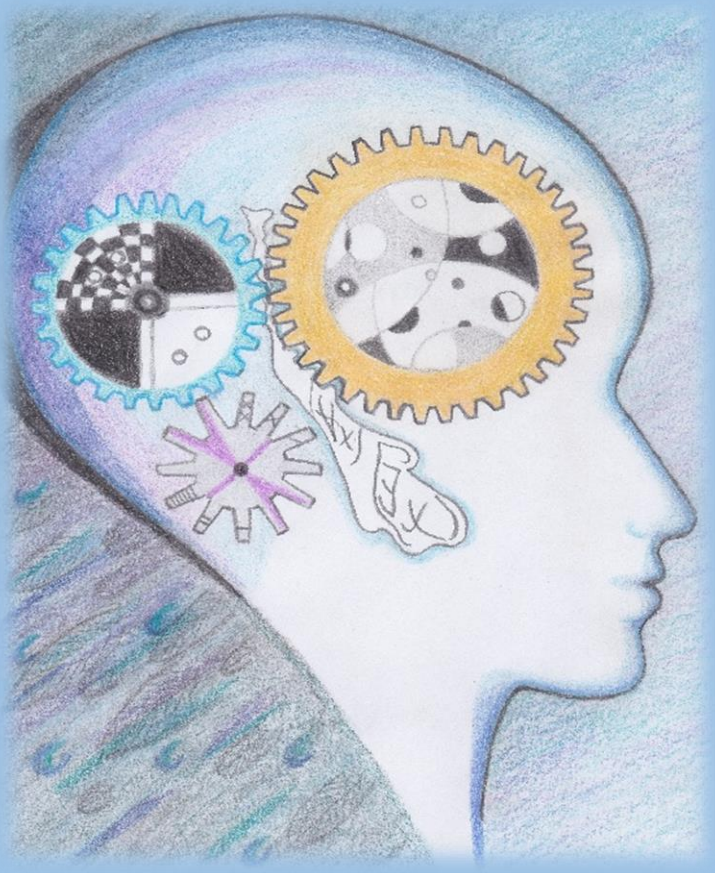
Univerza v Mariboru

Filozofska fakulteta

Študentski filozofski zbornik 2017

PHILOSOPHY STUDENTS' COMPENDIUM 2017



Zbornik študentov oddelka za filozofijo Filozofske fakultete Maribor

Maribor, 2018

ŠTUDENTSKI FILOZOFSKI ZBORNIK 2017

PHILOSOPHY STUDENTS' COMPENDIUM 2017

Zbornik prispevkov z mednarodnega študentskega simpozija/Compendium of works from the international philosophy symposium

Urednika/Editors:

Neja Kaiser, dipl. fil. (UN) in dipl. um. zgod. (UN)
Tadej Todorović, magister anglistike in magister filozofije

Recenzenti/Reviewers:

red. prof. dr. Bojan Borstner
izr. prof. dr. Janez Bregant
izr. prof. dr. Friderik Klampfer
izr. prof. dr. Rudi Kotnik
ml. razisk. Alen Lipuš
red. prof. dr. Danilo Šuster
izr. prof. dr. Boris Vezjak

Risba na naslovnici/Cover drawing: Tina Ritlop, dipl. um. zgod. (UN) in dipl. ang. (UN)

Lektorja/Proof-readers:

Tadej Todorović, magister anglistike in magister filozofije

Tina Ritlop, dipl. um. zgod. (UN) in dipl. ang. (UN)

Oblikovanje in prelom/Design and layout: Tadej Todorović, magister anglistike in magister filozofije

Izdala/Published by: Oddelek za filozofijo, Filozofska fakulteta, Univerza v Mariboru in Društvo za analitično filozofijo in filozofijo znanosti

Kraj/Published in: Maribor, Slovenija

Leto izida/Year of publication: 2018

ISSN 2464-0468

Zbornik je dostopen na spletnem naslovu/The compendium is accessible at: https://daf.splet.arnes.si/2018/06/26/studentski-filozofski-zbornik-2017/ .

Table of Contents

Preface

What you are perceiving in front of you (or what the Evil Demon is showing you at the moment) is the compendium of papers, presented at the International Students' Symposium 2017, which was organized by the Department of Philosophy, University of Maribor, and the Slovenian Society for Analytic Philosophy (DAF) in April 2017.

It is with great pleasure that I am able to say that we have been able to, with a commendable response to our invitation from colleagues from other universities, expand our, now almost traditional, philosophical symposium to an international event. As such, it seemed only fitting that the compendium is presented to the reader in the English language.

I would like to make use of this situation to thank everyone who have contributed to the creation of this compendium. First and foremost, many thanks to our reviewers, who have allocated their precious time to this cause; to my co-editor, Tadej Todorović, who, next to doing his fair share of editorial duties, also proof-read the articles and designed the compendium; to Tina Ritlop, who has allowed us to use her work as our cover for the second year in a row; and, naturally, to all the other organizers who have helped with the organization of the symposium.

I hope that reading this compendium will furnish you with an idea of what interests those, who are still just at the start of their academic careers, and perhaps even offer you some fresh perspectives and new knowledge.

Neja Kaiser

Foucault and Analytic Philosophy: Two Comparisons Regarding Knowledge

Nikola Cerovac, University of Zagreb

1 Introduction

In this paper, I compare Foucault's approach known as the 'archaeology of knowledge' with that of analytic philosophers, such as Wittgenstein and Quine. My aim is to show how Foucault, despite being known as a continental philosopher, shares certain conceptual similarities with analytic philosophy, which calls for a re-evaluation of the continental-analytic divide. In the first part of the paper, I shortly outline the characteristics of the archaeology of knowledge; in the second part, I argue for some similarities between Foucault's concepts and those of Wittgenstein and Quine. In the third part, I point to the crucial differences and, in the end, draw the implications that these differences have for our understanding of the analytic-continental divide.

Several reasons make Foucault and his philosophy an interesting object of comparison: first of all, in American academic circles, Foucault is considered to belong to the 'French theory', a group of authors (which, together with Foucault, includes philosophers like Deleuze, Derrida, and Althusser) which made the biggest impact in the US in the departments of literary and culture studies. This led to the majority of philosophers disqualifying them as culture theorists (at best), or charlatans, relativists and obscurants (at worst) (Cusset, 2008). Although in Foucault's case, we find some serious philosophical reception at a very early stage – best-known examples being Dreyfus and Rabinow (1983), and Ian Hacking (1975) and his work on historical ontology – these cases are more of an exception than a rule.

However, if we turn to Foucault's original French context, we find that Foucault has made a name for himself dealing with epistemology and philosophy of science, following the path of French rationalist historical epistemologists (Canguilhem, Bachelard). Furthermore, we find numerous indications that Foucault read analytic philosophers of language, such as Wittgenstein, Austin and Strawson, with great interest. Taking that into consideration, it is valid to pose the following question: Are there any common points between Foucault and the analytic philosophers?

2 The archaeology of knowledge as an epistemological theory

The archaeology of knowledge is an approach Foucault developed in his first phase, spanning from his first book *Madness and Civilization* to the methodological book *The Archaeology of Knowledge*, in which he tries to systematize the method of his previous books. The archaeology of knowledge can be defined as an approach to the bodies of knowledge ('discursive formations', be it certain sciences, like linguistics or economy; texts about certain objects, like madness; or practices, for example, medical clinical practices) in which the aim is to describe the formations, analysing only the relations between the statements as they are materially manifest. Paraphrasing Foucault, four main principles can be stated regarding the archaeology of knowledge (Foucault, 2002: 155–6):

> 1) archaeology analyses discourses as practices following certain rules and does not try to define the thoughts that should be 'revealed' in them ("it does not seek another, better-hidden discourse");
>
> 2) it analyses them in their specificity, that is, it avoids descriptions based on the continuity of themes or historical discourses, it is not trying to reduce them to a set of principles or axioms of a stable science;
>
> 3) it is not psychology, anthropology or sociology of the individuals involved in the production of discourses; on the contrary, it sees authors and individuals as a result of the rules of discourse;
>
> 4) its aim is not the reconstruction of some sort of *Zeitgeist* or *Weltanschauung*, but a description of a discourse in its materiality.

Hence, the object of the archaeology is discourse, or more precisely, a certain discursive formation. Discussing the discursive formation, Foucault states the following:

> Whenever one can describe, between a number of statements, such a system of dispersion, whenever, between objects, types of statements, concepts, or thematic choices, one can define a regularity (an order, correlations, positions and functioning transformations), we will say, for the sake of convenience, that we are dealing with a discursive formation. (ibid.: 42)

The statement, claims Foucault, cannot be defined only by its form, its reference, or place of articulation, since a discursive formation includes various domains of knowledge: science, administration, education, etc. A statement cannot be reduced to a grammatical sentence, since it does

not depend entirely on the grammatical rules of a language, nor can it be reduced to a logical proposition, since it is not necessarily deducible by a deductive procedure. Using Deleuze's (2006) topological metaphor, it can be said that grammatical rules determine the rules of sentence production on a horizontal axis, the logical rules determine the rules of logical validity on a vertical axis, and archaeological rules are revealed if we try to analyse discourse on a *diagonal* axis, connecting the statements with various other discourses and domains.

To sum up, the object of archaeology is a discourse understood as a set of linguistic elements; however, in describing a discourse, archaeology aims to delineate the rules specific to that particular discourse, the rules somehow formed by the discursive practice alone.

3 The archaeology of knowledge with regard to Wittgenstein and Quine

In the previous section, I was unfortunately only able to point out the most important aspects of Foucault's approach, leaving out the specific details, problems and examples. Nonetheless, I think the rest will become clearer in the course of the second part: the comparison with Wittgenstein and Quine. At this stage, the most important thing to notice is that according to Foucault, every discourse has its own rules, and that statements are the *results* of the discursive practice observing those rules.

If we turn our attention to Wittgenstein, the first analogy is not hard to notice: the role of rule-following. For Wittgenstein, the construction of sentences involves taking part in a language-game (Wittgenstein, 2010: §23), since every sentence is produced with regard to the specific situation and activity in question. As mentioned earlier, for Foucault, the agents of discourse are not so much autonomous individuals as they are the structural positions occupied in accordance with certain rules. In Wittgenstein's case, not following the rules disqualifies you from the game; in Foucault's, it disables you from making statements.

The existence of rules is based on the existence of some type of knowledge available to the subjects. As Paltrinieri (2011) notices, there is an analogy between what Wittgenstein calls a 'form of life' (*Lebensform*) (Wittgenstein, 1969: §358) and what Foucault calls '*savoir*' (Foucault, 2002: 200–3). A form of life refers to the fact that every language-game determines a certain mode of ontological certainty, a network of possibilities in which it is possible to say something. Such a form is internalized to such an extent that it is hard (or even impossible) for a subject to *explain* its functioning: he or she can merely point to the examples of usage or admit that it is 'just so'. On the other hand, *savoir* constitutes the domain of possible objects of reference. That is to say, it is a set of practices which

enable and limit the subject in his enouncing, determining the rules of statement formation (or the 'conditions of possibility'). As such, *savoir* cannot be reduced to scientific knowledge, but it includes various discourses that represent some kind of knowledge (e.g. religious and folk beliefs, administrative regulations, etc.). Although Foucault favours the topological metaphor of a field or a space of statements, as opposed to Wittgenstein's biological metaphor of a life-form, they both espouse the idea of statements which precede the enunciations of an individual subject and enable them in the first place.[1]

This brings us to the commonality between Foucault and Wittgenstein that I find crucial: that which I call the non-normative view on knowledge. Wittgenstein writes:

> [A]sk yourself whether our own language is complete – whether it was so before the symbolism of chemistry and the notation of the infinitesimal calculus were incorporated in to it; for these are, so to speak, suburbs of our language. (And how many houses or streets does it take before a town begins to be a town?) Our language can be regarded as an ancient city: a maze of little streets and squares, of old and new houses, of houses with extensions from various periods, and all this surrounded by a multitude of new suburbs with straight and regular streets and uniform houses. (Wittgenstein, 2010: §18)

What this paragraph tells us is that formal languages represent an expansion of our everyday language, i.e. they also function as a language-game, striving to a certain ideal, but a language-game nonetheless. Previously, we have concluded that participation in a language-game implies the knowledge of the rules, which now leads us to conclude that scientific language can also be viewed as the knowledge of the rules specific to the scientific discourse.[2] Again, turning back to Foucault: in his case, we find a second kind of knowledge, *connaissance*, understood as the knowledge of a specific object, more or less scientific. Foucault says: "Science (or what is offered as such) is localized in a field of knowledge and plays a role in it" (Foucault, 2002: 203). His point is that science cannot be seen as a sole place of knowledge, but that it stems from the pre-scientific knowledge and has a privileged position in it.

Moving on to Quine, the similarities become more superficial and difficult to notice. After all, Quine is a philosopher who tried to assimilate philosophy into natural sciences, considering its role to be that

[1] Recently, some authors (Volbers, 2015) have pointed out the similarities in Foucault's and Wittgenstein's understanding of subjectivation and its ethical implications. Although related to the current topic, it is not possible to pursue the similarity further due to the limited format of this paper.
[2] Here I am following the so-called 'sceptical' interpretation of late Wittgenstein; *see* Kripke (1982).

of conceptual systematization. However, I think that there is one aspect in which Quine and Foucault meet, and that is the question of underdetermination.

In *Two Dogmas of Empiricism*, Quine (1961) analyses two presuppositions of empiricism: the existence of analytic truths and the existence of terms and sentences that refer to the immediate experience. In his refutation of the two dogmas, Quine comes to the conclusion that the whole of language and science depends on the connections between the terms, i.e. the truth-value of each statement depends on its relation to the whole of the theory. To quote Quine:

> A conflict with experience on the periphery [of a theoretical field] occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Re-evaluation of some statements entails re-evaluation of others /…/ Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. (Quine, 1961:42-3)

The second problem that Quine dealt with in his work, which is relevant here, is the problem of radical translation (Quine, 1960). For the sake of brevity, I will just mention the well-known example of 'gavagai'. Quine describes a situation in which a speaker of some language unknown to the observer shouts the word 'gavagai' at the sight of a rabbit running across a field. How should a stranger, not familiar with the language, interpret that statement, Quine asks. Does it refer to the rabbit? Or to the whole situation of the rabbit running? Or perhaps to some 'rabbithood' manifest in the situation? Or maybe just to the parts of the rabbit that are of interest to the speaker? There is no way, claims Quine, of positing a clear relation between the stimuli and the statement.

Foucault tackles a similar problem when he talks about 'immature sciences'. In his work, Foucault begins with a simple question: What is this 'madness' we are talking about? Does natural history of the 18th century talk about the same thing as biology? Starting from this, he goes on to prove that the objects that were considered as given, as unchangeable, are actually dependent on the discourse and the relations in which we talk about them:

> Discursive relations are not, as we can see, internal to discourse: they do not connect concepts or words with one another; they do not establish a deductive or rhetorical structure between propositions or sentences /…/ they determine the group of relations that discourse must establish in order to speak of this or that object /…/ These relations characterize not the language (*langue*)

> used by discourse, nor the circumstances in which it is deployed, but discourse itself as a practice. (Foucault, 2002: 50–1)

There is, however, one major difference between Foucault and Quine that limits the analogy. As Hacking states:

> Quine's [fabric of sentences] is a body of beliefs, a "lore", partly theoretical, partly practical, but such as could be entertained as a pretty consistent whole by a single informant. Foucault's discourses are what is said by a lot of people talking, writing, and arguing; it includes the pro and the con and a great many incompatible *connaissances.* (2002: 92)

Quine's understanding of language is that of an ordered whole which, depending on the circumstances, changes and regulates its homogeneity according to the empirical observations. On the other hand, Foucault's discourse is a messy, chaotic place, in which various objects and contradictory claims exist simultaneously. The difference can be articulated in the following way: Quine sees the underdetermination as an inherent part of language itself; Foucault sees the indeterminacy as a product of a specific discursive practice, not language itself.

4 The differences in perspective and Foucault's ambition

This brings us to the third part of the paper, the differences between Foucault on one side and Wittgenstein and Quine on the other. I believe I pointed out to certain similarities between the philosophers: the role of rule-following, a non-foundational approach to knowledge, and the interconnectedness of theory and the indeterminacy of reference. Far from the 'radical untranslatability', there are obvious structural analogies. However, in the following part of the paper, it is the causes of differences that interest me.

Hacking and Paltrieri would say that the main difference between Foucault and the other two philosophers is the ahistoricity of the latter. They analyse language-games and theories from an ideal abstract position, not delving into the wealth of empirical material. Following Foucault's terminology, I would sooner call it a matter of perspective. As one of the main principles of the archaeology of knowledge, Foucault puts forward the principle of exteriority. By exteriority, Foucault understands approaching the statement as 'it is said', i.e. as an anonymous, immanent and transparent object. In contrast to that, it is possible to call the analytic approach 'interior', analysing the internal logical and grammatical structure of the discourse.

Another term used by Foucault to characterize his approach is 'positivity'. The positivity of an analysis is a consequence of adopting the exteriority principle; it points to the fact that Foucault's interest lies only in the materially present discourse. He does not bother with the validity of analysed claims, but is interested in the conditions that made it possible for such a claim to be made. His objective is to define the rules that explain the production of given claims in an analysed discourse.

This is the reason why Foucault's analyses are always historical. Firstly, to analyse a certain discourse in its positivity, one must be able to survey enough material to establish the relations. Secondly, and more importantly, one must be able to distance oneself from the analysed discourse. Although this approach allows Foucault to avoid falling into metaphysics,[3] it brings out another problem. An analytic philosopher could dismiss Foucault's analyses as philosophically irrelevant, not bearing any relevance for the philosophical enterprise, since they are dealing with historical and contingent material ('the history of science as a history of mistakes'). For example, Quine's answer could be that Foucault analyses non-sciences, hence his analyses, however historically interesting, cannot have any impact on our understanding of real sciences, i.e. natural sciences. Compared to Wittgenstein, it could be said that Foucault's enterprise is only an application of a pseudo-Wittgensteinian approach to the history of science.

Although Foucault's analyses are far from trivial, Foucault himself took those objections seriously. For that reason, he developed, in his later phase, an alternative approach, which he dubbed 'genealogy'. In it, he tried to explain the discursive formations with regard to their connection with apparatuses (*dispositifs*), i.e. institutions, economic and social factors. In this way, he tried to explain why a certain discursive formation appears in its specific context and form. However, by doing this, he introduced the category of power into his analyses, which, though defendable as non-metaphysical, is partly irrational, and therefore does not have an analytic equivalent.

The point I am trying to make is the following: we have seen that Foucault, Wittgenstein, and Quine operate on the same conceptual level, adopting similar approaches. However, because of the difference in perspective, in Foucault's case, we see that the object of his interest forced him to expand his approach. Even though I do not claim to have a definitive answer, it seems to me that this transition from archaeology to genealogy indicates a certain incapability of a purely linguistic approach to grasp

[3] In this case, the otherwise cumbersome notion of metaphysics denotes the explanation relying on some absolute and fundamental principle, such as 'world spirit' or 'worldview'. The other aspect refers to the notion of an agentive subject, which, according to Foucault, is actually a result of a discourse and cannot be the privileged position of knowledge.

the complexity of social reality – an incapability which can, retrospectively, be ascribed to the other two authors.[4]

5 Conclusion

I started with an analysis of certain aspects of three non-foundational approaches to knowledge and ended up with a question regarding the relation between society and knowledge. This development can be understood as the result of the demands posed to the philosophical task by Foucault himself. Foucault's early approach, which relied on relations between statements, lacked the explanatory power expected from his perspective. While analytic philosophers mentioned in this paper were occupied with questions pertaining to the internal structure of language, Foucault's ambitions regarding human knowledge surpassed the analytic theoretical frame. This case points to the fact that the pre-conceived analytic-continental division turns out to be much blurrier upon closer inspection; however, the question of the nature of the division still remains to be answered.

Reference list

Cusset, F. (2008). *French Theory: How Foucault, Derrida, Deleuze, & Co. Transformed the Intellectual Life of the United States*. Minneapolis & London: University of Minnesota Press.

Deleuze, G. (2006). *Foucault*. Minneapolis: University of Minnesota Press.

Dreyfus, H. and P. Rabinow. (1983). *Michel Foucault: Beyond Hermeneutics and Structuralism*. Chicago: University of Chicago Press.

Foucault, M. (2002). *Archaeology of Knowledge*. New York & London: Routledge.

Foucault, M. (2010). *The Government of Self and Others. Lectures at the College de France 1982-1983*. London: Palgrave Macmillan.

Hacking, I. (1975). *Why Does Language Matter to Philosophy?* Cambridge: Cambridge University Press.

Hacking, I. (2002). "The Archaeology of Michel Foucault." In *Historical Ontology*, Cambridge & London: Harvard University Press, pp. 73-86.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.

[4] Later on in his lectures, trying to reconstruct the historical development of the Kantian critical project, Foucault delineates two strands of post-Kantian philosophy: one concerned with the conditions of the possibility of true knowledge (where he explicitly mentions analytic philosophy as its representative), and the other concerned with the historical present (or the 'ontology of present', 'ontology of modernity', 'of ourselves'), whose representatives are Nietzsche, Marx, the Frankfurt School, and among others, Foucault himself (Foucault, 2010: 20–1).

Paltrinieri, L. (2011). “Pratique et langage chez Wittgenstein et Foucault.” In *Foucault, Wittgenstein: de possible rencontres*, ed. Frédéric Gros and Arnold Davidson, Paris: Kimé, pp. 41-77.

Quine, W. v. O. (1960). *Word and Object*. Cambridge: The MIT Press.

Quine, W. v. O. (1961). “Two Dogmas of Empiricism”. In *From the Logical Point of View*. New York: Harper & Row, pp. 20-64.

Volbers, J. (2015). *Selbsterkenntnis und Lebensform: kritische Subjektivität nach Wittgenstein und Foucault*. Bielefeld: Transcript.

Wittgenstein, L. (1969). *On Certainty*. Oxford: Basil Blackwell.

Wittgenstein, L. (2009). *Philosophical Investigations*. Oxford: Wiley-Blackwell.

Importance of Common Sense in Analytic Philosophy: The Moorean Shift and The Popular View

Mehmet Taylan Cüyaz, Bilkent University, Ali Berk İdil, Middle East Technical University

Abstract

Common sense has always been one of the most central tenets of philosophy after the 'Linguistic Turn'. Many different views of philosophy have stated various positions regarding common sense, thus the definition and the importance of common sense is a matter of dispute. Analytic philosophers tried to determine the scope and the essential core of its constituents. Although their descriptions differ, what they all had agreed upon was the fact that common sense could be used as a fruitful concept to establish an epistemological framework that can cover various fields of modern philosophy. We believe that G.E. Moore's philosophical works are constituting a historical cornerstone for this understanding of common sense as well as the understanding of the analytic tradition. Therefore, in a nutshell, there could be four conclusions that we intend to draw: Moorean thought, thanks to its reconsideration of common sense, has provided an alternative for some popular views; it generally weakened the radical sceptic arguments given against common sense; strengthened its own comprehension of ontology, and also helped fellow analytic philosophers to construct their own understanding of common sense.

