenticity*, but also on *Reflective Authenticity*, *The Force of the Example* and the more recent *The Democratic Horizon.* In a manner not dissimilar to Charles Taylor's reading of Rousseau (which figures in Section 2 of the study), Ferrara views the Enlightenment author as standing at the wellspring of not one, but two principal normative-theoretical traditions of modernity: the Kantian deontological universalism (the paradigm of "equal rights") and the "ethics of authenticity" tradition that spans the works of Herder, Schiller, Kierkegaard, Nietzsche and Heidegger. Arguing against the clichéd misinterpretations of Rousseau as "primitivist" or "totalitarian", Ferrara demonstrates that Rousseau's ideal of personal authenticity, as formulated in *Julie, or The New Heloise* is also at the foundation of his republican vision of the good society laid out in the *The Social Contract*.

In Ferrara's analysis, the two dimensions of the Swiss philosopher's thought are brought together by being refracted through both prisms of Ferrara's own synthetic perspective – the late-Rawlsianism and the paradigm of exemplarity. In other words, the "*The Social Contract* Rousseau" (the political philosopher) is read, not only through the late-Rawlsian optic as a reader might expect, but also through the lens of Ferrara's paradigm of exemplarity, while the "*Discourses* Rousseau" (the critic of modernity) is analyzed not only as a theorist of authenticity but as one that anticipates contemporary anti-authoritarian strands of political thought. Ferrara first shows us that Rousseau the critic of modernity is not epistemologically authoritarian – relying on Harry Frankfurt's influential conception of "orders of volition",

he demonstrates clearly that Rousseau's ideal of authenticity is not essentialist, but involves a reflective interplay between "first-", "second-" and "third-order" volitions. Ferrara then ventures to convince us that Rousseau the political theorist is not contextually insensitive (or, in Richard Rorty's words, violent towards the "final vocabularies" of historically situated actors and collectivities), in the sense that his vision of the good society, epitomized by the self-determining "general will", does not crush the historical lifeworlds under the weight of an abstract, context-insensitive blueprint for utopia.

Woven into this fabric of Ferrara's reconstruction is a complex and layered argument that revolves around the conception of the "normativity of identity" that Ferrara sees as central to Rousseau's perspective. This argument is gradually elaborated over the course of the three sections, but a reader unfamiliar with Ferrara's work might to some extent fail to take its full measure. In a nutshell, Ferrara argues that a tension that potentially arises in the process of identity formation (personal as well as collective) between our striving to order our conduct around abstract principles, on the one hand, and a fundamental need that we be able to experience our lives as coherent "narratives", on the other, can be the source of strong emancipatory impulses. Ferrara's paradigmatic example is the case of Rousseau's Julie, a case of "failed exemplarity" since Julie resolves the above tension - between her "Kantian" striving not to be happy at the cost of others' unhappiness and the need to actualize her love for Saint-Preux – in favour of the "principled" side. As result, Julie's self gradually disintegrates due to inauthenticity. But what is crucial here is that Ferrara's Rousseauianism does not merely point to a tension between "societal expectations" (the Meadean "generalized other") and our "innermost needs".

When Ferrara transposes the argument onto the level of collectivities (through the analysis of the "legislator" in Rousseau's *The Social Contract*), he demonstrates that the fundamental tension (let us call it the "Kantian-Romanticist" one for want of a better term) resurfaces even as we engage in revolutions, trying to overcome the existing world of "societal expectations". Rather than imposing a blueprint for utopia on the citizenry, the role of the legislator is to "gesture" towards a vision of the good society that's best for us *in light of what we are*. Those who might be tempted to read conservative overtones into Ferrara's argument would fail to appreciate the extent to which this "in light of what we are" is understood in terms of "singularity" rather than essence. Ferrara suggests that the Rousseauian "situated normativity" of political communities might nowadays be thought of along the lines of Rawlsian "political identity" (p. 54). But there is another possible toolkit for deciphering Ferrara's Rousseauianism – notwithstanding his reservations toward Adorno's facile dismissal of Rousseau as a "primitivist", Ferrara's perspective can indeed be read as an Adornian corrective to the progressism that still believes today it can overcome a world of alienation through "administering" the right dose of a trans-contextually valid utopia.

CARL SCHMITT, *DER SCHATTEN GOTTES: INTROSPEKTIONEN, TAGEBÜCHER UND BRIEFE 1921 BIS 1924*, HRSG. VON GERD GIESLER, ERNST HÜSMERT UND WOLFGANG H. SPINDLER, BERLIN, DUNCKER & HUMBLOT 2014.

CARL SCHMITT, *TAGEBÜCHER 1925 BIS 1929*, HRSG. VON MARTIN TIELKE UND GERD GIESLER, BERLIN, DUNCKER & HUMBLOT, 2018.

Petar Bojanić
Željko Radinković

CARL SCHMITT, *DER SCHATTEN GOTTES: INTROSPEKTIONEN, TAGEBÜCHER UND BRIEFE 1921 BIS 1924*.

In seinem Buch *Carl Schmitt: Denker im Widerstreit* (Freiburg/München, Verlag Karl Alber, 2017) weist Reinhard Mehring darauf hin, dass Carl Schmitt (1888-1985) ein ambivalentes Verhältnis zur Tagebuchführung hatte. Denn Schmitt führte sein ganzes Leben lang Tagebuch, doch zugleich spottete er über die Praxis der Tagebuchführung als lebensfremde und -fliehende Tätigkeit eines „Buribunken" bzw. des modernen Menschen, der das Tagebuch füllt und für das Tagebuch lebt. Darüber hinaus finden sich in seinen Tagebüchern Stellen, die auf das Bewusstsein der Bestimmung seines Lebens und seiner Karriere hindeuten. In den hier zu besprechenden Tagebüchern des Lebensabschnitts 1921-1924 schreibt Schmitt folgendes: „Ich kümmere mich nicht um den kommenden Tag, mir gehört die Zukunft. Zwischen dem kommenden Tag und der Zukunft ist ein großer Unterschied".

Die Ausgabe der Tagebücher ist in drei Teile gegliedert. Teil I umfasst den Zeitraum vom August 1921 bis August 1922, Teil II die beinahe lückenlose Alltagsaufzeichnung der Jahre 1923 und 1924. Für Teil III wählte Schmitt selbst den Titel „Der Schatten Gottes" und darin lassen sich Einträge unterschiedlichster Art finden: Vorlesungsentwürfe, Selbstreflexionen, Aphorismen etc. Die Ausgabe enthält auch einen wertvollen Anhang mit Briefen, Dokumenten und Abbildungen aus diesem Abschnitt seines Lebens.

Die Jahre 1921 bis 1924, die diese Tagebücher umfassen, bedeuten für Schmitts akademische Laufbahn einen langsamen Aufstieg. In diesen Zeitraum fallen die Veröffentlichungen seiner Werke *Diktatur* (1921), *Politische Theologie* (1922), *Römischer Katholizismus und politische Form* (1923), *Die geistesgeschichtliche Lage des heutigen Parlamentarismus* (1923). In diese Zeit fällt auch, nach der Tätigkeit an der Handelshochschule München (1920), der Ruf auf

die erste Professur an der Universität im preußischen Greifswald (Wintersemester 1921/22), an der sich Schmitt nicht lange aufhalten und eher unmotiviert arbeiten wird, und die Professur in Bonn (1922), wo er bis zu seinem Wechsel nach Berlin im Jahr 1928 bleibt.

In der Bonner Zeit kommt Schmitt zu einigen seiner wichtigsten theoretischen Ansätze (die Unterscheidung Freund und Feind als Hauptmerkmal des Politischen) und einige der prägnantesten Formulierungen, wie jene: „Souverän ist, wer über den Ausnahmezustand entscheidet" (*Politische Theologie*). Reinhard Mehring bemerkt, dass die Jahre in Bonn als Schmitts beste Schaffensperiode anzusehen sind: „Hier arbeitet er seine Verfassungstheorie aus; hier formt er ein Seminar mit namhaften Schülern. Er war zweifellos ein charismatischer Menschenfischer" (*Carl Schmitt: Aufstieg und Fall*, München, C. H. Beck, 2009, S. 177.). In dieser Zeit ist Schmitt auch ein unmittelbarer Zeuge der politischen Krisen, wie etwa der im 1923 erfolgten französischen und belgischen Besatzung des Ruhrgebiets, und auch der wirtschaftlichen Krise und der Inflation. Im Tagebuch finden sich Einträge, die etwa von den Sorgen wegen der Entwertung des Dollars" zeugen (S. 159).

Der Beginn der zwanziger Jahre ist auch die Zeit des offen gelebten und in den Tagebüchern dokumentierten Selbstzweifels und der Ängste. Es ist die Zeit, in der seine erste Ehe mit Carita zusammenbricht und in der Schmitt 1926 zum zweiten Mal heiraten wird, aber auch die Zeit, in der er sich, von seiner Erotomanie getrieben, in zahlreiche Affären stürzt. Die Tagebücher zeugen hier von einer intimen Beziehung zur irischstämmigen Australierin Kathleen Murray, der er bei dem Abschluss ihrer Dissertation behilflich ist und mit der er, nachdem sie 1922 nach Australien zurückkehrte, noch bis in die 1960er Jahre Kontakt hält. In Bonn angekommen, wird sich Schmitt relativ schnell in das universitäre Leben eingliedern. Der Tagebucheintrag vom 22. Januar 1923 beinhaltet auch Eindrücke von Schmitts erster Begegnung mit seiner künftigen Ehefrau Dušanka Duška Todorović: „Nach dem Essen im Vormfelde zwei Dollar gekauft, dann zu einer Serbin, die mir die Angelegenheit Agram übersetzt. Nachher mit ihr bei Ritterhaus, sie ist sehr sympathisch, aber nicht schön. Dagegen sah sie auf der Straße elegant und graziös aus. Ich begleitete sie, die Vorlesung bei Spiethoff, in der sie gehen wollte, fiel aus, sodass wir einen Spaziergang in die Dämmerung am Rhein machen konnten. Wir sprachen schön, sie heißt Duschanka Todorowitsch. Seltsam. Ich war oft daran, mich zu verlieben, aber zu müde und gleichgültig. Sprach schön, philosophierte, es schien ihr zu gefallen. Begleitete sie in die Argelanderstraße nach Hause, verabredete mich für morgen Abend zu Schmitz. Zu Hause war nichts, die Lamberts waren auch nicht da gewesen. Eine Karte von Koellreutter. Gleichgültig, aber zufrieden, eine neue Bekanntschaft zu haben." (S. 145.) Bereits die Einträge vom 14. Februar und der folgenden Tagen zeigen eine steigende Zuneigung zu Duška: „Ich suchte natürlich nur Duschka, sie war da (ich war benommen, wie ich sie nicht gleich sah), ging mit ihr zur Pension, zum Essen, dann zum Schmitz. (...) Duschka war äußerst klug und gut. Ich gewann sie wieder lieb. Sie ist vernünftig." (S. 157–159).

CARL SCHMITT, *TAGEBÜCHER 1925 BIS 1929*.

Es lassen sich drei neue Momente in der Bonner Periode des Lebens von Carl Schmitt bzw. in der Zeit seiner Professur hervorheben. Das erste, obwohl keine ganz neue Charakteristik dieser Periode, ist die Vermehrung seiner Verpflichtungen, weil er in Bonn sehr komplexe Kurse anbietet. Schmitt lebt auch weiterhin sehr aktiv, seine Beziehung zu seiner Ehefrau ist sehr kompliziert infolge ihrer chronischen Lungenkrankheit. Die Anwesenheit mehrerer Frauen in seinem Leben ist immer noch sehr intensiv, jedoch führt die akademische Tätigkeit Schmitts dazu, dass das in diesen Tagebüchern am meisten erwähnte Wort, das am besten diesen Zeitabschnitt seines Lebens beschreibt, „Müdigkeit" ist.