Key words: G.E. Moore, common sense, scepticism, analytic philosophy

1 Introduction

For every philosopher it is nearly an obligation to ask whether an external world that is claimed to be the ultimate reality *is* as it is, or whether our 'common sense' is adequate or accurate to make any meaning out of it. These inquiries, without any doubt, have been a central part of philosophy throughout its entirety. Furthermore, the number of people who have at the very end claimed that common sense may not be the ultimate tool through which we can obtain the knowledge of an external world should not be underestimated. Nevertheless, 20th century analytic philosophy, especially with the undeniable influence of the Linguistic Turn[1] (Hacker, 2013),[2] has once again taken on the

[1] *See* also Saha, 1998, for an introductory analysis into Moore's philosophy and *his* linguistic turn.
[2] For a shorter summary *see* Wolf, n.d.

responsibility of investigating the nature of common sense and the plausibility of the propositions regarding the existence of an external world.

Analytic philosophers are of course not the only ones who have undertaken the task of investigating the nature of introductory common sense and the possibility of the existence of an external world. Two other prominent group of philosophers have also started to contemplate upon the issue, namely the pragmatists,[3] such as William James,[4] and the radical sceptics (Speaks, 2004: 1), such as Francis Herbert Bradley and John Ellis McTaggart.[5] Although the two views are too different from one another in terms of philosophical stances and comments, both had a mistrust against common sense knowledge. Therefore, this paper assumes that there is a contrast between the analytic tradition and the other two. Pragmatism, with its rising popularity, claimed that although common sense could be useful in means of developing the practical aspects of life, it is hard to attribute any intrinsic philosophical value to the common sense itself. Hence, common sense becomes instrumental in the view of pragmatism. Moreover, radical scepticism almost always claims that common sense is something to be doubted when it is considered as a tool to understand the world and obtain knowledge about it; hence it was basically concerned with the possibility of common sense being wrong even though it is commonly accepted (Speaks, 2004: 1).

However, analytic philosophers were much more optimistic in terms of trusting common sense as a plausible way of obtaining knowledge. Analytic philosophy has tried to circumvent the two popular beliefs that tried to explain the common sense; however, they did not consider it as a distinct epistemological concept from which true knowledge could be derived. In this paper, we aim to show that the early analytic tradition, despite having some differences amongst its prominent defenders, has carefully and strongly developed a philosophical framework in which common sense not only plays a significant role, but is also justified and supported. The aim is to evaluate and underline the works of G. E. Moore and some other analytic philosophers, to analyse how the proof for the existence of an external world works, and how common sense can have a significant place in philosophy besides being 'popular' and 'useful'.

[3] Moore's attitude on pragmatisim in general and William James's philosophy in specific is presented in his paper "Professor James' 'Philosophy" (Moore, 1907–1908).
[4] For a general survey of James' philosophy, *see* Goodman, 2017.
[5] *See* also Baldwin, 2015.

2 Moore and the Proof of an External World

It is nearly undeniable to think that the existence of an external world as we know it is necessary to make sure that common sense and its knowledge provides us with true knowledge. To see how an external world could exist and how its existence could be proven, one can look at Moore's eloquent and systematic arguments. Moore claims that an external world exists not only because it is an inference to the best explanation, but because he also has a logical proof and justification to infer so:

> I can prove now, for instance, that two human hands exist. How? By holding up my two hands, and saying, as I make a certain gesture with the right hand, 'Here is one hand', and adding, as I make a certain gesture with the left, 'and here is another'. And if, by doing this, I have proved ipso facto the existence of external things, you will all see that I can also do it now in numbers of other ways: there is no need to multiply examples. (Moore, 2013: 165-166)

Moore also continues by claiming that the argument is a rigorous one, which satisfies its conditions. (Moore, 2013: 165-166)

Therefore, in our opinion, Moore does not attack the sceptic's argument by claiming its invalidity; however, he tries to show that the burden of proof is on the sceptics' side. In other words, he claims, in his argumentation, that the sceptics could also be responsible for providing proof in favour of their argument, rather than just casting doubt upon arguments about common sense. Simply put, Moore claims that saying that the existence of two hands is not only a valid argument but also does not lack soundness when compared to the sceptic's argument. What is much more important to understand here is that Moore does not seek to refute sceptical arguments by simply stating that the ordinary experience and its straightforward statements are correct. Rather, he tries to show that the sceptical arguments are self-undermining (Baldwin, 2015: VI). The source of our knowledge, even the knowledge of the sceptics, usually comes from straightforward cases; however, by refuting the straightforward statements, sceptics radically undermine their own philosophical stance (Baldwin, 2015: VI).

3 Moore and Common Sense

As noted by Tom Baldwin, Moore's postulation of common sense as an argument against idealism and some forms of scepticism is first stated in his book, *Some Main Problems of Philosophy* (1910–11) (Baldwin, 2015: VI). However, almost all Moore scholars credit the article entitled "A Defence of Common Sense" (1925) as the most crystalized version of his thoughts. We believe that this article could be analysed in two different fashions. At the first glance, this article is a cardinal criticism against

scepticism or a refutation of idealism that results with different types of the sceptical thesis. This kind of analysis is fruitful for the understanding of how analytic philosophy, in the domain of ontology, makes use of common sense arguments that are directed towards a certain set of arguments. In other words, this paper posits a qualification that is a powerful opposition against the sceptical thesis through the usage of non-technical or ordinary knowledge. However, at a second glance, this article serves as a long-term agenda or a path for some analytic philosophers. This feature of the article is rather subtler, but significant, as one has the chance to indicate some pieces from the history of analytic philosophy and try to demonstrate the Moorean influence.

Firstly, Moore's article starts with a long inventory of propositions that are, labelled by Moore, as 'truisms'. Moore states that, these are

> /…/ at first sight, such obvious truisms that are not to be worth stating: they are in fact, a set of propositions, every one of which /…/ I know, with certainty. (Moore, 2013: 106)

The first part of these truisms consists of a set of propositions that are claims about the subject and its' relation with other subjects and other persons. In a nutshell, we could state the first part of the truisms as follows:

(i) I have an existing body. It has a history. In other words, it existed for a time. It undergoes some changes such as growing up. Also, it has certain relations between some objects that are coexisting with this body. These are three-dimensional objects that are in a certain spatial relationship with me.

(ii) Furthermore, this body also coexists and has certain relations with other bodies. These bodies are also having same conditions. Some of these bodies cease to exist. In other words, they die. In addition to that, the world that we live in, Earth, has existed for a long time, even when this body did not exist. Some bodies lived and died even before the subject was born.

(iii) As an existing human being, the subject has certain experiences. These are experiences that includes perceptions about the self and other objects, observations of facts that are about the circumstances and positions of spatiotemporal things, and awareness of the things that are not instantly perceived, such as past events. Moreover, he/she holds some beliefs regardless of whether these are true or false, and although he/she does not believe in them, he/she has some virtual images such as dreams and so on. Just like the subject, other bodies

> also have the same features, they have experiences of the same sort. (Moore, 2013: 107–108)[6]

After stating this set of propositions, he states a single proposition that "makes an assertion about a complete set of classes of propositions" (Moore, 2013: 106). In short, this proposition can be summarized in the following manner: many human beings know propositions that are "corresponding to each of the propositions" and are about themselves and their bodies (Moore, 2013: 109).

These propositions can be regarded as common sense propositions; as noted by Scott Soames, these are the "core of what Moore called the 'common sense view of the world" (Soames, 2003: 5). For Moore, we know these propositions in certainty; these are wholly true – despite the whole philosophical and linguistic confusion and complexity these propositions have in ordinary meaning (Moore, 2013: 110–111). In other words, these propositions are not *plausibly* refutable, and the terms that propositions include must be understood in an everyday fashion. Moore claims that philosophers could try to refute this claim in two ways: by refuting some common sense claims such as the reality of space, time, matter, and the self, or by refuting all of these (Moore, 2013: 112–115). As noticed by Soames, these claims are not "necessary truths", thus they are also refutable without a contradiction, these are not the popular beliefs or pragmatic standpoints for the philosophical enquiry (Soames, 2003: 6). However, here a question emerges: What is the source of Moore's confidence about his claims? The answer to this question corresponds a bit with the Moorean understanding of common sense. If these common sense claims are wrong, there is no way for a philosopher to hold any consistent beliefs. In addition to that, there is no point to support the sceptical thesis; because what philosophers do when they formulate the sceptical thesis also discredits the thing that they want to achieve. Soames summarizes this argument as follows:

> After all, Moore points out, philosophers live lives that are much like those of other man – lives in which they take for granted all the common sense truths that he does. Moreover, this is evidenced as much in their profession of scepticism as in anything else. In propounding their sceptical doctrines, they address their lectures to other men, publish books they know will be purchased and read, and criticize writings of others. *Moore's point is that doing all this they presuppose that which their sceptical doctrines deny*. (Soames, 2003: 7) [my emphasis]

This quotation clearly shows the very core of Moorean criticism of scepticism. In a Moorean framework, we use words in their ordinary sense to attain a more precise place and use of the common

[6] *See* Soames (2003: 6) for a short and proper summary.

sense argumentation to force scepticism into an inconsistent position that sceptics do not want to hold. As we said before, the first glance of this article gives us the refutation of scepticism by the common sense argumentation.

Still, the notions Moore introduced do have more merits. For instance, we claim that Moore was a pioneer philosopher who provided a realist ontology with minimal elements: his comments on the foundation of the sceptical thesis, which refutes space, time and self, also show us to what one needs to postulate a realist ontology with the least elements, which should ultimately be derived from a common sense perspective. This might again be a good enough reason for the analysis of this article. However, we believe that Moore's arguments also serve as a starting point for a different kind of analytic philosophers. This is mainly because Moore's view on common sense could be analysed as an agenda against a certain way of doing philosophy. As noted by Soames, Moore was against an idealized type of philosophy that was introduced by his teachers; for him, philosophy should *explain* how we know certain propositions, and this project cannot be achievable without giving common sense or ordinary view of our world a cardinal importance. This view certainly influenced some analytic philosophers; however, here we will deal with only three of them, namely Bertrand Russell, Alfred Jules Ayer, and Ludwig Wittgenstein.

4 Influences on Russell and Ayer: A Starting Point for Contemporary Analytic Philosophy

The analytic tradition has obviously been influenced by Moore's work. Aside from the personal influence of Moore on other analytic philosophers, the most famous example being Bertrand Russell; Moore's work had a significant impact upon the direction of modern philosophy. This impact was articulated by Russell with his own words:

> Moore led the way, but I followed closely in his footsteps. /…/ [Our rebellion centered upon] the doctrine that fact is in general independent of experience. Although we were in agreement, I think that we differed as to what most interested us in our new philosophy. (As cited in Irvine, 2015: III)

These words from Russell serve as a testimony of how our contemporary analytic philosophy started its journey. However, it is true that Moore's thoughts on common sense and the external world are not systematic, it is possible to say that his thoughts developed, evolved and became more convincing through time (Baldwin, 2015: VII). It is important to note here that his ideas about common sense and the external world are distinct to a degree, yet complementary. This detail becomes significantly more

essential when it comes to evaluating the role and the place of Moore's work within the history of the analytic tradition. Analytic tradition's ontology benefits a lot from the Moorean understanding of the common sense *and* its proof of the external world. Another prominent analytic philosopher, A. J. Ayer, has developed an ethical theory that is at many points in contrast with Moore's ethical theory.[7] However, Ayerian ethics still stems from the complete elimination of metaphysics and an undeterred praise of empirical data. With this regard, it is nearly impossible for anyone to think that Ayer has formulated his ontological arguments without Moore's influence, and thus his ethical views are also under the influence of Moore's philosophy. This suggests that Moore has been influential within the history of the analytic tradition and even amongst philosophers with views that are different from his.

5 Influence on Wittgenstein: Ordinary Language and Personal Experience

Moore's influence on Wittgensteinian philosophy is beyond all doubt. In *On Certainty* (1969), even though he accepts Moore's account on truisms that was presented in "A Defence of Common Sense"; Wittgenstein deals with some misunderstandings (especially on the meaning of the word 'know') – that was due to Moore (Gibson, 2011:88)[8]. Yet we believe that in *Philosophical Investigations* (1953), some of Wittgenstein's arguments can be associated with an influence from Moore's philosophical agenda – most notably in two positions. First, we can see this influence in the matter of ordinary language and its particular importance for Wittgenstein's philosophy; and second, we can say that there is an agreement between these two philosophers on the nature of sense-datum, which is also a cardinal point of Wittgensteinian philosophy. To begin with, Wittgenstein's therapeutic approach was formulated to cure philosophers and their problematic propositions, which share the same form, i.e. "I don't know my way about" (Wittgenstein, 1953: §123). One should treat the philosophers by treating their grammatical form that they use for postulating new questions. But how can this be possible? Here, Wittgenstein takes an approach similar to Moore. He claims that what one should be doing is "to bring words back from their metaphysical to their everyday use" (Wittgenstein, 1953: §116). This tendency of making use of ordinary language to clear the grounds of language on which the metaphysical theories stand could be credited to Moore. Therewithal, Wittgenstein's thoughts about sense-perception are heavily influenced by Moore's thesis of transparency (Sa Pereira, 2016: 143). In *Philosophical Investigations* he basically argues that pointing a sensation (such as blue) is not something that different from pointing that sensations' phenomenal character itself. In a similar

[7] *See Language, Logic and Truth* (1952)
[8] For a detailed analysis of Wittgenstein's arguments on Moore *see* Salvatore's article "Wittgenstein: Epistemology" on IEP.

fashion, Moore states that "when we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous" (as cited in Sa Pereira, 2016: 144) Thus, we can say that Moore had a profound influence on Wittgenstein's philosophy in both (but not limited to) of these subjects; namely, ordinary language and its power as a tool for correcting philosophical errors, and the content and understanding of sense-datum.

6 Moorean Shift and its Consequences

One of the most important consequences of the Moorean shift is that it has managed to circumvent the popular and the radically sceptical views of the 20th century, not to mention that it has also established a different and an efficient understanding of common sense. It has, to a certain degree, evaded the deadlock created by the radical sceptic arguments and halted the incredibly fast evolving idea of common sense becoming a mere practical benefit and not a solid philosophical concept. Hence, Moorean influence has managed to create an alternative for some of the popular views of the 20th century.

There have been attempts to evade popular beliefs and radical scepticism by enhancing and augmenting the ontology of the analytic tradition. This development of course cannot be attributed to Moore's works only; however, one could easily see its significance in a historical context. Strengthening the argument in favour of the existence of an external world managed to create an atmosphere in which a solid foundation on which the contemporary analytic philosophy could be built upon. In this regard, the paper tries to point out to this situation through its review of the two famous articles from Moore himself. Therefore, Moorean thought strengthened its own ontological position.

In addition to Moore's personal contributions to his argumentations, one can see in the fifth section of the paper that Wittgenstein, one of the most prominent philosophers of the 20th century, took a personal interest in Moore's argumentations and contributed to them. This alone proves that Moore is not an isolated figure within the analytic tradition, but that his philosophy is a dynamic one which has affected the philosophical views. Thence, it can be claimed, with absolute confidence, that Moore has also made a significant progress within the analytic tradition in means of aiding other philosophers, whether directly or indirectly, in creating and advancing their own philosophical argumentations.

To sum up, considering the argumentation given above, there can be four conclusions that we intend to draw. Moorean thought has, to a certain degree, evaded some popular views; in general, weakened the radical sceptic arguments against common sense and the existence of an external world;

strengthened its own ontological comprehension; and helped other analytic philosophers to construct their own understanding of common sense.

Reference list

Baldwin, T. (2015). George Edward Moore. *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved June 11, 2017, from https://plato.stanford.edu/entries/moore/

Gibson, R. (2011). Quine, Wittgenstein and Holism. In A. Orenstein & P. Kotako (Eds.),*Knowledge, Language and Logic: Questions for Quine* (pp. 81-94). Dordrecht: Kluwer Academic.

Goodman, R. (2017). William James. *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved October 1, 2017, from https://plato.stanford.edu/entries/russell/

Hacker, P. (2013). The Linguistic Turn in Analytic Philosophy. In M. Beaney (Ed.), *The Oxford Handbook of The History of Analytic Philosophy* (1st ed., pp. 926-947). Oxford: Oxford University Press.

Irvine, A.D. (2015). Bertrand Russell. *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved June 14, 2017, from https://plato.stanford.edu/entries/russell/

Moore, G. E. (2013). A Defence of Common Sense. In T. Baldwin (Ed.), *G.E. Moore: Selected Writings* (pp. 106-133). London: Routledge.

Moore, G. (1907-1908). "Professor James's 'Pragmatism'." *Proceedings of the Aristotelian Society*, 8, 33-77.

Moore, G. E. (2013). Proof of an External World. In T. Baldwin (Ed.), *G.E. Moore: Selected Writings* (pp. 147-170). London: Routledge.

Saha, D. (1998). Moore's Lingustic Turn. *Indian Philosophy Quartertly*, 25(3), 341-348. Retrieved October 1, 2017.

Sa Pereira, R.H. (2016). "What We Can Learn about Phenomenal Concepts from Wittgenstein's Private Language." *Nordic Wittgenstein Review*, 5(2).

Salvatore, N. C. (n.d.) Wittgenstein: Epistemology. *Internet Encyclopedia of Philosophy*. Retrieved June 14, from http://www.iep.utm.edu/witt-epi/#H1

Soames, S. (2003). *Philosophical Analysis in the Twentieth Century* (Vol. 1). Princeton, NJ: Princeton University Press.

Speaks, J. (2004, September 9). *Moore's Response to Skepticism*. Retrieved October 1, 2017, from https://www3.nd.edu/~jspeaks/courses/mcgill/370/Moore-skepticism.pdf

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell.

Wolf, M. (n.d.). Philosophy of Language. *Internet Encyclopedia of Philosophy*. Retrieved September 30, 2017, from http://www.iep.utm.edu/lang-phi/#SH2c

To Quantify or Not to Quantify: Montague *versus* Quine

Vanja Subotić, University of Belgrade

Abstract

W. V. O. Quine argues in his papers "Reference and Modality" and "Quantifiers and Propositional Attitudes" that it is inappropriate to quantify into referentially opaque contexts. Referentially opaque contexts are linguistic contexts which fail to conform to the principle of intersubstitutivity *salva veritate*. The feature that will be crucial for my paper is the following: referentially opaque contexts are made out of 'that'-clauses. The constituting verb of 'that'-clause is called an opacity verb. Such verb can be taken to express belief, desire, modality, etc., which means that the main reason for opaqueness of modal contexts and propositional attitude reports is the verb in question. According to Quine, the problem of quantifying into can be tackled by restituting the intensions, but, as he proceeds to show, remedy is as bad as the disease. On the contrary, Montague, in his paper "That" proposes a different solution to the problem of quantification: why not get rid of 'that'-clause? Using solely the means of the extensional language of first-order predicate logic, Montague aims to show that one can avoid referential opaqueness if one turns to naming functions for eliminating 'that'-clauses. In this paper, I shall give arguments in favor of Montague's *modus operandi* by showing that parallels can be drawn between Montague's elimination of 'that'-clauses and between the independent subjunctive in Latin.

Key words: quantification into, 'that'-clause, propositional attitude reports, naming function, independent subjunctive

1 Introduction

The aim of this paper is to illustrate an alternative to Quine's argument that it is inappropriate to quantify into opaque contexts made out of 'that'-clauses. I have found Richard Montague's paper "'That'" to be a valuable alternative, given the fact that Montague proposed to syntactically eliminate 'that'-clauses and that such a move gave him an opportunity to completely avoid the problem of opacity.

Since Richard Montague is mostly known for his idea of the unique treatment of natural and artificial languages, I will also try to show that the traces of this revolutionary idea can be found in his previously mentioned paper. My hypothesis is the following:

if it is possible to find at least one small portion of natural language which can be treated in the same way by the means of both natural languages' syntax and artificial languages' syntax, then the very idea of using a formal technique to describe a natural language may be plausible and worthy of notice for philosophers of language and linguists alike.

My paper is composed of the following four chapters. In the first chapter, I shall state the relevant remarks in natural languages' grammar and philosophy of language, and thus elucidate the terminology and the basic issue with opaque contexts. In the second chapter, Quine's argumentation shall be elaborated, whereas Montague's argumentation is the topic of the third chapter. In the fourth chapter, I shall consider the case of the independent subjunctive in Classical Latin, which is arguably a *rara avis* of natural languages' syntax, since it 'omits' 'that'-clauses.

2 Preliminary Grammatical and Terminological Remarks

2.1 On 'that'-clauses in natural languages' grammar

In natural languages' grammar, the most important syntactic 'building blocks' for predication are sentences. Usually, sentences comprise several agent-predicate links. The so-called sentence in the narrow sense comprises only one agent-predicate link, and it is referred to as clause. We can distinguish between two types of clauses: independent and dependent ones. My concern will be just with the dependent or subordinate ones. This type of clauses has the following characteristics: a) such a clause is always part of a longer and more complex sentence, b) such a clause is embedded in a sentence by means of various subordinating conjunctions. *Ipso facto*, 'that'-clauses are subordinate clauses that are embedded in sentences by means of conjunction 'that'. Consider the following examples:[1]

(1) Nero really did *believe that* Rome would be better off without ugly, old architecture which his predecessors had somehow found impressively noble, so he burned it.
(2) The Imperial Praetorian guards *demanded that* stuttering, old and foolish Claudius should become emperor.

Evidently, emphasized verbs in (1) and (2) are constituting verbs of 'that'-clauses, and it is common that these verbs express belief, willingness, desire, doubt, or modality. It is also worth noting that (1) and (2) are declarative sentences – they are conveying certain information.

[1] Unless explicitly indicated by adequate reference, examples are mine.

2.2 On ‘that’-clauses in philosophy of language

If the constituting verb in ‘that’-clause expresses either agent’s belief, doubt, willingness, wish, or attitude towards some kind of action, then the ‘that’-clause, by which the agent’s psychological state is expressed, may be characterized in the philosophy of language as a propositional attitude report. A historical curiosity is that we owe gratitude to Bertrand Russell for the nomenclature of propositional contexts. In his lectures, published under the title *Philosophy of Logical Atomism*, he posed the following question:

> What sort of name shall we give to verbs like 'believe' and 'wish' and so forth? I should be inclined to call them 'propositional verbs'. This is merely a suggested name for convenience, because they are verbs which have the form of relating an object to a proposition /.../ Of course you might call them 'attitudes', but I should not like that because it is a psychological term, and although all the instances in our experience are psychological, there is no reason to suppose that all the verbs I am talking of are psychological /.../ (Russell, 2010: 103)[2]

Another historical curiosity is that the first author in analytic philosophy who noticed the problem with ‘that’-clauses is Gottlob Frege. In his seminal paper, “On Sense and Meaning”, Frege remarked that in the cases of reported speech substitution of the content of ‘that’-clauses cannot be made without violating the principle of substitutivity *salva veritate*. By this principle, two co-referent terms are said to be synonymous if the substitution of one for the other does not change the truth value or meaning of any context in which either term appears. Therefore, on Frege’s account:

> If words are used in the ordinary way, what one intends to speak of is what they mean. It can also happen, however, that one wishes to talk about the words themselves or their sense. /…/ In reported speech one talks about the sense, e.g., of another person’s remarks. It is quite clear that in this way of speaking words do not have their customary meaning /…/ In reported speech, words are used indirectly, or have their indirect meaning. (Frege, 1984: 158-159)

For Frege, the meaning of a sentence is its truth value, and the sense of the sentence is the thought it conveys. Thus, when reporting another person’s remarks, a sentence should not have its customary

[2] Despite Russell’s inclination, in contemporary literature propositional contexts are nearly without exception being referred to as propositional attitude reports (*cf.* McKay & Nelson, 2014). In any case, Russell’s remark is beneficial for Richard Montague, who showed throughout his work that there are many verbs that give rise to opacity when used in certain grammatical constructions, even though these verbs are not at all psychological (*cf.* Partee, 1974).

reference, for it is not relevant whether these remarks coincide in truth conditions with facts; rather, in that case, I am reporting another person's thought about some matter. Because of this, Frege says that "the indirect meaning of words is their customary sense." (Frege, 1984: 159)

Following this stream of thought, Willard Van Orman Quine argues that failure of substitutivity *salva veritate* reveals merely that the occurrence of a term in some statement is not purely referential, that is, that the statement depends not only on the object but on the form of the term by which we name the object in question (Quine, 1953/1961: 140). Contexts in which substitutivity fails are contexts made out of opacity verbs, which express belief, willingness, desire, doubt, or modality, and Quine designated such contexts as "referentially opaque contexts".[3] Linguistic contexts that do conform to the principle of substitutivity *salva veritate* are designated by Quine either as 'referentially transparent', or as 'purely referential'. Taking into account that previously specified verbs (expressing belief, modality, etc.) are forming part of propositional attitude reports, it is clear that propositional attitude reports are referentially opaque. Consider the following examples:

(3) I believe that Nasica was called that way because of his pointy nose.
(4) Publius Cornelius Scipio = Nasica
(5) I believe that Publius Cornelius Scipio was called that way because of his pointy nose.