Zwei weitere Momente, die interessant sein könnten, sind Schmitts erste Jahre mit

der zweiten Ehefrau und eine wichtige, im September 1925 unternommene Reise nach Jugoslawien, das Geburtsland seiner Ehefrau. Diese Reise und die Forschungsarbeit in der Bibliothek von Valtazar Bogišić in der nahe Dubrovnik gelegenen Stadt Cavtat werden zusätzlich seine Zuneigung zu diesen Gegenden und zur serbischen Poesie und Kultur überhaupt orientieren. Diese Reise wird Schmitt unmittelbar zum im selben Jahr erfolgten Verfassen des Textes "Illyrien" inspirieren. Das größte Ereignis dieser Jahre ist jedoch seine Arbeit an dem Begriff des Politischen und die erste Redaktion des Textes *Der Begriff des Politischen*. In diesem Kontext ist besonders wertvoll die Korrespondenz mit Hermann Heller, die sich als Zusatz in diesem Buch befindet (S. 500–504). In dem im Wintersemester 1925-1926 an der Universität Bonn gehaltenen Seminar "Philosophie der Politik" spricht Schmitt zum ersten Mal über den Feind. Damals entsteht die erste Version des Textes "Der Begriff des Politischen", die erst zum Ende des Sommers 1927 in Heidelberg in der Zeitschrift *Archiv für Sozialwissenschaft und Sozialpolitik* (LXII, 1927) veröffentlicht wird.

Auf Einladung von Hermann Heller, der zu dem Zeitpunkt an der Berliner Hochschule für Politik unterrichtet, hält Schmitt am 10. Mai 1927 den Vortrag *Der Begriff des Politischen* (zweite Version). Dies geschieht also vor der Veröffentlichung der ersten Version des Textes. Über die enge Freundschaft zwischen diesen beiden zeugt auch der Brief von Haller vom 17. April 1927, in dem er sich für den mehrtägigen Aufenthalt im Bonner Hause der Schmitts bedankt: „(...) *für die beglückenden Tage in Ihrem Haus von Herzen zu danken*" (Paul Noack, *Carl Schmitt. Eine Biographie*, Berlin/Frankfurt am Main, Propyläen, 1993, S. 118). Diese zweite Version des Textes *Der Begriff des Politischen* wurde zusammen mit dem Text von Heller Ende 1928. Hallers Vortrag, bei dem Schmitt nicht anwesend war, hielt er eine Woche später am 17. Mai 1927. Heller hat das Manuskript des erwähnten Vortrags von Schmitt vor sich gehabt, worauf ein präzises, auf den Seiten 37-38 zu findendes Zitat hinweist. Dem auf 22. Dezember 1928 datierten, letzten Brief an Heller zufolge, hat Schmitt demgegenüber den Text von Heller vor seiner Veröffentlichung nicht gekannt (bis jetzt gibt es keine Hinweise über seine Redaktion). Anstatt dass er diesen Text an seinen Freund Schmitt schickt, entschied sich Heller, ihm im Dezember 1928 ein kleines, mit der Widmung versehenes Manuskript unter dem Namen „Bemerkungen zur staats- und rechtstheoretischen Problematik der Gegenwart" zu schicken.

(Dieser Text wurde in *Archiv für öffentlichen Rechts*, XVIII (1929), S. 338 veröffentlicht. Auf dem Manuskript befindet sich die Widmung und die Anmerkung, dass er 1927 geschrieben wurde.) (*Nachlass Carl Schmitt*, Siegburg, Respublica-Verlag, 1993, S. 575.) Offensichtlich möchte Heller mit dem einen oder anderem alusiven Satz, in dem er überhaupt nicht erwähnt wird, Schmitt und seine Bereitschaft zur Kritikakzeptanz prüfen. Hier ist der Abschnitt:

> Dann ist der Tat die Grundkategorie des Politischen das Begriffspaar Freund-Feind, wobei der Nachdruck auf dem existentiell anders gearteten und im Konfliktfalle zu vernichtenden Feind liegt. Der Sinn aller Politik und Geschichte ist dann der Kampf um die nackte Macht... (Hermann Heller, *Gesammelte Schriften*, Band 2, Tübingen, J.C.B. Mohr, Paul Siebeck, 1992, S. 264.)

(In dem posthum veröffentlichten Buch *Staatslehre* (1934), das er vor seinem Tode im Jahre 1933 schrieb, notiert Heller alle diese Sätze über den "*Advokat des deutschen Faschismus, Carl Schmitt*".) (Hermann Heller, *Gesammelte Schriften*, Band 3, Tübingen, J.C.B. Mohr, Paul Siebeck, 1992, S. 314.)

Zwei Briefe von Schmitt vom 18. Und 22. Dezember 1928 beenden diese Freundschaft. I, ersten Brief, der auch als Beilage zum Text von P. Tommissen „Problemen rond de houding van Carl Schmitt in en na 1933" (*Liber Memoralis*, Tien Jaar Economische Hogesschool Limburg, Limburg, 1979, S. 182–183.) veröffentlicht wurde, verlangt Schmitt, korrekt zitiert zu werden, weil er

das, was Heller suggeriert, nirgendwo geschrieben habe. Auch erinnert sich Schmitt nicht, dass er in einem der Gespräche mit Heller gesagt hätte, der Feind solle vernichtet werden.

Die Vernichtung ergibt sich erst aus der Fälschung politischer Begriffe, durch eine Moralisierung und Juridifizierung, und es war gerade der Sinn meines Aufsatzes, gegenüber dieser Verwirrung die einfache Wahrheit zu restituieren (S. 182).