While (3) and (4) are true respectively, substitution of the term "Nasica" for the term "Publius Cornelius Scipio" does not yield a true sentence necessarily, because I may not know that Nasica and Publius Cornelius Scipio are the same person. Nevertheless, an important difference between Quine and Frege should be noted:

> Failures of substitutivity of identity /…/ were in Frege's view unallowable; so he nominally rectified them by decreeing that when a sentence or term occurs within a construction of propositional attitude or the like it ceases to name a truth-value, class, or individual and comes to name a proposition, attribute, or 'individual concept'. /…/ I make none of these moves. I do not disallow failure of substitutivity, but only take it as evidence of non-referential position; nor do I envisage shifts of reference under opaque constructions. (Quine, 1960: §31)

[3] Barbara Partee, in her paper "Opacity and Scope", distinguishes between two strategies of locating opacity (Partee, 1974: 1-2). According to the first strategy, opacity is a feature of the sentence as a whole when the terms are within the propositional attitudes, and this strategy is Quine's strategy. According to the second strategy, opacity is a feature of various grammatical relations between constituents of the sentence, and this is Montague's strategy.

3 Quine on Quantifying Into Propositional Attitudes

Quine aims to show that belief and other propositional attitude constructions are semantically and syntactically ambiguous. Semantic ambiguity is marked by the difference between the relational sense of the verb and the notional sense of the verb. In Romance languages, such as French and Spanish, such difference may be made obvious by the usage of different moods of either indicative or subjunctive mood in the subordinate clause. Quine takes into account the following examples in Spanish (Quine, 1956/1976: 184):

(6) Procuro un perro que *habla*.

(7) Procuro un perro que *hable*.

The emphasized verb in (6) is used in the indicative mood. In grammar, indicative is characterized as *modus realis*, which means that by using this mood the situation we want to describe is factual. Thus, by uttering (6) I am assuming that, somewhere out there, there is a talking dog which I am desperately seeking at this moment. On the other hand, the emphasized verb in (7) is used in the subjunctive mood. In grammar, subjunctive is characterized as *modus irrealis*, which means that by using this mood, we are merely considering some possible situation which may or may not happen, but the main point is that we have some kind of attitude towards that situation. By uttering (7) I am not committing myself to presupposing whether a talking dog exists somewhere out there, rather I am expressing that I would definitely want to find such dog, for talking dogs surely are a mystery to me. In English, (6) & (7) can be formalized "with some violence both to logic and grammar" by making quantifiers explicit:

(6') $(\exists x)$ (x is dog $\wedge$ x talks $\wedge$ I seek x).
(7') I strive that $(\exists x)$ (x is dog $\wedge$ x talks $\wedge$ I find x).

In (6'), the verb "seek" has, according to Quine, a relational sense, for I am in relation to a particular talking dog. In (7'), on the contrary, the verb "strive" has, according to Quine, a notional sense, for I am in relation to a proposition. This semantic distinction is accompanied by a difference in quantifier scope. I shall analyze the following examples:

(8) $(\exists x)$ (Poirot believes that x is a killer),
(9) Poirot believes that $(\exists x)$ (x is a killer).

In the sentence (8), it is clear that Poirot believes for a specific person that she has homicidal intentions. On the contrary, (9) tells us that Poirot believes that there are killers to be found in the world. Quantifier in (8) is understood as 'passing through' the propositional attitude report, so Quine proposed to call such quantification 'quantification into' (Quine, 1956/1976: 184-185). Quantifier in (9) is within the propositional attitude report, after 'that'-clause, so Quine proposed to call such quantification 'quantification within' (Quine, 1956/1976: 184-185). Quantifier in (8) has a wider scope and binds the free variable inside 'that'-clause. Quantifier in (9) has a narrow scope because its domain lies inside the verb 'believe'. The difference between these two sentences is tightly intertwined with *de re* and *de dicto* distinction. Namely, a sentence is syntactically *de re* just in case it contains a pronoun or free variable within the scope of an opacity verb that is bound by a singular term or quantifier outside the scope of that verb. Otherwise, it is syntactically *de dicto* (McKay & Nelson 2014). Quantification in (8) is problematic, because we can imagine a situation in which it would seem that both sentences are true for Poirot:

(10) Poirot believes that Mister T is a killer.
(11) Poirot does not believe that the man with a bowler hat is a killer.

These sentences give rise to a paradox: Poirot believes and disbelieves for the same man that he is a killer, because Poirot is ignorant of the fact that it is the same man in question. Quine proposes a solution to this paradox: If we understand 'that'-clause as isolating the propositional attitude from the rest of the sentence, then (10) and (11) are not in opposition. But instead of quantifying into as it was the case in (8), we should paraphrase the sentence so we could treat the verb 'believes' as univocal and as standing for a dyadic relation between a believer and an intension, or as standing for a triadic relation between attribute (intension which names 'that-clause'), Poirot (believer), and object (Mister T) (*cf.* Quine, 1956/1976: 186-187). In any case, Quine does not think that any clarity can be gained by restituting the intensions, for as Quine says, "intensions are creatures of darkness" (Quine, 1956/1976: 186). The trouble with intensions is that it is not possible to give rigorous extensional criteria for synonymy between intensions, which entails that they cannot be specified adequatly and thus remain dubious intermediary entities between reference and objects of reference. According to Quine, the key reason for staying clear from intensions is that the alternative is presented in form of truth functions:

> Perhaps the only useful modes of statement composition susceptible to unrestricted quantification are the truth functions. Happily, no other mode of statement composition is needed, at any rate, in mathematics; and mathematics, significantly, is the branch of science whose needs are most clearly understood. (Quine, 1953/1961: 159; cf. 1943: 125)

Let us return to the main issue in Quine's paper: why is the quantification into deemed improper and meaningless? Before I answer the question I will briefly remind you of the difference between objectual and substitutional interpretation of quantifier. Some atomic sentence *A* in first-order language *L* can be read in two different ways: '$\exists x$ (Px)' can mean that either there exists an object which satisfies the condition Px, in which case the quantifier is interpreted objectually; or that there exists an expression x which may be substituted by another expression, say q, and still yield a true sentence, namely Pq would then be equivalent to Px. This is, as you have probably noticed by now, a syntactical counterpart to relational and notional sense. Basically, existential quantifier in (8) is to be understood as objectual, which means that quantification into must be objectual. If that is the case, then if it is a person that (8) tells us Poirot believes to instantiate homicidal intentions, then this must be because there is a relation between these persons. But if this is the force of the quantifier in (8), then that force requires that it must be Mister T who is the value of the variable x. And that, in turn, requires that 'the man in the bowler hat' must also refer to the killer, namely Mister T. Thus quantifying into requires referential transparency. Conversely, referential opacity prohibits quantifying into. The problem is that quantifying into seems indispensable in general, but quantifying into opaque contexts is nonsense. *Summa summarum*, Quine must find a way to insure that the opacity induced by propositional attitudes is confined to a part of a sentence that is not quantified into. As I will render clear in next section of this paper, Montague also shared this idea, but instead of trying to solve the problem by means of intensions, he tried to give a completely extensional solution.

4 Montague's Naming Function for Eliminating 'That'-Clause

Montague proposes his completely extensional treatment of 'that'-clauses in one of his earliest paper – from 1959. In that paper, Montague had explicitly criticized Quine, although Quine never really paid any attention either to Montague's treatment of 'that'-clauses or to his work in quantified modal logic, which is not so strange given his despise towards modal logic (*cf.* Quine, 1943). Nevertheless, that is not the reason to sanction Montague by *damnatio memoriae*.

Firstly I shall deal with his notion of naming function. Naming function as used in arithmetic can be defined in the following way: for every number in a set of natural numbers there is exactly one name such that the ordered pair (*number*, *name*) is contained in the subset defining the naming function *N*. *Simpliciter*, naming function *N* will assign an appropriate name to an appropriate number. According to Montague, Quine's treatment of 'that'-clauses is at least incomplete, because it makes no provision for the joint use of 'that' and quantifiers, and such use is quite common in technical writing (Montague, 1959/1974: 87). As Montague states, the following sentence is an important mathematical fact: *For each prime number x, it is provable in arithmetic that x is prime*. So, by using previously explained naming function and clause "it is provable in arithmetic that", Montague develops his idea by means of the following example (Montague, 1959/1974: 88-89):

(12) N_0 (9) = '9'.
(13) It is provable in arithmetic that N_0 (9) is not a prime number.
(14) The result of substituting the constant '*a*' by N_0 (9) in the proposition '*a* is not a prime number' is provable in arithmetic.

The most important step is from (13) to (14), because in (14) we can see how Montague intends to eliminate the 'that'-clause. But, given the indeterminacy of (13), what if we introduce naming function N_1, which is congruent with N_0 in all instances except when it comes to naming number 9: N_1 (9) ='number of planets'? The sentence (13) is then, strictly speaking ambiguous, because it is not exactly determined which sentence is asserted to be provable in arithmetic. To remove the ambiguity, we should somehow indicate exactly what naming function is to be employed by determining the context in which N_0 will be used. Mind you that not only should the naming function be specified, but it must also be indicated which parts in the clause following 'that' are to be controlled by the naming function. According to Montague, the following formula will be of great help for removing ambiguities (Montague, 1959/1974: 91):

^{x}that *N* (Φ) = the result of substituting all free occurrences of the variable 'a' by *Nx* in Φ.

Now, given the formula, we may reformulate sentence (8) *à la* Montague:

(8*) N_2 (killer) = 'Mister T'.
(∃*x*) (Poirot believes that ^{killer}that N_2 ('*x*=*a*')).
(∃*x*) ('*x*=Mister T' is Poirot's belief).

The opaque occurrences are those standing within the quoted formula which follows 'that'; and the transparent occurrences are those occurring as superscripts to 'that'. Obviously, referential opaqueness must be forbidden in the part that is controlled by naming function N_2, for it seems that any free occurrence of a variable in 'that'-clause should be construed as proper; otherwise vacuous quantifiers will unexpectedly appear (Montague, 1959/1974: 92). In Quine's own terms – it seems that Montague has tried to preserve the notional sense of verb without rejecting 'quantification into', and without restituting intensions. However, it remains to examine whether his analysis departs from common parlance constrained by natural languages' grammar.

5 Final Remarks: The possibility of using a formal technique to describe natural languages' syntax

In the same way that Quine invokes the peculiarities of natural language grammar for the purpose of solidifying the basis for his argumentation in "Quantifiers and Propositional Attitudes", I will invoke peculiarities of Classical Latin for the purpose of showing the virtuosity of Montague in the domain of extensional first-order language. It is worth noticing that I am not trying to simply present an evidence for similarities between natural languages' syntax and artificial languages' syntax. In fact, I shall present evidence in favour of the hypothesis proposed in the introduction of this paper, and while doing so I shall emphasize the 'visionary aspect' of Montague's early paper – arguing against Barbara Partee that the paper has no connection to the project of Montague's 'Universal Grammar' (Partee, 2012: 428).

In the syntax of Classical Latin, we may discern between two types of subjunctives, the dependent and independent subjunctive (*cf.* Ernout & Thomas, 1964: 291-321). Independent subjunctive characterizes the same verbs that are constituting for 'that'-clauses, especially for 'that'-clauses which play role in constructing propositional attitude reports and modal contexts. In reference to an

appropriate verb, there are various types of independent subjunctive. Such a subjunctive can express exhortation or command, concession, wish, question or doubt, or possibility or contingency. The curios thing about the independent subjunctive is that it completely omits 'that'-clause. Interestingly enough, this type of subjunctive cannot be accurately translated into English, but only paraphrased (by using 'that'-clause!): '*maluissem allium obulisses*' can be paraphrased as 'I would like that you smell of leek'.

I believe that I have shown that it is possible to find at least one small portion of a natural language which can be treated in the same way by the means of both natural languages' syntax and artificial languages' syntax. Therefore, similarities between natural languages' syntax and artificial languages' syntax indicate that the same technique employed to create formal languages can be used to describe natural languages in mathematically revealing ways, and even to clarify argumentation in traditionally philosophical problems – which is exactly the vision behind 'Montague's grammar'. Instead of enriching extensional language with formalized parts of the natural language, Montague envisaged an original project of making natural language equally precise. Along these lines, my conclusion is that natural language should not be undermined when dealing with problems in the philosophy of language and that the perspective of comparative syntactic analysis of various languages can make these problems seem even more challenging.

Reference list

Ernout, A. & Thomas, F. (1964). *Syntaxe Latine*. Paris: Librairie C. Klincksieck.

Frege, G. (1984). "On Sense and Meaning" (translation by Black, M.). In: *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell Publisher Ltd, pp. 157-178.

McKay, T. & Nelson, M. (2014). "Propositional Attitude Reports". Retrieved (2017, February 14) from: https://plato.stanford.edu/archives/spr2014/entries/prop-attitude-reports/

Montague & Kalish. (1959/1974). "'That'". *Philosophical Studies*, 10, pp. 54-61. Reprinted in Thomason (ed.), pp. 84-94.

Partee, B. (1974). "Opacity and Scope". In: *Semantics and Philosophy* (ed. by Munitz, M. & Unger, P.). New York: New York University Press, pp. 1-28.

Partee, B. (2012). "Montague's "Linguistic" Work: Motivations, Trajectory, Attitudes". In: *Proceedings of Sinn und Bedeutung* (ed. by Chemla, E; Homer, V. & Winterstein, G.). Paris: École normale supérieure, pp. 427-453.

Quine, W. V. O. (1943). “Notes on Existence and Necessity”. *Journal of Philosophy*, 40, pp. 113-127.

Quine, W. V. O. (1956/1976). “Quantifiers and Propositional Attitudes”. *Journal of Philosophy*, 53. Reprinted in *Ways of Paradox and Other Essays*, Harvard: Harvard University Press, pp. 183-194.

Quine, W. V. O. (1953/1961). “Reference and Modality”. In: *From a Logical Point of View*, New York: Harper Torchbooks, pp. 139-160.

Quine, W. V. O. (2013). *Word and Object*. Mansfield Centre, CT: Martino Publishing.

Russell, B. 2010. *The Philosophy of Logical Atomism*. London: Routledge.

Thomason, R (ed.). 1974. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven & London: Yale University Press.

Linguistic Relativity and Philosophy: From the Beginning to the Analytic Approach

Filippo Batisti, University of Venice

Abstract

The Linguistic Relativity idea, also known as the Sapir-Whorf hypothesis, has lived many lives; at the present day, there is the opportunity to study it following a common trend in several disciplines. Language, as well as cognition, are now conceived as intrinsically social. It is argued that LR effects should be looked for in social realms, and that analytic philosophy has helped in this task.

Key words: Linguistic Relativity; Sapir-Whorf; Philosophy of Language; Psycholinguistics; Extended Mind

1 Introduction

This paper will address a single line of research within the many ways in which the language-thought relationship has been studied, namely the so called *linguistic relativity* idea (LR), also known as the 'Sapir-Whorf hypothesis'.[1] Let us define it in an expanded fashion:

> (1) Linguistic relativity is the idea in accordance to which speakers of different specific varieties of natural languages, which differ in a number of respects studied by linguistics (such as phonetics, syntax, semantics and pragmatics), could experience[2] the same objects and activities of the world (such as but not limited to physical objects perception, colour perception, space relationships, discourse interaction, calculus, shaping of categories, decision making) in different ways, on grounds of that very linguistic diversity – and not because of other factors such as explicit cultural elaboration or cognitive deficiencies or deviations.[3]

[1] This label, albeit popular, should be not be preferred because, as Lee (1996) stated, there is simply no such thing as a 'hypothesis' formulated by Sapir or Whorf, let alone jointly.
[2] Even if the word 'experience' surely rings a phenomenological bell, the intent was to cover a vast number of aspects of human life (see *infra*) with a single term.
[3] See Everett (2013) and Casasanto (2016) for an extensive overview of recent empirical research of various kinds.

More concisely, speakers of two languages that do not have similar linguistic structures in an identified respect could be influenced by this asymmetry in the way they think or experience that respect. As (1) shows, there is a wide variety in the domains supposedly interested by such LR effects.

This paper aims to shed light on the relationship between philosophy and the study of LR in the last two centuries or so, analysing especially the last few years in which the whole branch has gained new vitality in its aims and methods, also – I argue – thanks to analytic philosophy. I mean to do so by a brief overview of on the most interesting paths recently taken by scholars.

2 A brief history of linguistic relativity

2.1 Theoretical premise

Before engaging in sketching a history of the treatment of this line of work, I wish to explain the criteria upon which the following partition has been organised. As a premise, I need to state that I follow the opinion that language sciences, and philosophy of language as well, should not try to treat their object of inquiry as something abstract from its actual use in everyday contexts. Hypostatizing certain features of linguistic structure may have the countereffect to make us stray from the ultimate scope of investigating language itself, namely to understand how and why humans *use* it. The concrete patterns of interaction – and of action in solitude as well, even if language arguably first originated as a tool of communication (Tartabini, 2011; *contra* Humboldt, see Koerner, 2000: 10) – should be the starting point of an enquiry on its functioning, as well as its arrival point. Certainly, conceptual analysis and theoretical knowledge require some degree of abstraction, but especially in psycholinguistic research, the output of scholarly elaboration should, eventually, describe the state of affairs without overlooking any of the actual situations in which language is used by (and among) individuals.

	APPROXIMATE DATE	REPRESENTATIVE SCHOLARS	*LR* EXISTS / RELEVANT	*LR* AFFECTS ACTUAL BEHAVIOUR / ACTION	PHILOSOPHICAL INFLUENCES
1st wave	18-19th cent.	Hamann, Herder, Humboldt	Yes	Yes	Romantic Philosophy
	1890s - 1950s	Boas, Sapir, Whorf	Yes	Yes	(Theosophism - Whorf)
2nd wave	1960s - 1980s	Berlin & Kay, Rosch, Penn	No	No	/
	1990s - 2000s	Lucy, Slobin, Boroditsky, Levinson	(Mostly) Yes	No	/
3rd wave	2000s …	Michael, Enfield, Zinken	Yes	Yes	Analytic Philosophy

Figure 1

Figure 1 summarises the criteria upon which I look at the history of LR: first, a small number of representative scholars from each wave or sub-wave were chosen; then it was noted whether they supported the existence (and relevance) of LR. Thirdly, it was assessed if their approach was consistent with the idea that LR effects influence speakers in their everyday life and not only in artificial settings. Finally, the broad philosophical influences for each (sub-) wave were indicated. Let us now examine each one of them in greater detail.

2.2 The 1st wave: the origins

In the 19th century, German intellectual Humboldt was the first to give some in-depth insights into the relationship between natural languages and the way in which one sees the world (*Weltansicht*). Humboldt wrote that "the world in which we live /…/ is exactly that into which the language we speak transplants us" (Humboldt, 1904: 332) – meaning that every language brings a world-view that (mostly unconsciously) 'mirrors' the way in which language categories "construct the world" (*see* Koerner, 2000: 10). Again, language is seen as something that strongly mediates the external world and the subject that gets to perceive it afterwards:

> /.../ there resides in every language a characteristic *world-view*. As the individual sound stands between man and the object, so the entire language steps in between him and the nature that operates, both inwardly and outwardly, upon him /.../ Man lives primarily with objects, indeed, since feeling and acting in him depend on his presentations, he actually does so exclusively, as language presents them to him. (Humboldt, 1988: 6)

As Koerner has accurately shown, there exists a line of thought that unites German philosophers (Hamann, Herder and Humboldt) and linguists and anthropologists based in North America (Boas, Sapir, Whorf). Oddly enough, each of these scholars was the teacher of the next in line – or at least the two had been in contact for academic reasons. Sapir was the first, in 1924, to use the term

'relativity' to name the 'linguistic relativity hypothesis' as it was popularized by Whorf's papers, who also took advantage of the analogy with Einstein's theory of relativity[4] in physics.[5]

However, in the historical partition that I am trying to sketch, the German–North American circulation stage of LR still falls into the first of the three waves. This is due to the circumstance that the actual implications of linguistic diversity were described in terms of 'action', 'behaviour' and 'habits' - which are alien to the second phase of debate. But first let me clarify what Whorf, as the most prominent representative of the first phase, meant with the aforementioned notions. He wrote:

> [the grammar] of each language is /…/ itself a shaper of ideas, the program and guide for the individual's mental activity, for his analysis of impressions, for his synthesis of his mental stock in trade. (Whorf, 2012: 272)

Furthermore, he adumbrated a definition of

> a new principle of relativity, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic background are similar /…/ (ibid.: 274).

Language helps us 'organize' the world as we perceive it. Then, on the basis of this mental organization, we get an already (at some level) elaborated blueprint for making decisions and acting. As Whorf clarified, he did not "wish to imply that language is the sole or even the leading factor in the types of behaviour mentioned /.../ but that this is simply a coordinate factor along with others" (Lee, 1996: 153). The point I want to make clear is that, in Whorf's view, language-driven perception is something that is linked in a causal chain to behaviour, that is, to action.

2.3 The 2nd wave: Chomskianism and the Whorf renaissance

This last link of the chain had been missing in the LR debate since, *grosso modo*, Whorf's posthumous publications in the 1950s and until the last decade. Thus, phase two began as a consequence of the renovated *milieu* in psycholinguistic research due to the hegemony gained by Chomsky's 'Universal Grammar' theory. Universalist interpretations of the language-thought problem were generally preferred over 'relativistic' ones (Berlin & Kay, 1969; Rosch, 1972). Meanwhile, experimental cognitive psychology procedures and techniques were improved and fine-tuned, so that perceptual domains, conceptualisation, or orienting in space were the dominant themes in LR research. Such a

[4] This claim is consistent with the one made in Footnote 1, as speaking of a 'Sapir-Whorf Hypothesis' entails many factors that are not necessarily true, e.g. that the two had the same view on the matter.
[5] See Zinken (2008) for a repertoire of the various metaphors used in the language-thought debate.

trend had the effect of lowering the interest about the cognitive consequences of linguistic diversity, because if it was nothing but a superficial phenomenon and there existed a cognitive unity of mankind, then LR ought to be false or, in the best case, irrelevant (*see* Penn, 1972: I; or Pinker, 1994: 57).

However, in the early 1990s, LR received new attention thanks to the seminal work by Lucy (1992) and Gumperz & Levinson (1996), who represented a different stream in LR research, as they *confirmed* the effects of LR (e.g. Boroditsky, 2001, on time and space; or Imai & Mazuka, 2003, on objects and substances; Levinson, 2003, on spatial frames of reference; *see* Casasanto, 2008, for an overview). Multiple perspectives were adopted in relation to a range of ontological domains, though 'mostly nonsocial', as linguistic anthropologist Enfield points out:

> research that has been done /…/ has covered only a thin slice of the possible scope of this topic because Neo-Whorfian work has been fairly consistent in its narrow interpretation of the three key concepts. *Reality* has been taken to mean the realm of objective, nonsocial facts: "concepts of 'time,' 'space,' and 'matter'". *Thought* or *mind* has been taken to mean general, nonsocial cognition: forms of categorization, reasoning, and memory about reality as perceived. And *language* has mostly been taken to refer to structural and semantic features, synchronically framed, with a focus on the referential functions of words /…/ Restricting the scope in this way has delivered valuable progress. But it is time to consider the larger space of things that could or should be regarded as instances of linguistic relativity. (Enfield, 2015: 213)

2.4 The 3rd wave: the expansion phase

The last quotation could serve as a manifesto for a new generation of LR researchers as it represents the third phase, which I propose to call the 'expansion phase'. This choice of words is justified by a common trend shared by many different disciplines' recent developments, namely, the *extension* of their object of inquiry – branches such as philosophy of mind, the so-called 4E-cognition in psychology, linguistic anthropology, linguistic pragmatics, or conversation analysis.

More precisely, the focus of their investigation is shifting from the individual, taken 'in isolation', to the individual as an agent who interacts with the environment she happens to inhabit; namely, when she deals both with other people, and with the so-called cognitive artefacts, that is, the artificial devices that influence human cognition (Norman, 1993; Clark, 2003; Heersmink, 2013).

All the approaches of this kind seem to fit well with the theoretical concerns expressed *supra* (section 2.1). In order not to overlook, while doing research, the actual linguistic practices that we are normally engaged in, it is useful to conceive language as a tool that primarily exists for communicating with other humans (Enfield, 2010), and only secondarily for self-improving one's cognitive operations (Everett, 2012). Consequently, a new wave in LR studies could find fruitful suggestions and notions apt to pursue the aforementioned goals. Let us see how.

4E-cognition lies on the assumption that every thought process is not entirely abstract, but is grounded on contextual axes related to the physical bonds on which mental characters are realised (DiFrancesco & Piredda, 2012). Then, interaction with the environment is a factor that contributes to defining the ongoing mental processes. In this perspective, of all the factors that influence cognitive processes, the first should be our body: in fact, low-level processes, such as perceptual and motor ones, seem to be in strict continuity with high-level ones, such as reasoning and cognition in general (Lupyan & Clark 2015). Thus, 4E-cognition employs a *situated* – and not abstract – notion of cognition, which conforms to the faithful picture of psycholinguistic processes sought here.

Linguistic anthropology has also always considered it crucial to study language in ordinary and daily contexts (Everett, 2012; Lupyan, 2012). Since language is intrinsically social (Enfield, 2010), it seems clear that this proposition supports the notion of distributed cognition (Michael, 2002), which serves as a *trait d'union* between research on the functioning of thought and on the nature of language. In this last vein, conversation analysis (Sidnell & Enfield, 2012; Enfield & Sidnell, 2015) falls within those approaches in philosophy of language and language sciences that study ordinary language and all its possible functions. The aim is not to lose the dynamic features which define the actual use of language.

Now my point should be clearer: there are new domains in which LR effects should be looked for. Nonetheless, LR researchers should be informed of the latest trends in all of these disciplines that share their core interests: language, cognition, what kind of relationship links these two elements of human life, and how speaking two different languages can affect this relationship.

This point is not that original *per se*; however, previous attempts (see Enfield 2015: 214) were apparently not adequately followed up among scholars – philosophers in particular. This new frame for the research on LR should bring into play a plurality of disciplines. This is not a simple purpose and, rather, it may seem more perilous than promising. Nonetheless, I argue that it is the matter involved that demands such a complex approach, without which our understanding of the language-

thought problem is bound to remain incomplete. My further claim is that philosophy needs to consolidate its role in this expansion phase.

3 Philosophy in past and present LR research

3.1 New paths in research

Let us illustrate a few examples of (future) LR research that could benefit from a philosophical contribution. Michael's attempt to reformulate LR may be a starting point: his is an example of empiric research that goes beyond the cognitivist paradigm, thanks to two 'theoretical shifts':

> first, from a concern with grammar to a concern with discourse in the context of face-to-face interaction; and second, from individual, isolated cognition, to socially-distributed cognition among a group of individuals. (Michael, 2002: 108)

This new paradigm unwraps many challenges. First, it is recognized that, so far, the conversational approach to culture has been tied to an individualist model of cognition – which should be integrated. Andy Clark's 'Extended Mind' model (Clark & Chalmers, 1998) fulfils this prescription, as it posits that, in Michael's words, cognition is "rarely, if ever, a process bounded by the skull" and "involves interaction with other individuals, and with semiotic artefacts such as texts and maps" (ibid.).

Indeed, according to Clark and others, humans live in a language-permeated environment (Clark, 2003; Steffensen, 2009; Enfield, 2010). This has consequences for their epistemic access to the world, if we acknowledge that cognitive artefacts play a critical role in enhancing cognition (Heersmink, 2016: 78), and that language is "in many ways the ultimate artefact" (Clark, 1997: 218). If we accept that, then language is 'central' for human cognition (Lupyan, 2016), both for the high-level processes of abstract prediction and for the perceptual level, which is "cognitively penetrable" (Lupyan & Clark, 2015).

It is clear that LR studies need to redefine the role of cognition in the light of this different paradigm. Should we look for LR effects in distributed cognition situations? The answer is yes. This kind of collective cognitive processes will at some rate depend on the features of the means allowing such a communicative act. Therefore, the linguistic features of cognitive artefacts could be relevant: different languages may have different feedbacks from the artefacts involved, depending on the quality of the linguistic diversity between the two, thus generating LR effects.

Following Michael's suggestions on the linguistic side of the problem, linguistic interaction should represent the basic scenario in which LR has to be studied. Scholars belonging to the first two waves were concerned with grammatical structures (e.g. Lucy, 1992) and ignored the multiple ways in which

they could have been *used* in linguistic interaction. Language has many more functions than the referential one, which has been privileged for many (contingent) reasons (Enfield 2015: 215). This trend had the result of hiding one of its fundamental traits: language is a *social* tool for *action*, as well as for communication, and for cognition. The distributed approach to cognition then seems a promising frame for investigating the nature of human language. Language turns out to be a no longer isolated or individual tool but a situated and intrinsically social one, given that "human sociality is at the heart of language" (Enfield 2010). In conclusion, experimental research that has subjects abstracted from "real-life contexts" (*ibid.*) in which an everyday social action happens cannot claim to be depicting the actual state of affairs.[6]

For example, it has been shown that in a situated social interaction, different languages may have different effects on the kinds of social actions that can be achieved, thanks to their different linguistic and pragmatic paths to construct the conversational schema. Here, relativity is about "the different rights and duties that speech acts /…/ can give you. /…/ Language-specific side-effects on normative obligations in a next conversational move arise because of the unavoidable introduction of collateral effects when communicative tools have multiple functional features" (Enfield 2015: 218). Speakers of different languages are thus lead "to linguistically relative collateral effects, which lead in turn to differences in our very possibilities for social agency" (Sidnell & Enfield, 2012: 320-21).

Crosslinguistic differences may have dramatic relevance in domains such as heuristics, because decision making is often a less rational process than we may think (Gigerenzer et al., 2011). Since it must be efficient and quick, we rely on simple cues to take decisions, and language sometimes plays a role in this task, because "concepts are sieves" (Enfield 2015: 210) that filter what is brought to our attention. In fact, categorisation is one of the most powerful and frequently exploited functions of language. Categorisation is what builds up the concepts that are the basic units of many everyday actions (Clark, 1998, Diodato, 2015, Enfield, 2015).

3.2 A philosophical stance

Among the notions that have appeared so far, some are of obvious philosophical interest, e.g. 'cognitive artefact', which, according to Heersmink (2016), needs to be better understood from a metaphysical point of view, in addition to the existing literature in analytic philosophy of technology.

[6] See Björk (2008) for a detailed elaboration on LR empiric research and its 'segregation' from real-life contexts.

Obviously, this notion is embedded in the 'Extended Mind' paradigm, which has started one the most relevant discussions in analytic philosophy of mind.

Considering social interaction as a new 'locus' for LR, as Sidnell & Enfield did (2012), the concept of *action* is one that has received philosophical attention since Aristotle, but, in philosophy of language, most notably by Austin (1962). However, this notion needed further revision, at least in the opinion of the authors: the Austinian notion of illocutionary act has been judged insufficient to properly explain how interaction works, as in such a situation, "a person's primary task is to decide how to respond, not to label what someone just did" (Sidnell & Enfield, 2017). That this kind of philosophical-linguistic analysis may be labelled as analytic is argued, among others, by Glock (2008: 54), also considering Searle's and Grice's work.

Another philosophically relevant notion in analytic ontology is 'social reality', as postulated by Searle (2007), according to whom 'social reality' is only created through language and can be approached by linguistic means only. It should be investigated, from a crosslinguistic perspective, if different languages create different social realities, as well as 'social selves', to which different degrees of accountability or, in general, rights and duties, correspond (Enfield, 2015: 216).

To be fair, the very notion of analytic philosophy has not a single univocal nor universally accepted definition. Or, at least, even lengthy attempts at finding strict criteria to define it have somewhat failed (Glock, 2008; *see* Marconi, 2014: II). According to Glock, some of the features that an analytic philosopher *may* have are the willingness to answer substantive questions rather than historical ones following "universally applicable standards of rationality"; the clarity and rigour of argumentation (Beckermann, 2004: 12); adhesion to the linguistic turn; and rejection of speculative metaphysics;[7] just to name a few.

Defining the pure essence of analytic philosophy is clearly too vast of a task for the present scope, if it is possible at all[8] – rather than that a quick comparison between the third wave in LR studies and the previous two will be explicative. Humboldt and the other German romantic philosophers' interest in

[7] Even though since the second half of the 20th century analytic philosophers have expanded their area of interest to other branches of philosophy, including metaphysics. Simons (2013: 709) states that the analytic anti-metaphysicism was not even the case at the beginning of this tradition: "Among those with an outdated or partial conception of analytic philosophy, the whole movement is associated with the rejection of metaphysics. But such rejection, however motivated and justified, was never the sole prerogative of analytic philosophy, nor was it ever the majority view within that movement." In fact, "it was only during the 'middle period' of the 1930s–1950s that, under the influence of logical positivism and ordinary language philosophy, metaphysics was first rejected and later marginalized."
[8] In Pietarinen's words, "Such a task will invariably be frustrating" (Pietarinen 2009).

linguistics was linked to the idea that the 'inner form' of language of a community was an expression of a people's "national mind and unfolding, in line with the Romantic concept of history" (Koerner, 2000: 1). Whorf, for example, thought that every culture "carries with it an implicit metaphysics, a model of the universe, composed of notions and assumptions organized into a harmonious system" (Whorf, 2012: 361), arguing for an "implicit metaphysics", nestled "in the very structure and grammar" of a given language, "as well as being observable in /…/ culture and behavior" (*ibid.*: 75).[9] The minimum standards by which analytic philosophy has been defined here would probably not be satisfied by such claims. Instead, I have argued that the new paths recently taken (or that should be taken) by LR research have much to gain from philosophical debates on the topics that are addressed.

4 Conclusion

Michael, more than 15 years ago, wrote:

> the long-standing controversy over linguistic relativity has been only modestly impacted by two significant developments in our modern understandings of language and cognition – namely, the now commonplace position that both language and cognition are fundamentally interactional and socially-situated practices that cannot be reduced to isolated, abstract knowledge structures. (Michael, 2002: 107)

Unfortunately, little has been done since. Whatever the reasons, both linguistics and philosophy of language have limited themselves in regarding the referential function of language as its core function, giving a biased view of the actual use of language, and directing LR research only on specific trails. This trend has been recently inverted, but there is another force that can foster this change, namely a philosophical analysis of the notions involved in this paradigm shift. Thus, it will be possible to give an increasingly more accurate account of how language works in real-life contexts. Certainly, a multidisciplinary approach is needed, and analytic philosophy appears to be the most appropriate companion.

Reference list

Austin, J. L., (1962). *How to Do Things with Words*. Cambridge, MA: Harvard Univ. Press.

Beckermann, A. (2004). 'Einleitung', in P. Prechtl (Ed.), *Grundbegriffe der Analytischen Philosophie*. Stuttgart: Metzler, pp. 1–12.

[9] Not to mention his eclectic interest in theosophism (*see* Whorf, 2012: 23-25).

Berlin B., Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*, Berkeley: University of California Press.

Björk, I. (2008). *Relativizing Linguistic Relativity: Investigating Underlying Assumptions about Language in the Neo-Whorfian Literature*, Studia Linguistica Upsaliensia 6. Uppsala: Uppsala Univ.

Boroditsky, L. (2001). "Does Language Shape Thought? English and Mandarin Speakers' Conceptions of Time". *Cognitive Psychology*, 43, 1, pp. 1–22.

Casasanto, D. (2008). "Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought". *Language Learning*, 58(suppl. 1), pp. 63-79.

Casasanto, D. (2016). "Linguistic relativity", in Riemer, N. (ed.*) Routledge handbook of semantics*, pp. 158-174, New York: Routledge.

Clark, A., (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, Ma – London: MIT Press/Bradford Books.

Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence, New York:* Oxford University Press.

Clark, A., Chalmers, D. (1998) "The Extended Mind". *Analysis*, vol. 58, 1, pp. 7-19.

DiFrancesco, M., Piredda F. (2012), *La mente estesa*, Milan: Mondadori Università.

Diodato, F. (2015), *Teorie della categorizzazione*, Naples: Liguori.

Enfield, N.J. (2010). *Human Sociality at the Heart of Language*, prolusion, Nijmegen: Radboud University Nijmegen.

Enfield, N.J. (2015). "Linguistic Relativity from Reference to Agency". *Annual Review of Anthropology*, vol. 44, 1, pp. 207–224.

Enfield N.J, Sidnell J. (2015) "Language Structure and Social Agency: Confirming Polar Questions in Conversation". *Linguistics Vanguard,* 1, 1, pp. 131–143.

Everett, C. **(**2013). *Linguistic Relativity. Evidence Across Languages and Cognitive Domains.* Berlin-Boston: De Gruyter Mouton.

Everett, D. L. (2012) *Language: The Cultural Tool.* London: Profile.

Gigerenzer, G., Hertwig, R., Pachur, T. (2011). (Eds.) *Heuristics: The Foundations of Adaptive Behavior*. New York: Oxford Univ. Press.

Glock H.-J., (2008). *What Is Analytic Philosophy?* Cambridge University Press, Cambridge.

Gumperz. J. J., Levinson, S. C. (1996). (Eds), *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Heersmink, R. (2013). "A Taxonomy of Cognitive Artefacts: Function, Information, and Categories", *Review of Philosophy and Psychology, vol.* 4, 3, pp. 465-48.

Heersmink, R. (2016) "The Metaphysics of Cognitive Artefacts". *Philosophical Explorations*, vol. 19, 1, pp. 78-93.

Humboldt, W. v. (1904). "Essai sur les langues du Nouveau Continent". *Gesammelte Schriften* III, pp 300-341.

Humboldt, W. v. (1988) *On Language: The diversity of human language-structure and its influence on the mental development of mankind*, Heath, P. (Ed.), Cambridge: Cambridge Univ. Press.

Imai, M., Mazuka, R. (2003) "Re-evaluation of linguistic relativity: language-specific categories and the role of universal ontological knowledge in the construal of individuation". In: Gentner D., Goldin-Meadow, S., (Eds.) *Language in Mind: Advances in the Issues of Language and Thought.* Cambridge, MA: MIT Press, pp.430–464.

Koerner, E. F. K. (2000) *Towards a 'full pedigree' of the 'Sapir-Whorf hypothesis'. From Locke to Lucy*. In *Explorations in linguistic relativity*. Pütz, M., Verspoor, M. (Eds.), John Benjamins: Amsterdam-Philadelphia, pp. 1- 24.

Levinson, S.C. (2003). *Space in Language and Cognition.* Cambridge: Cambridge University Press.

Lee, P. (1996). *The Whorf Theory Complex: A Critical Reconstruction*. Amsterdam-New York: John Benjamins.

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Lupyan, G. (2012). "What Do Words Do? Toward a Theory of Language-Augmented Thought". *Psychology of Learning and Motivation*, 57, pp. 255–97.

Lupyan, G. (2016) "The Centrality of Language in Human Cognition". *Language Learning*, vol. 66, 3, pp. 516–553.

Lupyan, G., Clark, A. (2015). "Words and the World Predictive Coding and the Language-Perception-Cognition Interface", *Current Directions in Psychological Science*, vol. 24, 4, pp. 279-84.

Marconi, D. (2014). *Il mestiere di pensare*, Torino: Einaudi.

Michael, L. (2002) "Reformulating the Sapir-Whorf hypothesis: Discourse, interaction, and distributed cognition", *Proceedings of the Tenth Annual Symposium about Language and Society—Austin*. Citeseer.

Norman, D. (1993). *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine.* New York: Basic Books.

Penn, J. (1972). *Linguistic Relativity versus Innate Ideas: The Origins of the Sapir-Whorf Hypothesis in German Thought*. Paris: Mouton.

Pietarinen, A. (2009). "Significs and the Origins of Analytic Philosophy". *Journal of the History of Ideas*, 70 (3), pp. 467-490.

Pinker, S. (1994). *The Language Instinct*. New York, NY: Harper Perennial Modern Classics.

Rosch, E. (1972). "Universals in Color Naming and Memory", *Journal of experimental Psychology*, 93, pp. 10-20.

Searle, J. R. (2007). "Social Ontology and the Philosophy of Society". Margolis, E., Laurence, S. (Eds.), *Creations of the Mind*, Oxford, New York: Oxford University Press.

Sidnell, J., Enfield, N. J. (2012). "Language Diversity and Social Action: A Third Locus of Linguistic Relativity". *Current Anthropology*, 53, n. 3, pp. 302–33.

Sidnell, J., Enfield, N. J. (2017). *The Concept of Action*. Cambridge: Cambridge University Press.

Simons, P. (2013). "Metaphysics in Analytic Philosophy", in Beaney, Michael, *The Oxford Handbook of the History of Analytic Philosophy*, pp. 709-728. Oxford: Oxford University Press.

Slobin, D.I. (1987). "Thinking for speaking", *Proceedings of the Annual Meeting of the Berkeley Linguistic Society*, 13, pp. 435–445.

Steffensen, S. V. (2009). "Language, languaging, and the Extended Mind Hypothesis", *Pragmatics & Cognition*, vol. 17, 3, 677-697.

Tartabini, A. (2011). "Aspetti evolutivi della comunicazione animale ed umana". In: Grandi, N. (Ed.), *Dialoghi sulle lingue e sul linguaggio*. Bologna: Pàtron Editore, pp. 127-134.

Whorf, B. L. (2012). *Language, Thought, and Reality*. Cambridge, MA: MIT Press.

Zinken, J. (2008). "The *metaphor* of '*linguistic relativity*'", *History & Philosophy of Psychology*, vol. 10, 2, pp. 1-10.

Words and Meaning

Dorijan Dobrić, University of Belgrade

Abstract

One of the main goals of philosophy of language is to secure the point of having utter and undisputed certainty that some word means exactly what it is needed from it to mean. Furthermore, philosophy of language seeks uniformity: that everyone agrees on what that particular word means, and that it means exactly that, and not something else. The main fear of any philosopher or even scientist is that we do not understand each other, and that is where all the confusion and disagreement emerges from: we are playing with words instead of using them carefully. The ideal picture of human language to a logician is that the relation between word, sense, and object can and must be 1:1:1, which means that in that case, we all agree and truly know what one means when one says, e.g. *butterfly*. On the other hand, the ideal picture of human language to a linguist is variety, since it is the nature of human language itself. It is well known that logic and linguistics, in general, have different approaches to the problem of meaning. By using Frege's idea of *sense* and *reference* (more precisely, using his model) and thus giving new interpretations of his theory, we will try to combine these two approaches and give possible prospects of solving certain issues in philosophy of language that concern meaning.

According to Frege's idea, there are *name*, *sense*, and *bedeutung*, which will be, in this paper, translated from the German original text as *nominatum*, which will here be used as 'what is named'. In this paper, we will argue that the distinctions between names and words in general is obsolete, since words are articulated noises that can be written. We will try to point out that it isn't of any importance for the topic discussed whether we use the proper name *James*, or the noun *butterfly*, or literally any other word in any language.

Furthermore, to avoid certain problems, we will, for the sake of the topic, use *sense* as *meaning*, even though there is a difference between these two, especially for native English speakers. We will use them here as two equivalent words and the reason for that remains to be explained. We will argue that words can have a *possible infinite set* of meanings when used as a single word, extracted from the reality of a sentence (e.g. word in a dictionary), and that when used in a sentence, they have a *finite set of meanings*, and that adding more meanings (or words which carry them) helps with finding *nominatum*.

Finally, we will try to prove idea that *object* (nominatum) in language can be anything, that for language and exclusively for language, it is not important whether object is some abstract concept such as *good, existence, god, beautiful*, or any empirically available object, such as *butterfly*, *Pink Floyd's song "Money", Maria's notebook, Sagrada Familia*, etc.

It is important to emphasize that I am only using Frege's model, which is found to be very convenient for given ideas, and that implications of these ideas can be in conflict with Frege's results.

Key words: word, meaning, nominatum, awareness, language

1 Introduction

The goal of this paper is straightforward and simple: to refute the idea that some word, or even proposition, can have universal and undisputed meaning that is agreed upon by the entire community. We usually represent this as the ideal picture of human language; to have some sort of undisputed meanings of words which could represent the frame of reference and a sort of 'guide' for other, problematic meanings. As it is said, the goal of this paper is to refute that that we are able to find those undisputed meanings and to defend, in lack of better description, the position of the relativity of meaning. The method of this paper is not strictly analytical; as it is promised in abstract of this paper. We will combine two different approaches, because using *only* logic in discussing meaning and language renders the given discussion poor and always incomplete; which will be discussed further in the text. However, any scientific work without using logic results in a lack of dignity of given research and its validity becomes immediately questionable.

After discussing the previously mentioned subject, we will continue with a possible refutation of the statement that language entirely determines our existence and our apprehension of reality.