BENJAMIN MOFFITT, *THE GLOBAL RISE OF POPULISM. PERFORMANCE, POLITICAL STYLE, AND REPRESENTATION*, PALO ALTO, STANFORD UNIVERSITY PRESS, 2016.

Agustin Cosovschi

During recent years, several countries have witnessed the rise of political movements and parties which question to some extent many of the conventional values and rules of liberal democracy. Frequently referred to as "populist" by the media, public intellectuals and by other, more traditional politicians, political forces such as MoVimento 5 Estelle in Italy and Podemos in Spain have been identified by many as a symptom of a deeper and more general crisis in contemporary democratic politics. And yet, despite the use of the same concept to describe them, differences among these so-called "populists" are sometimes remarkable: for instance, one could hardly put Evo Morales' left-wing government in Bolivia, prone to economic redistribution and to the reinforcement of social and cultural rights, in the same group as Viktor Orbán's right-wing and authoritarian administration in Hungary, fervently opposed to immigration and always keen to undermine minority rights. And the same goes for Cristina Kirchner's government in Argentina, Marine Le Pen's Front National in France and Andrés Manuel López Obrador's party in Mexico. So, if the actors and forces that fall under this common dominator of "populism" are indeed so different and diverse, why use the same concept to name them? What is the content and the utility of the concept of *populism* today? In other words, what does one say when one talks of "populism"?

Drawing from the existing literature in social sciences, Benjamin Moffitt's book *The Global Rise of Populism* addresses this and other key questions, going beyond the purely nominal discussion and trying to reach an updated, complex and comprehensive definition of contemporary populism. The main hypothesis of the book is that populism today has been changed by the developments in new media technologies and by shifting modes of political representation and identification. As a result, the author claims, it is more productive to analyze the phenomenon not as a fixed entity, but rather as a political style that is performed, embodied and enacted across different contexts. This definition, he sustains, not only unifies the aesthetic, discursive and ideological dimension of politics, but it also allows for populism to be thought of as a truly *global* phenomenon, beyond regional variations.

The study has both a strong theoretical ambition as well as an aspiration to set a number of operational criteria on which a comparative approach to contemporary populism could be built. With that aim, the author deploys a two-phase analysis, starting with a conceptual and terminological

definition of the object and subsequently developing an inquiry of the main features of contemporary populism.

Firstly, Moffitt resorts to developments in the field of political theory and political sociology to develop a notion of populism capable of overcoming the biases and shortcomings of the existing literature. In the first chapters of the book, Moffitt builds on the thoughts of authors such as Cas Mudde, Kirk Hawkings, Ernesto Laclau, Pierre Taguieff, Robert Hariman, Frank Ankermit and Dick Pels to finally advance a concept of populism as *a political style* that is performed by different actors in different contexts. The advantage of such a focus, the author claims, is that it does not only overcome the limitations of previous notions that conceived it as either an ideology, a discourse or a political logic, but that it also recognizes the constructed character of political performance, something that the author considers of key importance due to the ever more mediatized character of today's politics". Secondly, after having established his theoretical premises, the author draws from an examination of twenty-eight cases of political leaders coming from North America, Europe, Latin America, Africa and Asia that have commonly been identified as "populists" by the specialized literature in English language in order to establish the main features of contemporary populism.

To begin with, the author focuses on one of the most salient characteristics of populist politics: the supremacy of one central character, the political leader. According to Moffitt, these mediatized times make the leader the dominant character of populist politics and it is through him or her that one can determine some of the main characteristics of populism. Among others, the author underlines the constant tension that populist leaders have to deal with: having to show him or herself as ordinary and as part of "the people", for example through the use of "bad manners", and at the same time having to present himself as extraordinary, for instance through the exaltation of strength, power and virility.

Moreover, if the leader is the main character of the theatrical representation of populism, claims the author, the media should be considered its central stage, and a very wide one indeed. Moffitt holds that the use of the media and communication strategies by populist leaders should be in the center of our reflection on contemporary populism, especially due to the emergence of new decentralized media that allow for leaders to reach their audiences ("the people") in a low-cost and efficient manner. What is more, the author claims, the importance of the media in contemporary politics not only explains the intensive use and concern for the media by populist leaders, but also the tendency of many traditional and non-populist politicians to imitate or emulate features of populists in order compete with populist leaders and try to increase their own popularity. At the same time, according to the author, the analysis of the media should go beyond the mere study of how leaders use and abuse political communication, and should in exchange scrutinize what he called "the process of mediation": in other words, the countless ways through which "the people" is rendered present in a never-ending process of representation that involves leaders, constituencies, audiences and media, and in which populists leaders and other political actors struggle over who "the people" actually are and what are their true aspirations.

Furthermore, Moffitt's study focuses on a key dimension of contemporary politics: the meaning and the role of crises in the political process. The author advances here an interesting and challenging notion: that crisis is thus not *external*, but *internal* to populism. According to Moffitt, a key component of contemporary populism is the ability of populist leaders to interpret certain political and economic failures as a "crisis" that demands their intervention, them being the spokesmen of the true "people" and the only ones capable of confronting "the elite" and any other actor that poses a threat to society. In other words, contrary to authors who view the emergence of populism as a result of a

preexistent crisis, the author proposes to analyze populism not as a consequence of crisis, but as its main performer.

Finally, the author also addresses one of the main debates surrounding populism: its unclear and controversial relation to democracy. As most of the literature has claimed, by underlining majoritarianism and the sovereignty of "the people", populism can endanger the rights of the individual and thus democracy itself. At the same time, as other authors have noted, populism can be not only a threat, but also a corrective to democracy, since it encourages the participation of the masses in the political process. According to Moffitt, interpreting populism as a political style allows us to a certain extent to deconstruct these dilemmas and it allows us to see it as a phenomenon that is torn between two competing directions: it can both render politics more understandable for everyday citizens and offer effective critiques of the shortcomings of democratic systems, as it can also lead to a radical personalization of politics and deny the complexity of problems and offer simplistic solutions. The populist character of a political project, claims the author, actually tells us very little about its democratic content, and populism as a political style is thus able to endanger democracy as it is capable of making it richer and more inclusive.