A lot of modern, prominent thinkers and philosophers were ready to reduce the entire philosophical inquiry on problems of language and meaning, which is, in my opinion, very wrong, and that is largely due to the great impact of the English language on world, not just philosophy – since it is a language which is 'contaminated' with nominalism. Greek, German, and Latin are the languages that are capable of grasping abstract concepts more effectively than the English language, as Nikola Tanasić once stated. And, if the reader would allow some emotion from the author of this paper: it is even demeaning for philosophy. I am not saying that language is not problematic or that many difficulties that history of philosophy had did not emerge from misunderstanding; history of philosophy suffered several times in its change of the vocabulary that was used – Hegel, for example, has very much different

understatements that he uses in his philosophy when he writes about the *abstract*; his meaning of the word *abstract* differs from what, for example, Kant or any other philosopher thought. Here, different usage of philosophical terms will not be explored, it is just important to emphasize that the problem of meaning is, and always has been a problem for philosophy, but it is not the *only* problem, similarly to how there are many different things in human life that determine human existence besides language. This will be a discussed further in the last subsection of this paper, under the subtitle *Conclusions*.

However, for modesty's sake, in this paper only the first goal that we have stated will be discussed and hopefully proven: that singular word, singular proposition, and even a singular sentence, cannot have undisputed, universal, and singular meaning.

2

By using Frege's model, which was presented in his famous work, *On Sense and Reference*, I will present my insights, opinions and the groundwork concepts of philosophy of language; it is important to state that I will be only using Frege's model and nothing else. My results can be different from Frege's, but that does not change the fact that his model is very convenient when approaching problems of meaning. There are three pillars of every word that we use in every language – those are *word*, *meaning*, and *nominatum*. Each will be discussed in a different subsection.

2.1 Words

What are words? Defining them seems difficult, but after contemplation, the answer to the given question that I came up with and that I found the most satisfying is the following one: *words are articulated sounds*. This, and nothing more. This definition is so simple and truthful that we could probably put an equals sign between *word* and *articulated noise*. That is, amongst other things, what differs us from animals, they create noises and thus communicate, while humans make noises, they also communicate, only humans do it with intentional articulation. Words have no connection with either meaning or *nominatum*, at least there isn't any ontological connection or necessary connection, only the connection given by the community in a certain time and place. Here is one example which supports the given claim. Ferdinand de Saussure in his *Course in General Linguistics* pointed out the flexibility of words in French language and how they can change through time. In the 11th century, the ancestors of modern French people used 'rei' (as it was both pronounced and written) for the word 'king'; in the 13th it was pronounced as 'roi' and was written as 'roi' (De Saussure, 1969: 38). Then things started to get a little bit more difficult when the discrepancy between writing and speaking

started to appear (and we all know how famous the French language is for that discrepancy). In the 14th century, 'king' was pronounced 'roe' and written 'roi' (De Saussure, 1969: 39), and finally in the 19th century, which is the current state of the word 'king', it is pronounced 'rwa' and written as 'roi' (De Saussure, 1969: 38). We can see that written language endures the time much better than spoken language and it is harder to change it, which is, in my opinion, healthy for preserving a language. The reason why I presented this example is to show how words and their sounds always depend on the community at certain point in time. Moreover, words are connected with their meaning only through social agreement, which means that in the future the French can start using the word 'roi' to not represent 'king', but define 'law' or 'presidency'. "How come?" one might ask. It is quite simple. We do not alter the meaning or the articulation intentionally, we do that through time. As the events in history change, the society changes, and so do its words, their sounds, and their meanings. Let me illustrate this with another example. In Serbia, there was a very famous brand of calculators called 'Digitron'. They produced calculators for the entire country. The brand became so famous that the word *Digitron* was later used for *calculator*. And the word *digitron* was only the invented name for the brand, which was based on the connection of the product with the concept *digital*, nothing more. If this is possible, (to completely change the word in one community) then this scenario is also possible: Alexander the Great had this famous horse called 'Bucephalus'. If we keep in mind the '*digitron* event' that occurred in Serbia, why would it impossible for English people to start using the word 'Bucephalus' for the word *horse*? Or even for Greeks to start using 'Bucephalus' instead of 'álogo', which is their current word for *horse*. The way the words sound in any language, and their meaning, completely depends on the social agreement. The possible counterexamples for this statement are those words that resemble the sound we use to pronounce them, such as 'round' in English or 'oblo'[1] in Serbian. When we say 'oblo', the way our speaking apparatus moves and behaves can resemble a round object; even the roundness of the letter *o* reminds us of the shape of our mouths. This is rendered completely unimportant once we realize that this 'likeness' of sound and object is completely arbitrary; it may happen in English or Serbian language, but in Chinese none of its logograms resemble any object; furthermore, the sound of the word 'round' in Chinese doesn't resemble the appearance of a round object. Possible counterexamples for my statement are also onomatopoeic words. However, every language differs in the interpretation of the sound of trees, animals, etc. For example, for the sound that a rooster makes in the morning, Serbs use 'kukuriku!', whereas in some parts of Spain

[1] Which is a Serbian word for *round*.

‘kikiriki! is used. It is similar, but not the same. It is different enough that it cannot become a general rule that the whole world equally ‘hears’ and transcribes the sounds of nature.

To say once again what was the point of this subsection – it was to show that the sound of words completely depends on the social agreement. This follows the next statement – that the meaning of every word in any language also completely depends on the society; this will be shown in the next subsection.

2.2 Meaning

The Roman verb for ‘to awake’ is ‘excito’. This word, just like many other words, came into the English language, but its meaning changed. Now that word is used for *excitement*, to feel *thrilled*, etc. We have in this an example of how language is used in different societies – completely arbitrarily. Meaning also depends on which language community we live in. We borrow words from other languages and we alter their meaning in the way that we find suitable for us. Furthermore, that means that *any word can mean anything*, depending on the community. When a word is created, there is no general rule for how it will be used, or the strict boundaries of how large a domain of a certain word is, and there is no point in searching for one specific meaning of any word, because there is none; there are only common usages of words, some of them less common, and some of them more common. One cannot say, “This x means this and exactly that and nothing more.” That’s why I divided words in two different sorts, based on their *quantity* of possible meanings, namely, there is: 1.) *Infinite set* of possible meanings, and: 2.) *Finite set* of possible meanings.

2.2.1. Infinite set of possible meanings

Words, when extracted from linguistic reality, as they are used in the dictionary – for example, when one says ‘*butterfly*’ – can possibly mean anything. Words in dictionary are completely extracted from reality and are not used properly, because they lack linguistic reality. So when one asks, “What does ‘*butterfly*’ mean?”, what one is actually given as the answer is an infinite set of meanings, which is frankly impossible. Of course, it would be wrong to respond with, “’*Butterfly*’ is butter which flies.”, just like it would be wrong to respond with, “’*Butterfly*’ is an insect that flies and is pretty.”, since there are many other insects that can fly and are possibly pretty. To grasp the domain of what ‘*butterfly*’ can mean, one must first be an expert in entomology, one must have clear and undisputed reasons according to which one can make a distinction between butterflies and other flying insects. That would be too much to ask from one person, and even this would only just be the beginning of giving a word

its meaning. What about the symbolism that can derived from the word '*butterfly*'? The change, the beauty of that change – butterfly can be used poetically in various ways, and that is also the domain of the meaning of the word "butterfly". This means that I completely subscribe to the notion given by Wittgenstein, "Don't ask me for meaning, ask me for use." (Ryle, 1957)

When it comes to grasping the meaning of a certain word, things get a little bit easier once they are put in linguistic reality – we understand each other because when we communicate, we count on a certain finite set of possible meanings for every word.

2.2.2. Finite set of possible meanings

Let's have a look at the following sentence: "James is a very lucky man." Anyone can be named *James*, which isn't a problem, since participants in the conversation know who *James* was, otherwise the sentence wouldn't be understood and would require further explanation – and that's how the infinite possibilities of who can James be, or what *James* means here, are reduced to the ones who are named *James* and who are known to the participants.[2]

Things get a little bit more difficult when it comes to *verbs*. Here we have 'is' as the form of the *to be*. Verbs are the words that have meaning, but have no *nominatum*, because they serve to describe actions, such as *running*, *killing*, *worshiping*, *dying*, etc. They are specific, since they serve only humans – rocks, mountains, rivers, planets and everything else in the universe would exist even without naming them in different languages, but verbs serve only to describe different actions. That doesn't mean that, for example, that running wouldn't exist as a concept, animals would still run, but they don't have concept of running, since they change their speed of movement because they are forced by their instinct – either to run or to hunt. Verbs are probably the most abstract things that thrive in the human language and are one of the things that can carry the title of a pure creation of human beings. Language is used to describe reality, but verbs are the purest means of describing reality. I will not engage further in exploring the nature of verbs, since the subject requires and deserves a completely independent topic and it will be left for another paper. For now, let's continue with meaning.

What will happen if we add more words to the given sentence? "James is a very lucky man."; "James is a very lucky, blue-eyed man, born in Amsterdam."; James is a very smart, lucky, blue-eyed man, born in Amsterdam."; etc. Nothing specific, since we already know the *nominatum*, *James*, who is

[2] Note here that the meaning of *James* is the same as the *nominatum*, which usually happens with names. Whether that is always case or not is left for another inquiry.

known by the participants of the dialogue – by adding more words we can only help with the description and with nothing more. That leads to another example: in the case of '*author of Crime and Punishment*' and '*Fyodor Dostoyevsky*', we don't have a problem with *nominatum*, since the *nominatum* is the same, but we cannot put the equals sign between them, since *Fyodor Dostoyevsky* is much more than just the '*author of Crime and Punishment*'. For example he is also the author of *The Gambler*. '*Author of Crime and Punishment*' is just one of the descriptions of '*Fyodor Dostoyevsky*'.

In case that this wasn't clear enough, I will stress out the conclusions of the last two subsections. In an unreal scenario (namely in a dictionary), we have words extracted from reality, which can have an infinite set of possible meanings. Once the words are used in reality, in a sentence, they are reduced to a finite set of possible meanings, because, when we talk, we count on the meanings that we are used to and that are established by our language community. When we do not understand each other, this is either because we do not speak the same language, or because we do not have that one popular meaning of a given word in mind.

3 Nominatum

Nominatum can be anything: an abstract concept, or an empirically available object, it doesn't matter. *Nominatum* can be *time, the sound of glass breaking, a wall, street*, etc. The proof for this idea is the following: the human mind is always singular. Let's have a look at the abstract concept of *time*. Time is something that transcends singular, and that is the reason why it is so hard to understand what time is and why the notion of time is different in different language communities. Western civilizations perceive time differently from some Indian tribes in Northern America – for them, space and time are *linguistically* inseparable concepts, while us Europeans can make a distinction between the two. Whenever one asks about time and thinks about time, every thought and every sentence is always singular. The notion is always singular and not the quantity of a given thought. In language we have a concept of plurality, for example, we can think of *cows*, or *cups*, but that thought is always singular. This is, and I cannot emphasize this enough, the underlying problem of language in philosophy. That is why when we write texts or books we collect every information, because every thought in every given word, proposition and sentence is always singular – we, as humans cannot think *whole*, only partially. Every act of mind is singular; when one says *time*, one can think of a grandmother's old clock, or Einstein's theory of relativity, one can think of Dali's *Persistence of Memory*, or anything that can one be reminded of when one thinks of *time.* That is why we cannot grasp abstract concepts, because we only think in singular and never wholesome, and abstract concepts are always *whole*.

4 Conclusions

I understand that I opened a much larger number of problems, but than I actually solved a certain number of them, yet I believe that I've shown several things. First of all, sounds of words and meanings completely thrive on our social agreement. When it comes to *nominatum*, it can be anything, and it was wrong from Frege to even consider putting examples like '*Odysseus*' or '*unicorn*' away from having a *nominatum*. Everything that constructs our reality can have a *nominatum*, even '*unicorn*', since it is a part of our created reality. Does a '*unicorn*' exist? Not in nature, but it exists in our reality, in books, in drawings – it is part of our reality insofar as it is found in ourselves.

I would like to finish this paper with one last conclusion: language is a strong, powerful means that determines our reality, and a very large portion of it. That does not mean that it is the only thing that describes our reality and that is the implication of the statement given in the previous subsection – that our minds are always singular. In the case of time, or any other abstract concept, one is always aware that a given concept is always something more than it is in that one singular thought. That subject's awareness is incomprehensible, a kind of shady knowledge (but still knowledge) that there is something more about the given abstract concept than it is given in that singular statement, word, or even the entire sentence.

Reference list

Ryle, G. (1957). "The Theory of Meaning". In Muirhead, J. H. (Ed.), *British Philosophy in the Mid-Century*. George Allen and Unwin, pp. 239–264.

Saussure, F., Bally, C., Sechehaye, A., & Riedlinger, A. (1969). *Course in General Linguistics*. New York: McGraw-Hill.

Conceptual Analysis, Definitions, Wittgenstein

Petar Srdanović, University of Belgrade

Abstract

The basic hypothesis of this article is that conceptual analysis represents an important method of inquiry in analytic philosophy and, often, its distinctive mark. My main goal is to show that the practice of conceptual analysis should look different than it actually does. In order to fulfil that aim, I have divided this paper into four parts (besides Introduction and Conclusion).

The first part is a short history of conceptual analysis. I shall adduce examples from Plato's *Euthyphro*, Kant's *Critique of Pure Reason*, and Gettier's "Is Justified True Belief Knowledge". Then I argue how conceptual analysis has been understood differently throughout the history of philosophy. However, it seems that there is one non-trivial similarity between the mentioned views: they all understand conceptual analysis as a search for definitions.

The second part shows that seeking for definitions faces one serious problem: the Plato–Weitz's problem. The problem states that (1) we did not succeed in formulating definitions for most of our concepts, *and* (2) for some concepts, it is impossible to provide a definition. (2) comes directly from Wittgenstein's later philosophy.

In the third part, I will try to illustrate the Plato–Weitz's problem; namely, I will present the attempts, made by some philosophers of art, to define art and then explain how Weitz criticized them. Their definitions were inadequate, because they were either too broad or too narrow.

Finally, in the fourth part, I will present an alternative to seeking for definitions. That alternative does not abandon conceptual analysis; it only modifies it. The alternative is the so-called Prototype Theory, which says that, during the practice of conceptual analysis, we should focus on finding the properties that items which fall into the extension of analysed concept *tend* to have. Interestingly, the Prototype Theory is also rooted in Wittgenstein's later philosophy.

Key words: analytic philosophy, conceptual analysis, definitions, Wittgenstein, Weitz

1 Introduction

It would be an exceptionally hard task to precisely determine the meaning of 'analytic philosophy'. Philosophers and historians of philosophy are, however, able to avoid that problem by explaining 'analytic philosophy' either by using the term 'conceptual analysis' or through comparing analytic philosophy with various philosophical schools, such as British Idealism or continental philosophy.[1] I find the first approach (i.e. using the term 'conceptual analysis') to be more attractive. In my opinion, it helps us to understand what many analytic philosophers do and how they do it, as well as what are they doing wrong and how they could improve their work. Since this paper's aims are both, the reconstruction of the practice of analytic philosophy and the critique of that practice, I shall begin with some preliminary remarks on conceptual analysis.

2 Conceptual analysis

Conceptual analysis is, simply speaking, a philosophical method of inquiry, which aims to clarify the meanings of various concepts, such as 'knowledge', 'art', 'good', 'piety', etc.[2] However, even a brief look at the history of conceptual analysis suggests that philosophers used that method in different ways. In this paragraph, I will argue that, despite all the differences that those approaches have, there is one non-trivial similarity among them. In order to fulfil that aim, I will adduce some examples from the history of philosophy.

Perhaps the first philosopher who used some sort of conceptual analysis as a tool for solving various philosophical problems is Socrates, whose work is mostly preserved in Plato's dialogues. There we can read about a famous Greek potter, who annoys his interlocutors with questions such as 'What is justice?', 'What is beauty?', and so on. Conceptual analysis in Plato's dialogues begins with those questions, continues with dialectical refutation of various answers to questions which Socrates asks, and ends either aporetically or with precise determination of the analysed concept's meaning (justice, beauty, etc.). So, according to Plato, a valid conceptual analysis demands two persons, a conversation,

[1] These two approaches are not mutually exclusive; namely, one could argue that what differentiates analytic philosophy from other philosophical schools is nothing else than conceptual analysis. For further discussion of the role of conceptual analysis in analytic philosophy, see Beaney (2016).
[2] This determination of the term 'conceptual analysis' is good enough for this paper's purposes; although, I admit it is probably not a perfect one.

and intuitions, as well as ordinary language as a criterion for the refutation or confirmation of given answers.

We can find a good illustration of Plato's use of conceptual analysis in his early dialogue *Euthyphro.*[3] In this dialogue, Plato's Socrates meets Euthyphro, a young man who has decided to sue his own father for manslaughter. Euthyphro justifies his act through the claim that a pious man must not ever be blind to manslaughter. This encourages Socrates to ask his interlocutor what piety is – a question seemingly simple and easy. However, it will turn out that not a single one of Euthyphro's answers is actually a good one, for piety could not be what is pleasing to the gods, since gods themselves could take different things as 'pleasing'; and piety is also not simply what *all* gods love, because piety precedes its liking – something is pious, gods recognize it, and only then they start loving it; and lastly, the answer that piety is an instance of justice is insufficiently precise, since it is not clear what differentiates pious action from other actions which are considered just. Therefore, Socrates concludes that their discussion did not clarify the meaning of 'piety'.

The next philosopher I want to mention is Immanuel Kant. Kant (trans. 1999: 157) first poses questions such as 'What is space?', 'What is time?', etc. But, unlike Plato's Socrates, he prefers the means of logic and *a priori* reasoning rather than dialogues and everyday language. Considering the mentioned case of space, Kant first offers three possible answers to the question 'What is space?': Newton's, Leibniz's and his own. After that, he aims to show, consistently using *a priori* reasoning, that we should accept his answer. I do not intend to go into details of Kant's argument, since they are not significant for this paper. All I wanted to point out is that Kant's analytical approach to concepts such as space and time noticeably differs from Plato's Socrates' approach to concepts such as piety or justice. Kant's approach, as mentioned, does not require two persons (i.e. dialogue) and intuitions from ordinary language. I would say that the difference is clear and unquestionable; although, one could argue, perhaps even with strong evidence, that those differences arise solely from different philosophical aims and different subject of inquiry.

Lastly, I shall focus on the practice of conceptual analysis in contemporary analytic philosophy. Many contemporary analytic philosophers write about necessary and sufficient conditions for the successful use of X, where X represents a concept. For example, Edmund Gettier (1963: 121-23) argues that knowledge is not a justified true belief, because that determination does not express sufficient

[3] The following subparagraph is my brief and simplified review of dialogue *Euthyphro.* The reading of the whole dialogue is highly recommended. See Plato (trans. 2002).

conditions for the use of the concept of 'knowledge'. Maybe it expresses necessary conditions, but it does not express the sufficient ones. I would say that this almost permanent insistence on the terms 'necessary condition' and 'sufficient condition' represents one of the most distinctive features of today's practice of conceptual analysis. Although it is not its only distinctive feature and perhaps not even the most significant one, I chose to mention it for one particularly important reason: it will be useful for the illustration of the problem which has followed philosophy since Plato and which got a new dimension in the 1950s. But before that, I must say that there is one non-trivial similarity among all presented types of conceptual analysis; namely, all could be perceived as *seeking for a definition;* Plato, Kant, and Gettier would agree about one thing – a meaning of a concept is clarified when we can formulate a definition of that concept.

3 Plato – Weitz's problem

In the previous paragraph, I have inferred that many philosophers, including many analytic philosophers, use conceptual analysis as a tool which provides definitions, i.e. precise determinations of the meaning of various concepts. In analytic philosophy, those precise determinations include listing all necessary and jointly sufficient conditions for the adequate use of the concept being analysed. I believe that this approach faces one serious problem, which I call the Plato–Weitz's problem. This problem has two aspects. Now I shall try to explain each of them.

The first aspect is often marked as Plato's problem.[4] It can be formulated in the following way: "Perhaps the most basic problem that has been leveled against the Classical Theory is that, for most concepts, there simply aren't any definitions." (Laurence and Margolis, 1999: 14). Indeed, philosophers still do not have a generally accepted definition of concepts such as 'knowledge' or 'justice'. Of course, the very fact that they do not have it does not imply that seeking for definitions has not contributed to philosophy so far. It probably has some heuristic value even today, but it did not help us a lot to solve the majority of core philosophical problems. However, Plato's problem is not the ultimate trouble for all who try to formulate definitions of some concepts, since they can simply argue that definitions are possible but rarely formulated because of the very nature of philosophy and its complexity. In other words, one hundred unsuccessful attempts to define, for example, beauty, do not imply that the one hundred and first will also be unsuccessful – definitions will come, sooner or later.

[4] For example, Margolis and Laurence mark that problem as Plato's problem. See Margolis and Laurence (1999).

Unfortunately, the second aspect of the Plato–Weitz's problem gives reasons for doubting whether definitions of concepts (or at least, definitions of *some* concepts) are possible at all. Namely, Morris Weitz (1956) argued that some concepts are 'open' concepts. It means that the meaning of those concepts is changeable and, maybe even more importantly, that entities or activities which fall into the extent of those concepts simply do not share the features which are sufficient for the precise determination of the meaning of that concept, i.e. for a definition of that concept. All this could sound fuzzy. It will become clearer in the next chapter, where I intend to illustrate both aspects of the Plato–Weitz's problem. What I should emphasize in this chapter is, however, the fact that Weitz's insight that some concepts are 'open' is actually based on Wittgenstein's theory of family resemblances from his later book *Philosophical Investigations*.

Wittgenstein (1986) argues, *inter alia*, that we cannot list all necessary and jointly sufficient conditions for the use of the concept of game, since there is nothing non-trivial that all examples of game share. What we can do is recognize some family resemblances between those examples of game. For example, chess resembles poker, because they both require thinking, they are both 'table games', they are both games for more than one player, etc. On the other hand, poker resembles betting on roulette, because they are both played for money. But in what way does chess resemble roulette? They are both table games, but roulette does not really require thinking, since it is mostly about luck. Now we are already forced to admit that many properties which games could have, such as 'requires thinking' or 'is for more than one player', are not necessary for the use of the concept of game. There are simply some games which do not possess those properties. And we have listed only three types of games! Instead of an endless and perhaps futile search for properties shared by all games, why would we not simply accept that the concept of game is such that its meaning is not as firmly fixed as definitions suggest yet still not arbitrary, since it is constrained by entities which are represented by that concept?

Hence, the Plato–Weitz's problem basically says that (1) for most of our everyday concepts, we did not succeed in formulating a valid, satisfying definition of them, *and* (2) for at least some of our everyday concepts, it is impossible to formulate a valid, satisfying definition. The reason I call this problem the Plato–Weitz's problem and not Plato–Wittgenstein's problem comes from the fact that, in my opinion, Weitz is the first philosopher who truly realized the power of Wittgenstein's theory of family resemblances. Of course, one could argue that Wittgenstein himself was aware of that too. That claim, however, is at least partially based on what kind of interpretation of *Philosophical Investigation* we accept and apply. In the next paragraph, I will make an illustration of the Plato–Weitz's problem.

It comes from the contemporary philosophy of art, and I shall hopefully make everything I have said so far in this paragraph clearer.

4 The illustration of Plato – Weitz's problem: defining art

Philosophical interest in art is almost as old as philosophy itself. One of the most significant question in the field of philosophy of art is certainly the question 'What is art?'. Plenty of definitions have been offered. I will briefly list three of them, which were popular during the first half of the 20th century. After that, I will present Weitz's critique of those attempts, which directs us, in my opinion inevitably, to the conclusion that neither of them contains both necessary and jointly sufficient conditions for the use of the concept of 'art'. Finally, at the end of this paragraph, Weitz's reasons for the belief that the definition of art is impossible will be illustrated.

Three attempts to define art. The first attempt I want to point out is the so-called Emotionalist theory. It says that art is "…the expression of emotion in some sensuous public medium." (Weitz, 1956: 28). The second one is the Intuitionist theory. According to intuitionists, art is "…identified not with some physical, public object but with a specific creative, cognitive and spiritual act." (Weitz, 1956: 28). And the third attempt is the Organicist theory, which claims that art is "a class of organic wholes consisting of distinguishable, albeit inseparable, elements in their causally efficacious relations which are presented in some sensuous medium" (Weitz, 1956: 29).

Weitz's critique of mentioned attempts as an illustration of the first aspect of Plato – Weitz's problem. The first aspect of the Plato–Weitz's problem says that, for most of our everyday concepts, we did not succeed in formulating a valid, satisfying definition of them. We use the concept of art very often. So, if there is a way to show that the mentioned definitions of art are invalid or unsatisfying, I would say that we dispose of a good illustration of the first aspect of the Plato–Weitz's problem. But how could one show inadequacy of a definition? In principle, there are two customary methods. The first one is to argue that composed definition is too broad or too inclusive, i.e. that composed definition of concept X can be applied on much what is actually not an instance of X. The second one is to argue that composed definition is too narrow, which means that we cannot apply it on much what is actually an instance of X.

In ''The Role of Theory in Aesthetics'', Weitz uses both of mentioned methods; namely, he explicitly claims: "Some of them, in their search for necessary and sufficient properties, emphasize too few properties, like (again) the Bell-Fry definition which leaves out subject-representation in painting, or

the Croce theory which omits inclusion of the very important feature of the public, physical character, say, of architecture. Others are too general and cover objects that are not art as well as works of art." (Weitz, 1956: 29). Indeed, Weitz has not specified what makes every single definition of art, which he took into consideration, too broad or too narrow. However, he obviously believes that no definition is satisfying exactly, because it is either too broad or too narrow. Now all we should do in order to make a strong parallel between Weitz's critique and the first aspect of the Plato–Weitz's problem is to expand Weitz's critique and to apply it to all definitions of art that we can remember. This exercise is left to the reader.

Weitz's critique of the possibility of defining art as the illustration of the second aspect of Plato – Weitz's problem. The second aspect of the problem says that for at least some of our everyday concepts, it is impossible to formulate a valid, satisfying definition. How could one show that? According to Weitz, the key is that the concept of art is an open concept: "'Art', itself, is an open concept. New conditions (cases) have constantly arisen and will undoubtedly constantly arise; new art forms, new movements will emerge, which will demand decisions on the part of those interested, usually professional critics, as to whether the concept should be extended or not." (Weitz, 1956: 32) The basic idea is obvious. Definitions assume that the concept we define has a fixed, unchangeable meaning. That is clearly not the case in the example of art and many other concepts which Weitz calls 'open concepts'.

5 A new approach: The prototype theory

So far, we have seen that we can analyse concepts in various different ways. What they have in common is that they can all be perceived as seeking for a definition. However, it seems that philosophers did not succeed in formulating definitions for most of our everyday concepts. Moreover, Morris Weitz argued that definitions of some concepts are not even possible. Naturally, the following questions emerge now: Are we supposed to abandon conceptual analysis? Or should we try only to modify it in such a way that definitions are not the hoped-for result?

I believe that abandoning conceptual analysis would be too radical. As mentioned before, it has at least some heuristic value. It helps us to realize what is not the meaning of the analysed concept. For example, Socrates' conceptual analysis was able to show that piety could not be what is pleasing to the gods. But if we are to keep practicing conceptual analysis, we are obliged to explain in what way. If definitions are not the goal of conceptual analysis, what is? In this chapter, I want to present an

alternative to definitions – the so-called Prototype Theory. It is actually a theory of concepts that appeared in the 1970s. Just as Weitz's arguments, it has roots in Wittgenstein's philosophy. I would like to emphasize that my aim is only to briefly present it as a *possible* alternative, i.e. to show that we can analyse concepts in such way which does not require definitions as the final result. I do not intend to argue that we have to accept that theory. Accordingly, limits and problems of the Prototype Theory shall not be considered in this paper.

The fundamental claim of the Prototype Theory can be formulated as follows: "According to the Prototype Theory, most concepts – including most lexical concepts – are complex representations whose structure encodes a statistical analysis of the properties their members tend to have. Although the items in the extension of a concept *tend* to have these properties, for any given feature and the property it expresses, there may be items in the extension of a concept that fail to instantiate the property. Thus the features of a concept aren't taken to be necessary…" (Laurence and Margolis, 1999: 27). There is a clear analogy between this quote and Wittgenstein's theory of family resemblances. Wittgenstein's theory indicated that there are no non-trivial necessary and jointly sufficient conditions for the use of the concept of game. Moreover, it could be said that, according to Wittgenstein, games *tend* to be fun, to be played by more than one player, etc.

More important questions than the analogy between Wittgenstein and Prototype Theory are 'How do we analyse concepts?' and 'What is the goal of our analysis if concepts could be grasped as mental representations that contain a list of properties which items in the extension tend to have?' I believe that the answers to those questions are not hard to give, because they are implicitly contained in the formulation of the theory: 1) the goal of our analysis is to discover the list of properties that the items in the extension of analysed concept tend to have – for example, if we analyse the concepts of art, we want to discover what properties those items that we call 'art' tend to have; 2) we analyse concepts by carefully examining many[5] items which are considered to fall into the extension of the analysed concept and detecting their most important and most distinctive properties.

6 Conclusion

It seems that with the Prototype Theory, we have finally made a step forward. If we accept it, we will avoid the Plato–Weitz's problem. Naturally, the Prototype Theory has its own problems;[6] however,

[5] As much as possible
[6] For further discussion on this topic, see Margolis and Laurence (1999).

even if the Prototype Theory cannot provide satisfying answers to its critics, it is highly significant for the development of conceptual theory. Namely, it emphasizes one extremely significant fact: meanings of concepts are not precise in such a way that they could be formulated simply in one sentence, in terms of necessary and sufficient conditions. Natural language is something imperfect, but also something 'alive'. This fact will be, in my opinion, the basis of every future progressive conceptual theory.

Reference list

Beaney, M. (2016). "Analysis". *The Stanford Encyclopedia of Philosophy,* Zalta E. (Ed.), Retrieved (2017, June 23) from <https://plato.stanford.edu/archives/sum2016/entries/analysis/>

Gettier, E. (1963). "Is Justified True Belief Knowledge?". *Analysis,* Vol. 23, pp. 121-123

Kant, I. (1999). *Critique of Pure Reason,* Guyer, P. and Wood. A. (Trans. and Ed.). Cambridge: Cambridge University Press (Original work published in 1781/1787)

Laurence, S. and Margolis E. (1999). "Concepts and Cognitive Science". In Margolis E. and Laurence S. (Ed.), *Concepts: Core Readings,* Cambridge, Massachusetts: MIT Press, pp. 3-81

Plato. (trans. 2002). *Euthyphro.* In *Five Dialogues – 2nd ed.* Grube, G.M.A. (Trans.). Indianapolis/Cambridge: Hackett Publishing Company, Inc.

Weitz, M. (1956). "The Role of Theory in Aesthetics''. *The Journal of Aesthetics and Art Criticism,* Vol. 15, No. 1, pp. 27-35

Wittgenstein, L. (1986). *Philosophical Investigations – 3rd ed.* Oxford: Basil Blackwell Ltd.

Tone Kralj in Three Possible Worlds: Defence of the Ethical Value of Art

Neja Kaiser, University of Maribor

Abstract

At the end of the 19th century, the main task of any critic was to ethically evaluate a given work of art, while in the early 20th century, the viewpoint that prevailed stated that ethics and aesthetics are autonomous fields with their own values, which means that the evaluation of the aesthetic field should be free of ethics.

The quest for the autonomy of art, shortly called aestheticism, can be summed up in Oscar Wilde's thought: "There is no such thing as a moral or an immoral book." This can be supported by the 'common denominator argument', i.e. the fact that many works of art have nothing to do with ethics.

In this paper, we are going to show that even if some works of art have nothing to do with ethics and are therefore not the appropriate subject for ethical criticism, it does not mean that ethical criticism is unsuitable for all artworks. An existing ethical (dis)value of art will be corroborated through a thought experiment.

The thought experiment is called 'Tone Kralj in Three Possible Worlds'. Tone Kralj was a Slovenian painter, printmaker, and sculptor. At the time of Fascism, he affirmed his Slovenian nationality by including covert anti-Fascist allusions in his pictures and paintings of churches, for example on Mengore (1929/1930), Šentviška Gora (1941), and Hrenovice (1941/1942). This is one world, one reality. Besides this one, we are going to imagine two more, where his works of art become slightly different.

Key words: Tone Kralj, aesthetic value of art, ethical value of art, autonomy of art, aestheticism

1 Introduction

In 360 B.C., Plato pointed out the inadequacy of fairy tales that imply the imperfection of the gods, describe moaning heroes, and shame the underworld. According to him, those fairy tales were preventing good music education and the growth of children's souls. (Plato, *The Republic*: 376e–407d)

Even though, in the history of philosophy, aesthetics arose in the context of ethically motivated criticism of art, the ethical value of art and its evaluation faced disfavour at the beginning of the 20th century. Influenced by Kant's theory of aesthetics, or better, its (mis)interpretation, many movements were formed – for example, aestheticism and formalism – which caused the ethical criticism of art to become either irrelevant or conceptually illegitimate. (Carroll, 2000: 350)

Despite the feeling that the thesis of the autonomy and separateness of the aesthetic and ethical areas is still persistently gaining in popularity, we may say that recently, more and more humanistic and social critics are acknowledging the ethical values of art in criticism, especially when considering issues such as racism, sexism, homophobia, etc. (Carroll, 2000: 350)

However, it seems reasonable and useful to begin with a presentation of autonomism, which can be summed up in Oscar Wilde's thought: "There is no such thing as a moral or an immoral book. Books are well written or badly written. That is all." (Wilde, 1913: 5)

2 Autonomism

When we talk about autonomism, we talk about a belief that ethics and aesthetics are autonomous areas with their own values. Supporters of this view thus believe that when we evaluate the aesthetic field, we should not use the criteria from the ethical field. They claim that artworks are not valuable because they would serve some ulterior purposes, such as moral enlightenment of improvement; rather, they are valuable intrinsically. (Carroll, 2000: 351)

This approach may be called aestheticism, but this would require the emphasis on aesthetics as the main feature of artworks instead of the autonomy of ethic and aesthetic fields. One of the sociological explanations argues that the popularity of aestheticism is actually the art world's clever way to protect artworks from censorship. Their answer to Plato – who is also known for the fact that in his city-state, all existing melancholic and soft melodies should be prohibited – is that the assessment of art should not be based on enforced values. More historically specialized hypothesis of aestheticism states that aestheticism had arisen as a reaction to the triumph of the bourgeois culture in the 19th century. We are talking about a rebellion against the instrumentalization and commercialization of value, and the bourgeois thinking that it is possible to include the value of art in market value. Adversaries of popular art, which emerged with the rise of industrial design, also defend aestheticism. We may see the latter in the service of High Art and understand it as a gesture of cultural rebellion, for it seeks the value of an artwork that is separated from any instrumental and practical purposes. (Carroll, 2000: 351–352)

(i) ‘Common denominator argument’ and its criticism

The idea of the autonomy of art may be supported by the ‘common denominator argument’, i.e. the fact that many works of art have nothing to do with ethics. Four of Vivaldi’s concerts devoted to the seasons, Le Corbusier’s Villa Savoye, and many Monet’s paintings of water lilies from his garden in Giverny count as art, but they hardly promote an ethical dimension. The following question is how the mentioned works of art could thus be susceptible to ethical evaluation. In addition, this raises the question of a common moral criterion. When determining the artistic value, we must note that this value should be such as to apply to all artworks. Since ethical criteria cannot be universally applied to all artworks, it would be necessary to look for the appropriate value elsewhere. (Carroll, 2000: 352)

Where? At this point, the aesthetic experience is often highlighted. The criterion for the evaluation of art we are looking for is therefore the capacity to evoke the aesthetic experience. The aesthetic experience could take the given role, as it is explicitly defined in terms of disinterestedness, disinterested pleasure, or disinterested attention; and that confirms the aesthetics’ independence from ethics. Some autonomists, whom we may classify as formalists, say that the aesthetic experience revolves only around the assessment of formal aspects of art, while others have a slightly broader understanding of the concept of aesthetic experience. They speak about the aesthetic experience as the experience that is prescribed by an artwork and is important for its own sake (and not because of anything else, for example moral enlightenment or moral improvement). Autonomism also has supporters amongst the defenders of the essentialist perspective on the value of art. This view argues that the value of art, whatever it may be, is a value unique to art. Art is the only human practice the products of which primarily offer the aesthetic experience. Interestingly, sermons and ethical discussions can be aesthetically pleasing, but this is not their primary aim. Only artworks are meant to evoke the aesthetic experience and because of that, the aesthetic experience, not ethics, provides the appropriate evaluating frame for art. (Carroll, 2000: 352–353)

Autonomists warn us of another possible link between art and ethics. More specifically, they remind us that there are certain ethically dubious works of art that are brilliant and unsurpassable, for example Leni Riefenstahl’s film *Triumph des Willens*; and at the same time, there are works that cannot be morally arguable, but that is not enough to announce them as marvellous art pieces. (Carrol, 2000: 353)

Given the above, we can say that autonomists do not wonder whether or not there is such a thing as the ethical value of art, for the answer seems obvious. Rather, they deal with the ‘common denominator

argument', which assumes that there must be a suitable criterion for evaluating art, and it must be applicable to all artworks. But since many works of art are not directly related to morality, autonomism stresses that ethics cannot be the relevant criterion for evaluating art. But is it really so? The fact that some artworks are not concerned with morality directly, meaning they lack an ethical dimension or even ethical implications, really does show that some artworks could not be the subject of ethical criticism. However, it cannot be deduced from the above that ethical criticism is inappropriate for all artworks, since there are many art pieces with ethical content. (Carroll, 200: 357)

The 'common denominator argument' also says that we are looking for a single criterion for the evaluation of all works of art, with the emphasis on 'single criterion' and 'all'. Supporters of ethical evaluation of art have two possible answers. Firstly, they ask themselves whether the existence of a universal evaluating criterion for art would exclude the possibility of local criteria for evaluating individual artistic genres which are compatible with the selected universal criteria, and their answer is negative. Secondly, they doubt that there is only one criterion for evaluating all artworks. Ultimately, what do Malevich's *The Black square*, Palais Garnier, and Sex Pistols' song "Anarchy in the U.K" have in common? (Carroll, 2000: 358)

Considering this last challenge, the autonomists offer a very common and already discussed proposal that the aesthetic experience is the universal criterion for evaluating art that we are searching for. However, since the concept of aesthetic experience is unclear, the proposal seems problematic. When autonomists say that the aesthetic experience is the experience of a specific kind, this cannot serve as a universal evaluating criterion, since there are some pieces of art that are intended to come into conflict with such experiences. One of the artworks designed to undermine the perception of what is usually understood under the term 'significant form' is the infamous composition *4'33''* by John Cage, an American experimental composer. It offers four minutes, thirty-three seconds of silence. The pianist opens the score, but he or she never begins to play. The audience are listening to a composition that consists of the sounds of creaking chairs, coughing, whispering, clock-ticking, etc. (Carroll, 2000: 358)

Supporters of autonomism have avoided this problem with a new and different definition of aesthetic experience. When they talk about the aesthetic experience, they talk about the experience that is valued for its own sake. Moreover, they believe that any artwork can be assessed in terms of its intended capacity to offer such an experience. However, we must say that this proposal of the definition of aesthetic experience is not universally applicable to all artworks, for there are artworks that have no intention to offer that kind of aesthetic experience. Let us consider one of the monochromatic canvases

with cuts – the works by Lucio Fontana, the founder of Spatialism. His pieces, admired by many art lovers, were not created to hang in galleries and to provide the aesthetic experience. Those works were created as the result of a special process, which many famous photographers tried to capture in their pictures. In order to avoid this issue, autonomists may omit the requirement that the artworks are *meant* to offer an aesthetic experience (intentionally) and retain only the requirement that artworks *may* offer the aesthetic experience (unintentionally). Yet, this would not be in their interest, since it would mean that they should use this criterion of value – which is unique to art – for assessing natural phenomena and objects as well. After all, a rainbow or a waterfall may offer an aesthetic experience as well. (Carroll, 2000: 358–359)

So far, the autonomists have been unable to develop and offer any new proposals for unique and universal criteria for the evaluation of all art and only art. Because of that, we will now list three facts, with which we want to banish the autonomists' fear that ethical criticism would replace aesthetical criticism: i) supporters of ethical criticism are not required to assume that all art is susceptible to moral criteria; ii) in the case of artworks with an ethical dimension, evaluating them in terms of the quality of their moral perception (or misperception) does not include the criteria alien to their value as the kind of artworks they are – it is a matter of evaluating the artwork in terms of the norms (genre norms) of the art form to which it belongs; and iii) supporters of ethical criticism need not claim that ethical value is the only possible value for an artwork – even artworks that are appropriate for ethical evaluation may have formal dimensions which call for an independent evaluation. (Carroll, 2000: 359)

The said may be considered the ethical critic's response in the case of unethical artworks which are nevertheless excellent pieces of art. Many think that other qualities of Leni Riefenstahl's film largely outweigh its ethical defects. The same fact also allows the ethical critic to declare an artwork, highly praiseworthy for its ethical aspects but lacking the adequate formal and experiential aspects, insignificant. (Carroll, 2000: 360)

3 Ethical (dis)value of artworks

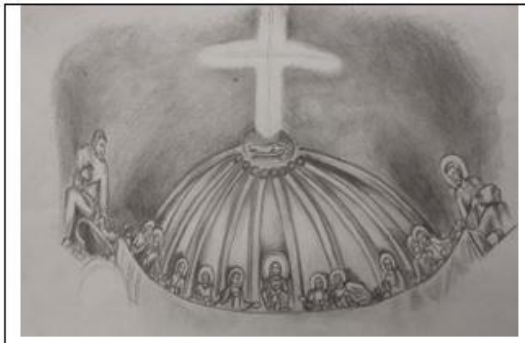
As we have shown, the follower of autonomism acknowledges that some artworks have an ethical dimension but thinks that ethical dimension is not associated with ethical (dis)value. With the following thought experiment, which we named ‘Tone Kralj in Three Possible worlds’, we are firstly going to explain what we have in mind when we talk about the ethical dimension of art. Secondly, we are going to show that ethical (dis)value is not alien to artworks; and finally, we are going to provide an example of an artwork with ethical value and an example of an artwork with ethical disvalue.

(i) Thought experiment: Tone Kralj in Three Possible Worlds

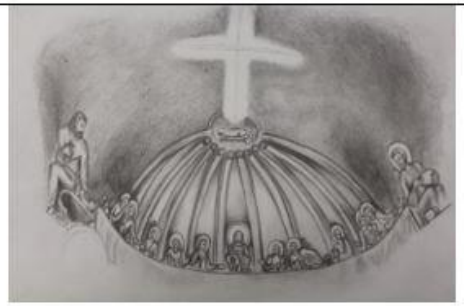
Possible world no. 1	Possible world no. 2	Possible world no. 3
Tone Kralj, a Slovenian painter, sculptor, printmaker, illustrator, and architect, was born in Zagorica on 23 August, 1900. Between 1906 and 1912, he was attending a public primary school in Videm. Then he entered the St. Stanislaus diocesan classical gymnasium. When he was seventeen years old, he was sent to the front line, where he witnessed the horrific events of the First World War. After the war, he continued with his education and graduated in 1920. As a gymnasium student, he created illustrations and developed his artistic talent with his older brother France Kralj and his gymnasium professor Gašper Porenta. He became a member of the Klub Mladih, later renamed the Slovenian Art Society, which was founded and led by his brother France. In the autumn of 1921, he went	Tone Kralj, a Slovenian painter, sculptor, printmaker, illustrator, and architect, was born in Zagorica on 23 August, 1900. Between 1906 and 1912, he attended a public school in Videm. Then he entered the St. Stanislaus diocesan classical gymnasium. He graduated in 1920. As a gymnasium student, he created illustrations and developed his artistic talent with his older brother France Kralj and his gymnasium professor Gašper Porenta. He became a member of the Klub Mladih, later renamed the Slovenian Art Society, which was founded and led by his brother France. In the autumn of 1921, he went to the Academy of Fine Arts in Prague to Professor Jan Štursa. He had studied sculpture until 1927. Then he came back to Slovenia. His catechist invited him to Julijska krajina. He painted churches' walls	Tone Kralj, a Slovenian painter, sculptor, printmaker, illustrator, and architect, was born in Zagorica on 23 August, 1900. Between 1906 and 1912, he attended a public primary school in Videm. Then he entered the St. Stanislaus diocesan classical gymnasium. He graduated in 1920. As a gymnasium student, he created illustrations and developed his artistic talent with his older brother France Kralj and his gymnasium professor Gašper Porenta. He became a member of the Klub Mladih, later renamed the Slovenian Art Society, which was founded and led by his brother France. In the autumn of 1921, he went to the Academy of Fine Arts in Prague to Professor Jan Štursa. He had studied sculpture until 1927. Then he came back to Slovenia. Hard times had begun. He did not get any orders. The circumstances

to the Academy of Fine Arts in Prague to Professor Jan Štursa. He had studied sculpture until 1927. Then he came back to Slovenia and eventually started working for a secret anti-Fascist organization, the members of which were mainly priests from Julijska krajina. He affirmed his Slovenian nationality by including covert anti-Fascist allusions in his pictures and paintings of churches in that area. In 1929, he painted *The Way of the Cross* in Mengore. While negative characters from biblical stories were painted as icons of the Fascist regime, faces of suffering Christians represented Slovenians. In 1939, he started working on his paintings in a church in Most na Soči, where he caught the attention of the Fascist police. He moved to Venice, where he applied for the study of architecture, but only under false pretences. In 1940, he continued with his studies in Rome. During the years of the occupation of the Slovenian territory, he	and ceilings. In 1929, he painted the church of the Holy Name of Mary. His work included the *Way of the Cross* and a plan for a painting in the presbytery. He followed the principle of a Slovenian philosopher and ideologist of political Catholicism, Aleš Ušeničnik: "The value of faith is above ethnicity – do not overturn the order of values." Tone Kralj decided to paint without any political view. Actually, he never specifically dealt with politics. He just wanted to create art. He wanted to know more about art, and because of that, he moved to Venice and began to study architecture. In 1940, he continued with his studies in Rome. Since he was not in conflict with the regime, unlike some intellectuals and representatives of political minorities, he came back to Julijska krajina. Among other things, he received an order for painting the church on Štentviška Gora (1941) and Hrenovice (1941/1942). The presbytery on Šentviška	pushed him to the margins of society. He violently reacted and protested at the Museum of Modern Art Ljubljana. That made him and his art even more undesirable. In 1918, the Italian authorities started replacing Slovenian priests with Italian clerofascists. In accordance with the labelling of the other newly acquired areas, they decided to mark the churches' paintings with Italian identity (italianita). They invited Tone Kralj, who had always fancied the Fascist ideas. He moved to Julijska krajina and started painting the churches' walls and ceilings. On Mengore, he painted the *Way of the Cross* and the motif of the 'Assumption'. Among the witnesses of the event, he magnificently portrayed Mussolini. In 1939, he moved to Venice, where he began to study architecture. In 1940, he continued with his studies in Rome. He radicalized his Fascist attitude. During the years of the occupation of the Slovenian territory, he returned and started

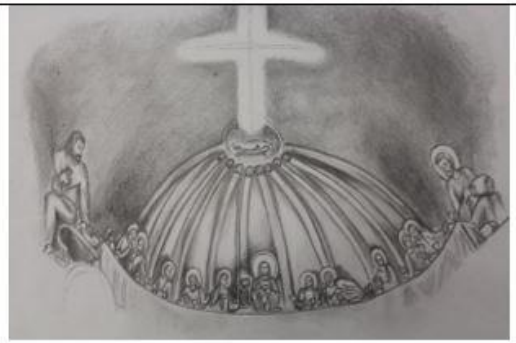
returned and painted in Julijska krajina. He faced the confiscation of property. The Fascists destroyed and lost many of his paintings and graphics matrices. He radicalized his anti-Fascist views, and he joined the patriotic liberty struggle. The paintings in churches, for example in Šentviška Gora (1941) and Hrenovice (1941/1942), contain anti-Fascist allusions and include ideological implications of Slovenian national identity. In 1942, he received an order for a painting in the church of St. Michael in Lokev. In the presbytery, he decided to paint the motif of Last Supper. He prepared the next sketch.	Gora was decorated with motifs like 'Solution of St. Peter' and 'Conversion of St. Paul'. He painted the presbytery, nave, and chapel with scenes from the 'Way of the Cross' and motifs like 'Resurrection'. In 1942, he received an order for painting the church of St. Michael in Lokev. In the presbytery, he decided to paint the motif of Last Supper. He prepared the next sketch.	to paint in the churches in Julijska krajina. The paintings in churches, for example in Šentviška Gora (1941) and Hrenovice (1941/1942), included Fascist symbols, which were not attributed negative, but positive connotation. In 1942, he received an order for a painting in the church of St. Michael in Lokev. In the presbytery, he decided to paint the Last Supper. He prepared the next sketch.