The book *The Global Rise of Populism* certainly presents interesting and challenging debates regarding one of the mostly discussed political phenomena of our times. Nevertheless, in spite of the many virtues of Moffitt's work, a number of problems arising from his treatment of populism should be underlined.

Firstly, one cannot but question the limited nature of his sources, which are composed exclusively of secondary sources in English. This poses at least two problems: one the one hand, analyzing a globally extended phenomenon through the lens of only one language will inevitably lead to some degree of bias; on the other hand, aspiring to overcome the problems of the existing literature on the basis of a theoretical framework that builds exclusively on that same literature also seems methodologically unconvincing.

Secondly, another objection should be directed at one of the book's fundamental hypothesis. The study is built on the notion that, since the media are much more developed now than ever before, and politics are to a large extent channeled through those media, politics have thus become more "stylized" and therefore the aesthetic and performance dimension of political representation has become more important than in the past. However, bearing in mind the works of historians such as Louis Marin or George Mosse, one could claim that aesthetics has been a key dimension of political representation all throughout modernity, in contexts as diverse as XVIII century France or 1930s Germany. Therefore, if current populism is essentially different from what it used to be, and it is partially as a result of the role of the media, the rising importance of aesthetics is not the key feature of this novelty. Analyzing the media's influence on contemporary politics should thus go beyond the simple confirmation that "aesthetics matter".

The aforesaid objections, however, do not question the great value of *The Global Rise of Populism* as an original and stimulating contribution to contemporary political theory. Benjamin Moffitt's book constitutes an excellent and challenging work: through its clarity and its drive to challenge old and dates notions of how politics work, the book pushes us not only to renew our thoughts on populism, but also to challenge traditional approaches to politics in the global age.

JOHN M. COOPER, *PLATO'S THEAETHETUS*, NEW YORK, ROUTLEDGE, 2015.

Aleksandar Kandić

This is the second edition of John M. Cooper's book on Plato's *Theaetetus*, and it appears to be unchanged from the first edition published in 1990. Cooper is a prominent ancient philosophy and Plato scholar, perhaps mostly known for the edition of Plato's *Complete Works*, together with D. S. Hutchinson back in 1997, as well as various philosophical studies such as *Pursuits of Wisdom: Six Ways of Life in Ancient Philosophy from Socrates to Plotinus* (2012), *Reason and Emotion* (1999), *Reason and Human Good in Aristotle* (1975), etc. Holding his study of *The Theaetetus* means holding an important piece of the history of philosophy in your hands - it is not a book which one leaves on the shelves after reading, it needs to be reflected upon and absorbed through time. Its polemic structure simply begs for a deeper analysis of the topics discussed. So, what are the main interpretive problems in Plato's *Theaetetus*, and what solutions does Cooper offer?

Above all, one has to notice that Cooper does not consider *The Theaetetus* in isolation from the other dialogues. His analysis of Plato's influential epistemological treatise seems to be a part of a much greater endeavor: showing that there are inconsistencies and disagreements between the most important metaphysical and epistemological doctrines of the middle period dialogues, particularly *The Republic*, and later dialogues, such as *The Theaetetus*, *The Philebus*, or *The Sophist*. Cooper claims there are two main groups of Platonic scholars - those who believe that the main philosophical conclusions of *The Republic* constitute the core of Plato's philosophy which is retained in *The Theaetetus* and other dialogues, and those who challenge this position by arguing that the conclusions of the late dialogues significantly differ from the middle period metaphysics (p. 3). In his book, Cooper allies with the second group of scholars, and identifies his main opponent in Paul Shorey, who follows F. M. Cornford, A. E. Taylor and other "orthodox", or "traditional" interpreters of Plato's philosophy. According to Cooper, Shorey strongly rejected *any* idea about revisions and developments in Plato's philosophy, he was against the theory that the dialectical dialogues criticize and reject some of the central philosophical conclusions of middle period Platonism, as well as that there is any alteration of Plato's position before or after *The Republic* (p. 4). One of the main concerns of Platonic scholars is, of course, the notion of Forms. In *The Republic*, Forms are fundamental to the explanation of knowledge, but in *The Theatetus*, this is not the case. Plato seeks a different definition of knowledge. Even so, scholars such as Shorey, do not find this sufficient to argue that *The Theaetetus* represents a significant divergence

from the middle period epistemology. Shorey attacks any "genetic" interpretation of Plato's philosophy, such as the one which developed in 19th century Germany. Cooper's main task is, therefore, to show that thorough analysis of Plato's argument in *The Theaetetus* reveals irreconcilable differences between the middle and the late period epistemology. In his opinion, this is sufficient to disapprove Shorey and the first group of scholars.

The book has a simple, well-thought structure which follows the structure of *The Theaetetus*. In the first chapter, Cooper deals with the "amalgamation" of the epistemological theories of Theaetetus, Protagoras, and Heraclitus (*The Theaetetus*, 151a-161a). The preliminary questions about the nature of knowledge are raised here, and Theaetetus proposes the first definition of knowledge as perception, only to be refuted by Socrates throughout the dialogue. In Plato's view, Theaetetus' understanding of knowledge as perception incorporates the most important aspects of the Protagorean and Heraclitean doctrines, the first being that "it is incorrect to say that something is true without saying to whom it is true" (152a), and the second that "all things are in flux" (152d). This way, Plato's Socrates actually argues against Protagoras and Heraclitus whose views are put to the mouth of "naive" Theaetetus. Cooper takes a great deal of effort to explain how could Protagaras' theory be reduced to the one of Heraclitus, as this is not immediately clear to Plato's readers (pp. 14-26, ff.). Some of the relevant passages from *The Timaeus*, concerning the notions of *genesis* and *ousia*, are recalled here by the American author (28a-b, 37e-38a, 51b, etc.). But, as we shall see, for Cooper, the main comparison is the one between *The Theaetetus* and *The Republic*.