Tone Kralj decided to paint Judas Iscariot, a very famous negative character from the Bible, as a Fascist icon. To be more precise, Judas Iscariot has a face of Gabriel D'Annunzio, a Fascist politician and poet. In the painting, we can see that the combination of the colours of his clothing imitates the Italian *tricolore*.



Tone Kralj decided to follow the words from the Gospel of Matthew. He painted a long table covered with white tablecloth. Jesus sits in the middle of the table with ten apostles surrounding him. At the table corners, Tone Kralj decided to highlight the eleventh and twelfth apostle. The one on the left edge has a pouch with thirty silver coins. He can be identified as Judas.



Tone Kralj decided to paint the fasces in the right hand of John the Evangelist. I believe you can see that Jesus is glancing towards one of the most well-known Fascist symbols. We may also say that he is blessing it with his right hand.

Do you notice the difference between sketch no. 2 and sketches no. 1 and no. 3? The answer should be obvious. While the latter two deal with Fascism, sketch no. 2 does not. Sketch no. 2 quite diligently follows the words from the Gospel of Matthew and appears to be an ordinary sketch of the 'Last Supper'. In other words, sketch no. 2 does not have an ethical dimension. Admittedly, the sketch is occupied with Christian symbolism and moral doctrine, but it certainly does not have any additional ethical dimension. This additional ethical dimension we were referring to at the beginning is definitely present in sketches no. 1 and 3.

Let us ask the next question. If you had to choose, which sketch would you pick for the realization, no. 1 or no. 3 – the sketch which shows a Fascist icon as something negative or the sketch which glorifies one of the most well-known Fascist symbols? Or to put it simply, the sketch which criticizes Fascism or the sketch which glorifies it?

Your answer is actually unimportant. What really matters is that you have chosen one. In fact, while you were choosing, you were evaluating – evaluating the (additional) ethical dimension of the sketch. But why? The details in which the ethical dimension is manifested represent the only difference between sketch no. 1 and sketch no. 3. Everything else, including the author and the formal characteristics, remains the same.

Therefore, you have attributed ethical value to one sketch and ethical disvalue to the other. Presumably, your answer to the question was sketch no. 1. We are guessing this has happened because of the sketch's critical attitude towards Fascism; thus, you applied ethical value to sketch no. 1. Consequently, we assume you did not select sketch no. 3. Why? Because of its glorification of Fascism. You identified sketch no. 3 as ethically disvalued.

Of course, there is a possibility that you selected sketch no. 3 and thus attributed ethical value to it. That means that you classified sketch no. 1 as ethically unacceptable. This is irrelevant to our goal in this article. At this point, we will not be working on how to properly get to the attribution of ethical (dis)value of the artworks we are talking about. For now, we just want to show that sometimes we cannot avoid the ethical evaluation of artworks. This means that the autonomists may be mistaken when they deny all connection between the ethical dimension of the artwork and the ethical (dis)value of the artwork.

Based on the findings above, we believe we may say that artworks with ethical dimension are familiar with ethical evaluation. Therefore, works of art with ethical dimension may be described as appropriate subject for ethical criticism.

4 Conclusion

In this paper, we endeavoured to i) present autonomism and its 'common denominator argument'; ii) show that even if some works of art have nothing to do with ethics and are therefore not the appropriate subject for ethical criticism, it does not mean that ethical criticism is unsuitable for all artworks; and iii) present the attempt to support the existence of ethical value of art through a thought experiment, namely 'Tone Kralj in Three Possible Worlds'.

Reference list

Carroll, N. (2000). "Art and Ethical Criticism: An Overview of Recent Directions of Research". *Ethics*, 110, pp. 350–387. Retrieved (2017, April 4) from <http://www.jstor.org/stable/10.1086/233273>.

Gaut, B. (1998). "The Ethical Criticism of Art". In Levinson, J. (Ed.), *Aesthetics and Ethics*, Cambridge: Cambridge University Press, pp. 182–203.

Kante, B. (2001). *Filozofija umetnosti*. Ljubljana: Jutro.

Pelikan, E. (2016). *Tone Kralj in prostor meje*. Ljubljana: Cankarjeva založba.

Plato. *The Republic*. (1968). Translated by Alan Bloom, New York: Basic Books.

Wilde, O. (1913). *The Picture of Dorian Gray.* London: Simpkin, Marshall, Hamilton, Kent and Co. Ltd. Retrieved (2018, May 28) from <https://babel.hathitrust.org/cgi/pt?id=mdp.39015092867749;view=1up;seq=9>.

What is Analytic in Analytic Aesthetics?

Aleksandar Jevremović, University of Belgrade

Abstract

My paper consists of four parts. In the first part, I analyse 3 distinctive attempts to determine the nature of analytic aesthetics. I take my position somewhere between those 3 approaches. In the second part, I point out the major oversight of all previous attempts to determine nature of analytic aesthetics – the ignorance of eighteenth-century aesthetics. In the third part, I argue that there are deep structural similarities between the enterprise of analytic aesthetics and eighteenth-century aesthetics by reconstructing the defining characteristics of analytic philosophy only by reference to eighteenth-century aesthetics. In the last part, I sketch a plan for some future inquiry which would not overlook the importance of eighteenth-century aesthetics.

Key words: Analytic aesthetics, analytic philosophy, eighteenth-century aesthetics, anti-essentialism, aesthetic judgments

1 On what has been done

The most obvious approach to defining analytic aesthetics is simply to say that it is not continental aesthetics or to say that it is British-American aesthetics. However, those propositions say almost nothing about the nature of the analytic aesthetics enterprise. In order to reveal the nature of X, one must articulate the definition of X. Therefore, if we are interested in the nature of analytic aesthetics, we have to give the definition of it. Nevertheless, it seems that defining any philosophical discipline (or philosophy itself) is an almost impossible task; moreover, it is hard to just name one thing that each specific analytic aesthetics enterprise has in common with the others. As it turns out, trying to define analytic aesthetics blurs the nature of it rather than revealing it. But are we to give up on our interest to better understand analytic aesthetics? I would say no. Although we can't define it, we can still say a lot about it.

I distinguish three different approaches to the problem. One is (I) to define analytic philosophy and then *a fortiori* define analytic aesthetics. The second is (II) to enumerate the main characteristics of analytic aesthetics. The third is (III) to reconstruct an early program of analytic aesthetics. All these approaches can be combined. In fact, I am going to focus on three influential articles that deal with the problem of the nature of analytic aesthetics. The first one is Shusterman's introduction to his book,

Analytic Aesthetics. There, he combines the first and the second approach. The second one is an article by Anita Silvers, "Letting the Sunshine in: Has Analysis Made Aesthetics Clear?" She combines the second and third approach. The third and final article is by Lars-Olof Ahlberg, "The Nature and Limits of Analytic Aesthetics", in which he mainly uses the first approach. In the continuation of this section, I will illustrate the main points of these articles, and then I will argue that they all have one thing in common. They all unavoidably refer to analytic philosophy at some point as the only historical foundation of analytic aesthetics.

2 Shusterman's determination through enumeration

At the very beginning, as I indicated, Shusterman refutes the possibility of the definition of analytic aesthetics. He thinks that it is "unreasonable to expect such definition" not because it is a "complex field", but because "the definition of any philosophical field is a problem itself and therefore shares the resolute irresolution of philosophy" (Shusterman, 1989: 2). One convincing reason can be named for that. Analytic aesthetics is still "alive", and it literally changes its nature with every new and different way of inquiry. If one way is not fruitful, an aesthetician can change the style or even method in order to improve the enterprise. I think that it is a valid argument. If the nature of X constantly changes, we can't capture it; and even if we do, we can't predict its future. But what can we do?

We can simply point out some characteristics that belong to most of the specific theories in analytic aesthetics. Shusterman points out exactly 10 such characteristics. I will not expose all 10 of them here, nor is it necessary. I find 4 characteristics to be fundamental:[1] 1) anti-essentialism, 2) self-conception as meta-criticism, 3) main interest in aesthetic judgments and concepts, and 4) main interest in art and not nature. Anti-essentialism refers to the idea that there is no common feature 'F' that, once discovered, serves as a standard of aesthetic judgment. What Shusterman has in mind here are so-called theories of art. Theories of art aim to discover the nature of art that can be formulated into the definition of it, which contains a set of necessary and together sufficient conditions. Theories of art were viciously criticized by early analytic aesthetician Morris Weitz in his paper, "The Role of Theory in Aesthetics". Weitz proposed there that art is an "open concept" and that it resists traditional

[1] I take these 4 to be fundamental because others, like Shusterman's number 2, refer to some external relations, like rebellion against Croce's idealist aesthetics. As I have indicated in the beginning, those characteristics reveal nothing about the nature of analytic aesthetics. Other Shusterman's characteristics can be considered the same class. For example, I have united his 4th and 6th because they speak of the same thing – self-conception as second order discipline.
[2] Same idea openness can be applied to our problematic concept of 'analytic aesthetics'. Perhaps that can help us better understand Shusterman's argument against definition from the beginning.

conceptual analysis of necessary and sufficient conditions. A concept is open when its conditions of application are "emendable and corrigible" (Weitz, 1956: 31).[2] The second main characteristic is self-conception as meta-criticism. What exactly is meta-criticism? Broadly speaking, it can be understood as a branch of "philosophy of art (or aesthetics) that takes as its object of inquiry the criticism of the arts" (Carrol, 2009: 1). Carroll argues that the main reason for the flourishing of meta-criticism was the direct consequence of anti-essentialism: "For, if you couldn't define art, how could you hope to develop a philosophy of art? In its stead, the next best thing seemed to be to construct a philosophy of criticism" (Carroll, 2009:1). Meta-criticism can be understood in many ways. I see it as an inquiry that aims to expose our (i.e. critics) presuppositions about the nature of art, value, interpretation, ontology, etc. in natural language. It is seen by some as making our language about art clearer, less obscure. I see it more as making our language more conscious about what it states. The third characteristic Shusterman derives directly from considerations about analytic philosophy. This is the place where approaches (I) and (II) meet in Shusterman's essay. By analytic philosophy here I mean an early model of philosophy, developed by Bertrand Russell and George Moore. The paradigm, or should we say 'the specific method', of this philosophy can be found in Russell's early essay "On Denoting" (see Russell, 1992: 55). Shusterman distinguishes two types of analyses. One is reductive analysis, breaking a concept on more basic components or properties, which are its necessary and sufficient conditions. Another type of analysis is simply the clarification of an ambiguous concept. According to Shusterman, analytic aesthetics deals with the second type of analysis, for it is full of vague concepts. Because of this notion of analysis, he says, analytic aesthetics can't be understood without the reference to analytic philosophy (Shusterman, 1989: 5). The fourth characteristic is the interest in art and not nature. That is an almost trivial consequence of the first three characteristics. If there is no first-order, normative discipline on nature, there can neither be a second-order discipline. Shusterman gave up on defining analytic aesthetics and accepted a more modest approach.

3 Determination through early program

Nevertheless, his characterization of analytic aesthetics as a unique enterprise can still be seen as too ambitious. Anita Silvers is less ambitious than Shusterman. She tried not to give the characterization of the entire enterprise of analytic aesthetics but merely to give some general characteristics of the

early analytic aesthetics program. By early program she means a series of books and papers from 1946 to 1962 (Silvers, 1987: 137-138). Three characteristics can be extracted from her analysis of those works. The first is the reformation of aesthetics by replacing obscure and confused ideas with clear ones. This characteristic is very similar to Shusterman's characteristic of clarifying analysis. The second characteristic of the early program is anti-essentialism. Here, she completely agrees with Shusterman. The third and last characteristic that Silvers names includes the prohibition of generalizing insights gained from experience of particular artworks and then expecting generalizations to function as rules in arguments. This is actually the instantiation of the general project of making aesthetics less obscure. The concepts of traditional aesthetics are obscure, but among the most obscures are formulations of rules. This is also connected with anti-essentialism, for essentialism provides not only the definition of art but also a standard for evaluation and interpretation that rests on the definition.

Do we accept Silvers' or Shusterman's method? I believe that we should accept both. Shusterman is more thorough. He names more characteristics and reveals more about the nature of analytic aesthetics. Silvers tells us too little. On the other hand, Shusterman is, indeed, too ambitious. His characteristics are mistaken if we take analytic aesthetics as a whole, but if we consider what Silvers calls early program, it is much more plausible. There is no reason why we shouldn't combine both ideas.

4 Problems with anti-essentialism

We have seen that both Silvers and Shusterman talk of anti-essentialism as a very important characteristic. Lars-Olof Ahlberg denies that. He thinks that it was widespread, but it cannot be regarded as a defining characteristic of analytic aesthetics. Furthermore, he thinks that a common denominator cannot be found 'among theories' of analytic aesthetics, for they are very diverse; among analytic aestheticians, we can find all kinds of contradictory theories, for example, both formalists and anti-formalist, emotionalists and anti-emotionalist, etc. (Ahlberg, 1993: 12). I am not going to consider the second objection here; I agree with Ahlberg that a common denominator cannot be found in all of the analytic aesthetics theories, but I disagree that it cannot be found in the early program, as I already indicated. We can talk of defining characteristics but with limited scope. Ahlberg's first objection is more vicious. He claims not only that there are essentialists in the analytic tradition, but

that there are essentialists in early analytic aesthetics. If that was true, my proposal to talk about Shusterman's characteristics in terms of Silvers' early program would fail, for at least one characteristic would not stand – anti-essentialism. Others would probably follow. As we have seen, all four characteristics are inter-connected and interdependent.

But is Ahlberg right? In a way, he is, and in a way, he is not. It is a conceptual mess. Lamarque distinguishes between two kinds of essentialism (Lamarque, 2013: 26.2) One refers to the idea that there is some inherent or perceptual property of an art object that represents the essence of art. That kind of essentialism, call it traditional, was the subject of Weitz's attack. The other kind of essentialism is the kind that Danto, Dickie, and Beardsley proposed. All three of them have very different conceptions, but they have one thing in common: the essence of art is not something perceptual or something inherent to the artwork. Danto proposed that properties that matter are relational, not intrinsic. Beardsley offered a functionalist definition. Dickie saw artwork as an artefact that is presented to the artworld public (for all three conceptions see Lamarque, 2013: 26.2). In a sense, all three of them did not strictly offer inherent essence of art as a definition, rather a function, role or relation of artworks. In that particular sense, they are not traditional anti-essentialist.

It turns out that we can save our fusion of Shusterman's and Silvers' work; even if we admit that anti-essentialism is not the defining property of the early analytic aesthetics program, we can still say that traditional anti-essentialism is.

5 On what has been overlooked

All the articles that we have considered by now have two things in common. First, they all, at some point, unavoidably refer to analytic philosophy. Second, they all completely ignore the 'golden age' of aesthetics – the 18th century. Several factors qualify the 18th century as the golden age. First of all, in the 18th century, aesthetics was at last recognized as a legitimate and autonomous philosophical discipline. Many authors in Britain, France and Germany have included aesthetics in their philosophical systems. For example, in Germany, Baumgarten included aesthetics in his system for the first time, and it was followed by Kant and Hegel. Secondly, many important problems have been formulated. For example, Hume and Hutcheson have enriched the debate about the objectivity of aesthetic value or the property of 'beauty', Hume has philosophically articulated the problem of interdependence of moral and aesthetic value, etc. Consequently, many plausible answers have been formulated as a solution to these problems. Most of them have shaped the discussions in the centuries

to come. Analytic aesthetics and analytic philosophy or eighteenth-century aesthetics can be connected in two ways. Historically and structurally. It is not the aim of this paper to deny the historical nor the structural connection between analytic aesthetics and analytic philosophy. But the aim of this paper is to point out that there are historical and, more importantly, deep structural connections between eighteenth-century aesthetics and analytic philosophy. I focus on the connection with the eighteenth-century philosophy because it has been, as we have seen, overlooked by all authors who wrote on the subject.

In fact, I will try to reconstruct the entire early program of analytic aesthetics without any reference to analytic philosophy. What is the use of this enterprise? Simply to point out that eighteenth-century aesthetics shouldn't be overlooked when we talk about analytic aesthetics. If I can show that we can reconstruct the entire program of early analytic aesthetics just by referring to the eighteenth-century aesthetics, that would mean that there most certainly exists some non-trivial connection that we must understand in order to understand the nature of analytic aesthetics.

In Shusterman's, Silvers', and Ahlberg's papers we see almost no positive reference to any aesthetics from the history of aesthetics. That surely sounds strange. Should we accept that analytic aesthetics started from scratch? I see no reason to do that. On the other hand, we see a lot of positive references to analytic philosophy. For example, Ahlberg points out that "what analytic aesthetics is depended in how we characterize analytic philosophy" (Ahlberg, 1993: 6). Shusterman points out that analytic aesthetics is "a consequence (though perhaps not a mere epiphenomenon) of twentieth-century analytic approach to philosophy introduced by Moore and Russell." (Shusterman, 1989: 4)

Therefore, I propose a new way of analysing analytic aesthetics. It is not radically new because I refer to the usual characteristics that are attributed to the early analytic aesthetics, but it is new because I explain the structural origins of those characteristics just by referring to the eighteenth-century aesthetics, without any reference to analytic philosophy. The historical picture is probably much more complex; if we analysed the historical origins of those characteristics we would probably find all kinds of relevant factors, from both analytic philosophy and non-analytic philosophy, and all eras of the history of aesthetics. However, the aim of this paper is not to find those historical connections. That would be too ambitious for this kind of paper. I merely want to point out to every future historian what should not be overlooked when we talk about the nature of the analytic aesthetics enterprise.

6 Reconstructing the 'early program'

First of all, I want to limit the broad term of the eighteenth-century aesthetics. What I will mainly focus on is the aesthetics of Immanuel Kant and David Hume. I will argue that all four characteristics of what is called early analytic aesthetics program can be reconstructed just by using the ideas of Kant and Hume combined. If I succeed in that quest, then deep structural similarities between early analytic aesthetics and late eighteenth-century aesthetics will be proven. I propose to reconstruct the characteristics one by one.

1.) *Anti-essentialism*, as said, refers to the idea that there is some inherent or perceptual property of an art object that represents the essence of art. Both Kant and Hume rejected this kind of essentialism. Although the regular interpretation of Hume as a subjectivist would be sufficient for my point, I will go one step further and present a slightly different interpretation. According to the philosophical tradition, the property of 'beauty' was both the defining characteristic of an artwork and the standard for evaluation. One part of that tradition is objectivism, i.e. seeing the property of beauty as inherent to the artwork. Beauty is literally in the artwork, just like its shape and size. Another tradition is subjectivism, which states that the property of 'beauty' is not in the artwork, but in the eye of the beholder. According to the interpretation of Leon Kojen, Hume rejects both of these interpretations (Kojen, 1989: 126). Hume takes the intermediate way, which is dispositionalism. Beauty really is in the artwork; however, not like shape and size, but dispositionally. When we say 'N is beautiful' we say that under certain circumstances C, the specific consequence E will occur.[2] Thus, Hume obviously refutes the idea of the inherent essence of art, and is therefore anti-essentialist. According to Kojen, Hume even formulated something similar to Wittgenstein's idea of family resemblance (Kojen, 1989: 126).[3] That particular idea, as we remember, inspired Morris Weitz's theory of art as an open concept, which is the paradigm of the analytic aesthetics' anti-essentialism. Kant's anti-essentialism is obvious. Kant reduces the property of beauty to the "aesthetic experience (which) involves a free play of our cognitive powers, but all but the first also involve an interplay between this free play of our cognitive powers and specific concepts of understanding or reason, that is, theoretical concepts of nature or moral concepts." (Guyer, 2014: 143)

[2] Solubility is a dispositional property. Salt will dissolve only if we put it in water.

[3] Here I refer to Wittgenstein's idea about the application of a general concept 'F' on more things *a, b, c, d* etc. In many cases, it doesn't presuppose the existence of a property F, which is common to all these different things, but only the existence of a chain of family resemblances among them.

2.) *The main interest in aesthetic judgments and concepts*. Considering the anti-essentialist position, while we are evaluating, interpreting or defining the artwork, we cannot refer to the common property 'F', but we can analyse distinct formal features of the judgments of art. This idea can be found, explicitly formulated, in Kant's third Critique. Kant states that we have to start from the analysis of the judgments of taste to find what is required to call an object beautiful (Kant, 2004: 73). As Wenzel points out: "Compared with previous aesthetic theories, Kant's approach is marked by a certain shift of focus, a shift from the object to the judgment about the object. Instead of giving an account of the nature and quality of certain kinds of objects (the objects that we find beautiful), Kant analyzes a certain kind of judgment, namely the judgment of taste." (Wenzel, 2005: 2).