The second and third chapter of the book discuss Plato's arguments against Protagoras and Heraclitus in detail. According to Socrates, Protagoras is bound to accept that some people's opinions are false, as well as that some people are wiser (or, less ignorant) than others, on the ground that classes of judgments exist in which error is possible (pp. 85-87). Protagoras relativistic conception of knowledge thus fails. On the other hand, Heraclitus' position is refuted by pointing out that if everything is in flux, then the witness himself is constantly changing - he is no more real than colors and appearances he's perceiving (pp. 91-92). Plato obviously thought that Heraclitean theory of flux is the underlying metaphysical basis for Protagorean relativism. Both theories introduce the aspect of instability into our notion of knowledge. It is also worth noting that authors such as Cooper, G. E. L. Owen, or H. F. Cherniss, pay much more attention to Plato's critique of Heracliteanism in *The Theaetetus* than, for example, W. K .C. Guthrie, or A. E. Taylor do. However, this is not Cooper's main concern. The main purpose of his argument is to show that Plato's insistence on the stable character of knowledge and its disassociation from perceptual world doesn't *automatically* amount to the middle period metaphysics and the Forms of *The Republic*. We have already seen that the Forms are not explicitly brought up as the objects of knowledge in *The Theaetetus*, but scholars such as Cherniss and Cornford, whom Cooper critizes in the third chapter of his book (p. 7, 121), stick to the claim that Plato's position in *The Theaetetus* is nothing but a continuation, or addendum to the epistemology of *The Republic*.

The fourth chapter of Cooper's analysis of *The Theaetetus* is probably the most significant ("The Refutation of the Sense-Perception Theory of Knowledge", pp. 118–140), as the reader is acquainted with the key premises of his argument against Shorey, Cornford, Cherniss, and other traditionalist interpreters of Plato's philosophy. There is no doubt that, for Plato, perception cannot be equated with knowledge, "with or without the aid of Heraclitean metaphysics" (p. 118). *The Theaetetus* 184b-186e ensures us that Plato's explanation of knowledge rests on certain entities which are not objects of immediate perception, but somehow constitute perception and enable the experience of knowledge. The nature of these non-perceptual

entities is questionable for the scholars, and the interpretations differ in great extent. Instead of *eidos* (usually translated as Form), Plato employs the term *koina* in order to denote such entities (185c), and *koina* are "common terms", such as existence and non-existence, similarity and dissimilarity, sameness and difference, etc. This is actually in line with the other late dialogues, such as *The Sophist*, or *The Timaeus*. In somewhat vague manner, *koina* are associated with *ousia*, which denotes permanent, pure, real existence, and without *ousia*, knowledge is impossible (152c, 186c). According to Cooper, traditional, or "conservative" Platonic scholars interpret these passages in *The Theaetetus* as consistent with the two-worlds argument found in *The Republic* (the passage 523b-524b is emphasized by the author of the book). As we know, in *The Republic*, knowledge is restricted to the unchanging world of Forms, and there cannot be knowledge of perceptual objects, only belief (*doxa*). But if Plato claimed this in such an elaborate, extensive manner in *The Republic*, how come Forms do not play any role in the explanation of knowledge in *The Theaetetus*? Did he change his mind, or did he wanted the Forms to be implicitly present in *The Theaetetus*? The "easier" approach to this interpretive problem is taken by Shorey and traditional scholars, for whom the argument in *The Theaetetus* represents some kind of weaker version of the argument in *The Republic*. On the other hand, Cooper is right to point out that such interpretations are not supported by Plato's writing. Even if *koina* are Forms, nowhere in *The Theaetetus* is knowledge restricted to *koina*, or intelligible entities (p. 121), which means that the dialogue is not a mere repetition of the theory of knowledge found in *The Republic*. Cooper's analysis appears to be in line with the aforementioned "genetic" school of interpretation, as well as, for example, S. Rosen's interpretation of *The Sophist* according to which Plato abandoned the concept of Forms in his late dialogues.

The rest of the book examines the last two definitions of knowledge proposed by Theaetetus: knowledge as true belief (187b-201c), and knowledge as true belief plus *logos* (201c-210b). Both are refuted by Socrates, and the dialogue reaches no definite solution to the problem of knowledge. Nevertheless, the definition of knowledge as *true belief plus logos* has become the cornerstone of traditional epistemology, being challenged in the last couple of decades only by, for example, contextualist theories. The impact of *The Theaetetus* on philosophical epistemology is therefore tremendous. In the final chapter, Cooper discusses Plato's conception of *logos* in *The Theaetetus* (pp. 234-279), which is supposed to be the explanation of true belief (*alethes doxa*) necessary for real knowledge. Four "senses" of logos are elaborated by Plato, and all four are found to be inadequate (some scholars, such as Guthrie, perceive three senses of *logos*, see: *A History of Greek Philosophy*, Vol. 5, Cambridge Univerity Press, Cambridge, 1978, pp. 117-120). Cooper considers the so-called "dream theory" a separate version of the first sense of *logos* (p. 237, ff.). According to this sense, the sufficient conditions for knowledge consist in the expression of thought in words. The next sense is enumeration of all parts, or elements of a thing, and the last one, the expression of a specific property by which something is differentiated from all other things. Being that all of these senses of *logos* are refuted by the end of the dialogue, we are back to the beginning of our inquiry.

Let us get back to Cooper's main argument. How successful is it in disapproving Shorey and the traditionalists? There is no doubt that *The Theaetetus*, *The Philebus*, or *The Sophist*, propose some kind of revision of middle period metaphysics. The dualistic position of *The Republic*, by which the world was split into the intelligible and physical realm is now softened and possibly abandoned. Plato's conception of knowledge does not revolve around transcendence anymore, but *objectivity* (p. 139). Even *The Timaeus* strives in this direction. While in *The Republic* knowledge was restricted to intelligible entities, in *The Theaetetus*, or *The Philebus* (61d-e), Plato

explicitly claims that knowledge of sensory, changing objects *is* possible. But, the fact that Plato's conception of knowledge remains grounded in certain intelligible, imperceptible principles, no matter if those are Forms, or *koina*, or something third, keeps the traditionalists in life. Although Cooper rightly claims that we cannot assimilate metaphysics and epistemology of the late period into the two-worlds ontology of *The Republic*, it is also true that Plato doesn't deny this solution explicitly by stating that his epistemological considerations in *The Theaetetus* are entirely separate from those of *The Republic*, as well as that Forms, from now on, do not play *any* role in the experience of knowledge. This topic is an ongoing discussion, and therefore, *Plato's Theaetetus* by John M. Cooper represents a highly valuable contribution to Platonic and ancient philosophy studies.

SUBMISSION INSTRUCTIONS

All submissions to Filozofija i društvo must conform to the following rules, mostly regarding citations. The Referencing Guide is the modified Harvard in-text referencing style. In this system within the text, the author's name is given first followed by the publication date and the page number/s for the source. The list of references or bibliography at the end of the document contains the full details listed in alphabetical order for all the in-text citations.

1. LENGTH OF TEXT

Up to two double sheets (60.000 characters including spaces), abstracts, key words, without comments.

2. ABSTRACT

Between 100 and 250 words.

3. KEY WORDS

Up to 10.

4. AFFILIATION

Full affiliation of the author, department, faculty, university, institute, etc.

5. BOOKS

In the bibliography: last name, first name, year of publication in parentheses, book title, place of publication, publisher. In the text: last name in parentheses, year of publication, colon, page number. In a comment: last name, year of publication, colon, page number. Books are cited in a shortened form only in comments.

Example:

In the bibliography: Moriarty, Michael (2003), Early Modern French Thought. The Age of Suspicion. Oxford: Oxford University Press.

In the text: (Moriarty 2003: 33).

In a comment: Moriarty 2003: 33.

6. ARTICLES

In the bibliography: last name, first name, year of publication, title in quotation marks, name of publication in italic, year of issue, in parentheses the volume number within year if the pagination is not uniform, colon and page number. In the text: last name in parentheses, year of publication, colon, page number. In acomment: last name, year of publication, colon, page number. Do not put abbreviations such as 'p.', 'vol.', 'tome', 'no.' etc. Articles are cited in shortened form only in comments.

Example:

In the bibliography: Miller, Johns Roger (1926), „The Ideas as Thoughts of God", Classical Philology 21: 317–326.

In the text: (Miller 1926: 320).

In a comment: Miller 1926: 320.

7. EDITED BOOKS

In the bibliography: last and first name of editor, abbreviation 'ed.' in parentheses, year of publication in parentheses, title of collection in italic, place of publication, publisher and page number if needed. In the text: last name in parentheses, year of publication, colon, page number. In a comment: last name, year of publication, colon, page number. Collectionsare cited in shortened form only in comments.

Example:

In the bibliography: Harris, John (ed.) (2001), Bioethics, Oxford: Oxford University Press

In the text: (Harris 2001).

In a comment: Harris 2001.

8. ARTICLES/CHAPTERS IN BOOK

In the bibliography: last name, first name, year of publication in parentheses, text title in quotation marks, the word 'in' (in collection), first and last name of editor, the abbreviation 'ed.' in parentheses, title of collection in italic, place of publication, publisher, colon, page number (if needed). In the text: Last name of author in parentheses, year of publication, colon, page number. In a comment: last name of author, year of publication, colon, page number. The abbreviation 'p.' is allowed only in the bibliography.

Example:

In the bibliography: Anscombe, Gertrude Elizabeth Margaret (1981), „You can have Sex without Children: Christianity and the New Offer", in The Collected Philosophical Papers of G.E.M. Anscombe, Ethics, Religion and Politics, Oxford: Basil Blackwell, pp. 82–96.

In the text: (Anscombe 1981: 82)

In a comment: Anscombe 1981: 82.

9. NEWSPAPER AND MAGAZINES ARTICLE

In the bibliography: last name, first name, year in parentheses, title of article in quotation marks, name of newspaper in italic, date, page.

Example:

In the bibliography: Logar, Gordana (2009), „Zemlja bez fajronta", Danas, 2 August, p. 12.

In the text: (Logar 2009: 12).

In a comment: Logar 2009: 12

10. WEB DOCUMENTS

When quoting an online text, apart from the web address of the site with the text and the text's title, cite the date of viewing the page, as well as further markings if available (year, chapter, etc.).

Example:

In the bibliography: Ross, Kelley R., „Ontological Undecidability", (internet) available at: http://www.friesian.com/undecd-1.htm (viewed 2 April, 2009).

In the text: (Ross, internet).

In a comment: Ross, internet.

UPUTSTVO ZA AUTORE

Pri pisanju tekstova za *Filozofiju i društvo* autori su u obavezi da se drže sledećih pravila, uglavnom vezanih za citiranje. Standardizacija je propisana *Aktom o uređivanju naučnih časopisa* Ministarstva za prosvetu i nauku Republike Srbije iz 2009. U *Filozofiji i društvu* bibliografske jedinice citiraju se u skladu s uputstvom *Harvard Style Manual*. U ovom uputstvu naveden je način citiranja najčešćih bibliografskih jedinica; informacije o načinu citiranja ređih mogu se naći na internetu.

1. VELIČINA TEKSTA

Do dva autorska tabaka (60.000 karaktera) s apstraktom, ključnim rečima i literaturom; napomene se ne računaju.

2. APSTRAKT

Na srpskom (hrvatskom, bosanskom, crnogorskom...) i jednom stranom jeziku, između 100 i 250 reči.

3. KLJUČNE REČI

Do deset.

4. PODACI O TEKSTU

Relevantni podaci o tekstu, broj projekta na kojem je rađen i slično, navode se u fusnoti broj 1 koja se stavlja na kraju prve rečenice teksta.

5. AFILIJACIJA

Puna afilijacija autora, odeljenje i fakultet, institut i slično.

6. INOSTRANA IMENA

Sva inostrana imena (osim u bibliografskim jedinicama) fonetski se transkribuju u skladu s pravilima pravopisa, a prilikom prvog javljanja u zagradi se navodi njihov izvorni oblik. Imena geografskih i sličnih odrednica takođe se fonetski transkribuju bez posebnog navođenja originala u zagradama, osim ukoliko autor smatra da je neophodno.

7. CRTA I CRTICA

Kada se navode stranice, od jedne do neke druge, ili kada se to čini za godine, između brojeva stoji crta, *ne crtica*.

Primer:
33–44, 1978–1988; ne: 33-44, 1978-1988.

8. KNJIGE

U spisku literature: prezime, ime, u zagradi godina izdanja, naslov knjige, mesto izdanja, izdavač. U tekstu: u zagradi prezime autora, godina izdanja, dvotačka, stranica. U napomeni: prezime autora, godina izdanja, dvotačka, stranica. U napomenama, knjiga se citira isključivo na skraćeni način.

Primer:
U literaturi: Haug, Volfgang Fric (1981), *Kritika robne estetike*, Beograd: IIC SSO Srbije.

U tekstu: (Haug 1981: 33).

U napomeni: Haug 1981: 33.

9. ČLANCI

U spisku literature: prezime, ime, u zagradi godina izdanja, naslov teksta pod navodnicima, naslov časopisa u italiku, godište časopisa, u zagradi broj sveske u godištu ukoliko paginacija nije jedinstvena za ceo tom, dvotačka i broj stranice. U tekstu: u zagradi prezime autora, godina izdanja, dvotačka, stranica. U napomeni: prezime autora, godina izdanja, dvotačka, stranica. Ne stavljaju se skraćenice „str.“, „vol.“, „tom“, „br.“ i slične. U napomenama, članci se citiraju isključivo na skraćeni način.

Primeri:
U literaturi: Miller, Johns Roger (1926), „The Ideas as Thoughts of God“, *Classical Philology* 21: 317–326.

Hartman, Nikolaj (1980) „O metodi istorije filozofije“, *Gledišta* 21 (6): 101–120.

U tekstu: (Hartman 1980: 108).

U napomeni: Hartman 1980: 108

10. ZBORNICI

U spisku literature: prezime i ime priređivača, u zagradi skraćenica „prir.“, u zagradi godina izdanja, naslov zbornika u italiku, mesto izdanja, izdavač i strana po potrebi. U tekstu: u zagradi prezime autora, godina izdanja, dvotačka, stranica. U napomeni: prezime autora, godina izdanja, dvotačka, stranica. U napomenama, zbornici se citiraju isključivo na skraćeni način.

Primer:
U literaturi: Espozito, Džon (prir.) (2002), *Oksfordska istorija islama*, Beograd: Clio.

U tekstu: (Espozito 2002).

U napomeni: Espozito 2002.

11. TEKSTOVI IZ ZBORNIKA

U spisku literature: prezime, ime autora, u zagradi godina, naslov teksta pod navodnicima, slovo „u“ (u zborniku), ime i prezime priređivača zbornika, u zagradi „prir.“, naslov zbornika u italiku, mesto izdanja, izdavač, dvotačka i broj stranice (ako je potrebno). U tekstu: u zagradi prezime autora, godina izdanja, dvotačka, stranica. U napomeni: prezime autora, godina izdanja, dvotačka, stranica. Skraćenica „str.“ dopuštena je samo u spisku literature.

Primer:
U literaturi: Nizbet, Robert (1999), „Jedinične ideje sociologije“, u A. Mimica (prir.), *Tekst i kontekst*, Beograd: Zavod za udžbenike i nastavna sredstva, str. 31–48.

U tekstu: (Nizbet 1999: 33).

U napomeni: Nizbet 1999: 33.

12. ČLANAK IZ NOVINA

U spisku literature: prezime, ime, u zagradi godina, naslov članka pod navodnicima, naslov novina u italiku, datum, stranica.

Primer:
U literaturi: Logar, Gordana (2009), „Zemlja bez fajronta“, *Danas*, 2. avgust, str. 12.

U tekstu: (Logar 2009: 12).

U napomeni: Logar 2009: 12.

13. INTERNET

Prilikom citiranja tekstova s interneta, osim internet-adrese sajta na kojem se tekst nalazi i naslova samog teksta, navesti i datum posete toj stranici, kao i dodatna određenja ukoliko su dostupna (godina, poglavlje i sl.).

Primer:
U literaturi: Ross, Kelley R., „Ontological Undecidability“, (internet) dostupno na: http://www.friesian.com/undecd-1.htm (pristupljeno 2. aprila 2009).

U tekstu: (Ross, internet).

U napomeni: Ross, internet.

CIP – Каталогизација у публикацији
Народна библиотека Србије, Београд

1+316+32

FILOZOFIJA i društvo = Philosophy and society / glavni i odgovorni urednik Rastko Jovanov. – Knj. 1 (1987)– . – Beograd : Institut za filozofiju i društvenu teoriju, 1987– (Beograd : Donat Graf). – 24 cm

Tromesečno. – Drugo izdanje na drugom medijumu: Filozofija i društvo (Online) = ISSN 2334-8577
ISSN 0353-5738 = Filozofija i društvo
COBISS.SR-ID 11442434