3.) *Self-conception as meta-criticism.* But what kind of judgments are we interested in? We are not interested in anyone's judgment about value or interpretation of an artwork. Analytic aesthetics' main interest is in judgments of art critics and art historians. Both art criticism and art history are first order evaluative and interpretative disciplines. Both praxes talk of art in various ways. When they are doing so, they necessarily presuppose some philosophical stances. Analytic aesthetics as a meta-criticism aims to elucidate, explicate and reveal those presuppositions. Let us see how meta-criticism works. Art criticism usually uses evaluative judgment that has the following form:

(1) This painting X is good…(judgment)
(2) …because it possess the quality K. (reason)

Some art critics consider the conjunction of (1) and (2) as sufficient. But some art critics feel that they should add a third judgment (at least potentially):

(3) Every painting that possess the quality K is inasmuch a good painting. (cannon)

Whether we need (3) or not is the typical meta-critical question. There is no way we could solve it inside the praxis of criticism. Criticism can just presuppose that we need or that we do not need (3). If we argue in favour of (3), we are stepping into the field of meta-criticism.[4]

This idea is explicitly formulated in Hume's essay "Of the Standard of Taste", where he postulates the art critic, a man of great knowledge and experience, as a standard of taste. Hume puts the art critic

[4] For the debate on critical reasons and aesthetic canons see Kojen, 1989: 13.

in the centre of aesthetics interest and deals with basic problems of aesthetics through the analysis of the practice of the critic's evaluation. (*see* Hume, 1991: 66)

4) The main interest in art and not nature. Kant is the exception here. He talks primarily about nature, not art, although he talks of art as well. But Hume and German aestheticians after Kant are not exceptions, they talk almost exclusively of art. I take this to be a mere consequence of 1)–3). From the fact that it analyses the judgments of a critic, it trivially entails that the main interest of analytic aesthetics is art, not nature, for the critic talks of art and not nature.

7 Conclusion

I hope that I have shown that we can reconstruct all 4 structural characteristics of analytic aesthetics just by referring to eighteenth-century aesthetic tradition, without any reference to analytic philosophy. This reconstruction aimed to show that there are deep structural similarities and analogies between these two enterprises. If there are structural similarities, we should at least consider the historical influence of eighteenth-century aesthetics on analytic aesthetics. This has often been overlooked by many authors. I believe that some future inquiry could prove non-trivial origins of analytic aesthetics in the eighteenth-century aesthetics. That would reveal some important aspects of the nature of analytic aesthetics, which was our original motive when starting this inquiry.

Nevertheless, if I still have not persuaded you, I will offer two more reasons why we should do that. Firstly, it is strange to explain the beginnings of analytic aesthetics by referring to analytic philosophy, because it strongly opposed aesthetics as a philosophical discipline. Secondly, the reconstruction of analytic aesthetics as the meta-criticism with reference to analytic philosophy is usually done by an analogy with philosophy of science as the meta-theory or meta-ethics. I think that a more intuitive approach is to reconstruct analytic aesthetics as a natural extension of the ideas that were started by Kant and Hume in the 18th century, which was interrupted by the analytic tradition, than by an analogy with philosophy of science.

Reference list

Ahlberg, L. (1993). "The Nature and Limits of Analytic Aesthetics". *British Journal of Aesthetics*, 33, pp. 5-16.

Carroll, N. (2009). *On criticism.* New York: Routledge.

Guyer, P. (2014). *A History of Modern Aesthetics: Volume I: The Eighteenth Century.* New York: Cambridge University Press.

Hjum, D. (1991). *O merilu ukusa*. Novi Sad: Izdavačka knjižarnica Zorana Stojanovića Sremski Karlovci.

Kant, I. (2004). *Kritika Moći Suđenja*. Beograd: Dereta.

Kojen, L. (1989). *Umetnost i vrednost.* Beograd: Filip Višnjić.

Lamarque, P. (2013). "Analytic Aesthetics". In Beaney, M., *The Oxford Handbook of the History of Analytic Philosophy*, Oxford: Oxford University Press, pp. 1655-1703.

Rasel, B. (1992). "O denotaciji". In Pavković A. and Lazović Ž.,*Ogledi o jeziku i značenju,* Beograd: Filozofsko društvo Srbije, pp. 51-61.

Shusterman, R. (1989). "Introduction: Analysing Analytic Aesthetics". In Shusterman, R., *Analytic Aesthetics*, Oxford: Basil Blackwell, pp. 1-20.

Silvers, A. (1987). "Letting the Sunshine in: Has Analysis Made Aesthetics Clear?". *TheJournal of Aesthetics and Art Criticism*, 46, pp. 137-149

Weitz, M. (1956). "The Role of Theory in Aesthetics". *The Journal of Aesthetics and Art Criticism*, 15 , pp. 27-35.

Wenzel, C. (2005). *An Introduction to Kant's Aesthetics: Core Concepts and Problems*. USA: Blackwell Publishing.

Arnauld in Anticipation of Contemporary Phenomenology of Colour

Niko Šetar, University of Maribor

Phenomenology, the study of the subjective experience of structures of consciousness, came to life as an autonomous philosophical discipline as late as in the early 20th century with the work of Edmund Husserl. There are earlier instances of phenomenological or at least pre-phenomenological thought, however. Some may argue that Brentano's classification of mental phenomena was already proper phenomenology. Others will go further and claim that even Kant's transcendental idealism was a form of the study in question. In this article, however, we will try to assert, that rather concrete roots of phenomenology may be found earlier still, in the works of 17th century Cartesian philosopher by the name of Antoine Arnauld.

It is commonly known that the works of René Descartes had major influence not only on philosophy, but on other sciences as well, including, for example, mathematics and optics. Yet less attention is usually given to a group of scholars a generation younger than Descartes, who operated in Port Royal in late 17th and early 18th century. One of these scholars was Descartes' student and faithful supporter, Antoine Arnauld, who significantly furthered his mentor's work.

It is in Arnauld's book *On True and False Ideas* (Originally: *Des vraies et des fausses idées*), where we can find the excerpt that this article will attempt to claim is phenomenological in nature. The book is a reply to Nicolas Malebranche's criticism of Cartesian dualism. At one point, specifically in chapter 23, Arnauld bases his defence of Descartes on dualistic view of secondary qualities. More precisely, he talks about colour perception and what it entails in terms of reality. He arrives at this topic via somewhat linguistically oriented path; he wanted to know why we speak of colours as though they are actually on objects. Cartesians claim colours are modifications of soul, *res cogitans*, and in no way modifications of body, *res extensa*, hence Arnauld's problematisation of attributing colours to objects. Descartes has no proper answer to this question. It is true he describes the physical attributes of light with astounding foresight and precision, claiming that the impression of colour requires a specific structure on the surface of the object it's being attributed to, which modifies the movement of light globules (one may claim the latter anticipate photons), and that, as a consequence, they somehow induce movements to the optical nerve. This description well approaches contemporary findings on colour perception, but it merely explains, on a physical level, how we perceive it. Arnauld, however,

was interested in why we perceive colours, or more specifically, why we perceive them on objects, when they are not actually there.

Arnauld finds that attributing colours to objects is in agreement with God's plan and benefits human beings – it helps us differentiate between objects with similar or same shapes et cetera. This idea has a certain evolutionistic nature, since it interprets colour vision as a biological advantage that serves the purpose of additional differentiation, for example between dangerous and harmless things in nature.

Since comparing the work of Arnauld to various authors in the field of phenomenology, such as Husserl, Heidegger, Merleau-Ponty etc. would be arduous and redundant, we shall attempt to determine the validity of our assertion by a single comparison. As Arnauld's alleged phenomenological thought is strictly limited to perception of colour and, on occasion, sound, we have chosen a contemporary phenomenological text, published in the journal *Philosophy and Phenomenological Research* dealing with that same specific type of perception.

The text chosen is titled *Seeing, Doing and Knowing: A Précis,* and is a work of contemporary cognitive approach philosopher, Professor Mohan Matthen of the University of Toronto. The text is a condensed summary of his book titled *Seeing, Doing and Knowing*, which is based on what Matthen calls the Sensory Classification Thesis – claim that perception systems assign distal objects to classes. The system of colour vision, for example, assigns objects to colour classes. Thus, says Matthen, sensory system operations give the organism a record about the state of the world.

In our analysis, we will work under the assumption that Matthen's text has sufficient phenomenological traits, which can be implied to also be featured in Arnauld's work if sufficient similarities are found during the comparison.

The first concept Matthen introduces is the concept of Sensory consciousness. He defines it as it follows:

> Sensory consciousness serves as an indication of how the senses have classified something. A thing that looks blue because once the sensory system has assigned it to that colour-class, it signals it has done so by tagging it with a blue 'look'. (Matthen, SDK: A précis, 2008: 392)

Therefore, sensory consciousness is a sub-branch of consciousness that specialises in interpreting the sensory input as a class designation. When a specific sensory input "arrives into" sensory consciousness, an impression of a colour, for example blueness, will form inside sensory consciousness, assigning that same colour to the object from which the sensory input originated.

With the exception of class designation or classification, Arnauld states much the same thing, albeit lacking the contemporary terminology:

> /…/ your soul is what is white, when you look at snow, what is black when you look at coal, and what stinks when you are near carrion. (Arnauld, Idées)[1]

By Cartesian definition, soul is functionally the same as consciousness; wherein the only distinction is the soul being in a dualistic relationship with the body – one could imply is that soul is in fact just consciousness, if that consciousness was separate from the body. As in Matthen's example, a sensory input arrives into consciousness (or soul), where an impression (idea, in Cartesian terminology) of a colour is formed, assigning that colour to the original object. For example, when you look at snow, the sensory inputs of everything that an image of snow entails arrive into your consciousness, where the impression of whiteness is formed, and that whiteness is simultaneously assigned to snow.

However, if we noted above, that it may be said that soul is a dualistic consciousness, would our above comparison not fall into some difficulties if it turned out that Matthen is in fact a physicalist? That might entail that the difference between soul and consciousness is indeed too great, since soul would be a concept too foreign, too removed from the concept of a physical consciousness. Luckily, Matthen and Arnauld's positions on physicalism are remarkably similar as well. Matthen notes:

> Physicalism leads to a puzzle. How can we have immediate knowledge of physical features? What kind of knowledge of sense-independent features do the senses provide? (Matthen, 2008: 393)

This is to say if it we claimed that sensory properties such as colours were physical in nature and existing outside the consciousness, which perceives them, in other words, were sense-independent, how could we know anything about them via our senses, and how could any knowledge we had about them be as immediate and as clear as it evidently is.

Descartes himself, as well as Arnauld have often argued for this point of view in various works. More precisely, in *The Ideas* Arnauld states:

> /…/ ideas of sensory properties are only confused and obscure when we compare them to the body, in the sense that they are modifications thereof. (Arnauld, 1986: 217)

[1] As I was not able to access the English translation of Arnauld's book during the creation of this article, I took the liberty of translating the quotes from French original to English myself. Therefore, quotes from Arnauld's book in this article may differ from quotes as found in other translations.

The terminology they use is different, but something being physical, for example, is the same principle as its being a modification of the body, and the notion of "clear and distinct" knowledge is very much analogous to the notion of knowledge being immediately accessible. It is evident both authors argue, that sensory properties cannot be physical and external in relation to the perceiving consciousness.

We find another match in both philosopher's views in Matthen's descriptions of sensory content. He identifies three types of property things that all have some sensory quality share. Firstly, there are real-world properties, which are presumably physical, and are detected by the sensory system, however the system does not recognise them independently from the impression they cause. This is what Matthen calls secondary sensory content. Secondly, they possess a property that they ought to be treated alike by epistemic operations. This is what he calls primary sensory content. We can exemplify this by thinking of one specific kind of sensory input – let us say colour. Secondary sensory content will be the wavelength of light arriving into one's eye, as well as probably the chemical changes in the receptor cells in the eye, while the primary sensory content will be that perhaps some objects with a specific shade of red will be recognised as hot – as having a high temperature. There is a third option, which states that they have the property that they ought to be avoided because of a negative effect they may have, which is implicable by one of their sensory properties (for example, something smelling foul might make you sick and ought to be avoided because of that). However, as Matthen notes, this is only a property of the effector mechanism, and can be disregarded. Even if one were to argue it is in fact sensory content, it would only be the furthering of primary sensory content at best, and in no way its own category.

We find Arnauld accurately describing the physical properties, which are required for secondary sensory content. Drawing from Descartes, he says:

> /…/ it [an impression of colour] is derived from particles on the surfaces of these two objects having different configurations, which causes the globules, by which the functions of light are transferred, to bounce differently from those two objects towards our eyes, where they cause different movements of optic nerve fibres. (Arnauld, 1986: 218)

It is obvious that this is an underlying physical property, which is causally related to phenomenal sensory experience. As Cartesian philosophy clearly states – and Arnauld agrees with that – physical things can never be perceived clearly and distinctly; that is to say they never can be immediately knowledgeable, and as Matthen states, they are detected, but not independently recognised, by the sensory system.

As for primary sensory content, Arnauld states that perceiving colours presents a simpler means of distinguishing bodies from each other. Thus, he says, we can avoid certain objects, or approach others etc. This reminds us of the third option Matthen discarded as being a part of the effector mechanism. However, an effector mechanism that operates based on distinguishing objects from one another by their sensory properties, should be based upon epistemic operations, if it is to be properly functional. Therefore we can say that Arnauld merits at least a partial, if not a full (albeit implied), recognition of having developed a basic notion of sensory content.

However, perhaps the most remarkable overlap in Matthen's theory and Arnauld's thinking lies in what Matthen calls coevolution. In *SDK: A précis*, we may find the following quote:

> Evidently, though, biological sensory systems are products of evolution, and as such they must have been of some use to their owners. /…/ it is not that sensory systems must converge upon some pre-existing scheme of reality (though they may do this); rather sensory systems and epistemic operations co-evolve so that they are useful together. (Matthen, 2008: 396)

Essentially, sensing and attributing sensory properties to objects exists because it provides the creature that possesses it a biological advantage. For example, a creature which has developed a keen sense of smell, will be able to sense the unpleasant smell of the fruit it is about to eat. The unpleasant smell will cause a certain epistemic operation to happen, for example it may cause a belief that eating that fruit will have negative consequences for the creature. Following this epistemic operation, an effector mechanism will kick in, having the creature avoid the fruit. Presumably, that fruit was toxic, and the survival of the creature is now due to its not eating the fruit because of its sensory property. Another creature whose sense of smell may not be as advanced or exist at all, will eat the fruit and die as a consequence.

Similar notions may be found in Arnauld's work:

> /…/ it presents a simpler means of distinguishing bodies we're observing from each other, or to distance ourselves from those whose presence bothers or obstructs us. It is therefore to accommodate our language to the intents of the Creator, calling bodies white, black or smelly, /…/ (Arnauld, 1986: 211)

One major difference is that Arnauld states that we attribute sensory properties to objects because it is the intention of the creator, which means that according to him, sensory systems, along with epistemic operations, must converge on a pre-existing scheme of reality. This, however, is not a problem, since

Matthen explicitly states that this kind of convergence is permitted. In case of the creature smelling the poisonous, foul-smelling fruit, everything still happens as in Matthen's case: the foul smell triggers an epistemic operation (although it must be noted Arnauld never explicitly describes an epistemic operation – it is simply implied from the notion that an effector mechanism based on sensory input cannot function without an epistemic operation "telling it what to do"), which in turn triggers an effector mechanism, because of which the creature avoids the fruit, surviving because of the biological advantage of being able to sense the smell of the fruit. The one difference is that in Arnauld's case, the biological advantage is not evolutionary, but rather "creationary" in a sense, and sensory systems and epistemic operations do not coevolve, but are rather perfectly coordinated in advance (simply because God has made them so). Again, this is still compatible with Matthen's theory, since he states that sensory systems MAY converge on a pre-existing scheme of reality.

A potentially important feature of Matthen's theory is also the Sensory ordering thesis. Within the Sensory classification thesis, if x is an object, to which we attribute certain sensory property, and F is that property, we can either say simply that "x is F" or at best "x and y are the same with regard to a determinable D". Herein, the determinable is for example a specific shade of a colour or a determinate shape, depending on whether colour vision or shape vision is in question. However, the above-mentioned options lead to the issue of intransivity of indiscriminability. Matthen observes that if x is indiscriminable from y, and y from z, it does not necessarily follow that x be indiscriminable from z. To avoid this, he proposes the sensory ordering thesis, within which the relations between objects and determinables can be formulated as follows "x is more similar to y than to z with regard to determinable D".

Arnauld, however, goes to no such extent. In his thinking an object x is perceived as having sensory property F because it is similar to the idea of F; the idea of F being the impression of that property. Even if it could be argued, admittedly by an extreme long-shot, that the idea of F can be analogous to the determinable D, because it represents a specific shade of colour (in colour vision) as a determinable does (of course we would have to disregard the fact that F is only a mental impression why D is no such thing – hence the long-shot), we would arrive at the fact that there is no way we could possibly imply or infer that Arnauld in any way made the transition between classification and ordering.

Inevitably, there are some other elements of Matthen's theory Arnauld fails to cover or they mismatch with Arnauld's view on them. One such element is the order of priority in which sensory and epistemic processes take place. Matthen quotes:

> A sensory system makes a thing look blue only after it has determined that it is blue. Sense-features such as blue or round must therefore be defined independently of the looks they present. (Matthen, 2008: 393)

That means an object has to actually possess the property we are attributing it, independently from our sensory consciousness. However, the question is, what does this mean? If it is required to be physically blue, if we take the example of blue, then we would arrive at a major mismatch between Arnauld and Matthen's views on the nature of secondary qualities, since Arnauld is deeply opposed to the notion of anything he considers modification of soul (or mind) being attributed to a physical object, and would have to disregard more or less everything we have established so far. However, as we have shown Matthen does not seem to be a proponent of physicalism, therefore the demand for a physicalist definition of colour would directly contradict his earlier stance. This means, the one remaining interpretation must be correct. The only thing that can be physical, be a sufficient condition for something to be considered having a property it relates to, and is not in fact colour (including its phenomenal experience), is what we have earlier called secondary sensory content. These are the physical processes that construe the sensory input, are not independently recognised by the sensory system, but are necessary to cause an impression of a sensory property. As we have shown, Arnauld has a working notion of secondary sensory content, therefore orders of priority of both authors are compatible.

The elements, which Arnauld lacks in his work, are a sensory class hierarchy and the notion of vision reference, as well as the concept of motion-guiding vision.

The former is a proposition that sensory properties are ordered from broader to narrower classification categories – for example from "cool colours" to "blue" to "aquamarine blue". While this is perhaps a helpful concept, it is not one that would be intrinsic to phenomenology or necessary for Arnauld to have in order for his work to be considered anticipatory of phenomenology.

Vision reference denotes "of what" is the visual consciousness. Most authors claim that visual consciousness is of a field, meaning a 2D or 3D array of places that accommodate visual features. Matthen, however, claims visual consciousness is of objects, and visual features are attributed to objects, still or in motion. Arnauld does not make an explicit stance on the matter, but if anything, he might again be closer to Matthen, since his discussion of any visual property is strictly bracketed to the object to which we attribute that property.

Lastly, motion guided vision is the notion of certain types of motion being caused directly by certain sensory experience, without or before us becoming aware of the causal link or even motion itself. Without contemporary grasp of neurology, elaborations on this kind of notions could not have been expected from, for example, Husserl, and much less can they be expected from Arnauld.

After analysing Arnauld's compatibility with specific contemporary author, one might furthermore append to this article Arnauld's capability of addressing a more general contemporary (albeit originally dating back to Locke) concept – the theory of inverted spectrum. The theory proposes a subject who perceives colours normally until one certain point in their lives, when their perception of the colour spectrum suddenly inverts; i.e. green is thereafter perceived as red and vice-versa. Supporters of colours (and indeed qualia in general) as completely non-physical argue that this can happen without any loss of identity of the perceived quality.

This, however, is not the case in Arnauld. While he does claim that colours, as long as they only exist only in our soul (mind), are merely obscure and confused ideas, they are intentional – serve a "God given", or in modern terms evolutionary purpose – and do have a direct and unalterable connection with a physical basis in the extensive substance, and it is only the inexplicability of this connections that renders them obscure and confused. The aforementioned connection, however, is undoubtedly present according to Arnauld, and thus the inversion of colour spectrum is not possible without the loss of identity. In fact, it has grave consequences for the intentionality of colour perception: if the sufferer of the inverted spectrum was previously acquainted with the impression of red, naming it "red", and the impression of green, naming it "green", he will henceforth interpret red as "green" and green as "red", attributing what is indeed green the properties associated with red, as well as the other way around, thereby defeating the evolutionary purpose of colour perception.

However, considering another variety of the theory, where a subject is born with the inverted spectrum, the loss of identity would arguably not happen according to Arnauld. That subject would perceive red things as green, yet still conventionally naming them "red", assigning them properties commonly associated with red, retaining the identity of the colour and intentionality of colour perception.

In any instance, Arnauld's philosophy on colour is evidently more than well enough equipped to consider such problems and answer them in a remarkably contemporary manner.

What we have demonstrated in this article is that in 17th century, Antoine Arnauld undoubtedly had notions and convictions, which remarkably coincide with contemporary phenomenological thought.

He implicitly provides answers to the questions of intentionality and qualitative experience of colours, as well as other less essential issues, such as why (in an evolutionary or, if you will, teleological sense) perceive colours at all. From everything shown in the comparison I believe that Arnauld may be considered an early phenomenological thinker, and certainly should be considered at least an early cognitive thinker, who in his works well anticipated what we today call phenomenology.

Reference list

Descartes, R. (1993). *Les principes de la philosophie*. Paris: Librarie Philosophique J. VRIN.

Arnauld, A. (1986). *Des vraies et des fausses idées.* Paris: Fayard.

Matthen, M. (2008). "Seeing, Doing and Knowing: A Précis". *Philosophy and Phenomenological Research*, vol. 76, no. 2, pp. 392–399.

Schumacher, R. (2008). "Locke on intentionality of Sensory Ideas". *Studies in the History of Philosophy of Mind*, vol. 6, pp. 271–283.

Kant's Cosmopolitanism and the European Union

Toni Milevoj, University of Rijeka

1 Introduction

Immanuel Kant, one of the greatest modern philosophers working in the field of practical philosophy, contributed greatly with his theories in the domain of theoretical philosophy, such as seen in his work *The Critique of Pure Reason.* At the end of the 18th century, Kant wrote *Perpetual Peace,* a work that is considered an epitome of his political philosophy. *Perpetual Peace* will be the main topic of this essay. I will propose arguments that Kant's cosmopolitan ideas, found in his aforementioned work, are implemented in the ideas and practices of the European Union, or that they could be implemented as such. I will approach this topic by comparing the author's cosmopolitanism to the cosmopolitan tendencies of the European Union. Furthermore, I will try to find the answer to one of the fundamental issues: are those ideas compatible with the political functions of the European Union? If they are not entirely compatible, can the European Union acquire the knowledge of some of these theories in the future and implement them, and is it wise to put the theory into practice? By studying the 6 Preliminary Articles on peace and the 3 Definitive Articles of *Perpetual Peace,* I will find those ideas proposed by Kant that can be compared and implemented into the European Union to be useful in the future. Apart from his already mentioned work, I will use several articles that deal with similar topics in the field of cosmopolitanism and Kant to get a wider spectrum of options and knowledge about the articles on perpetual peace.[1] However, some shortcomings of the European Union's politics are to be expected when comparing them to Kant's central cosmopolitan ideas. Although the ideas in *Perpetual Peace* are ahead of their time, the work is not always adequate for an argument in modern times, thus I will try to adapt it to obtain information that is more useful and get finer conclusions. I will elaborate this essay in two parts. In the first part, I will deal with the 6 Preliminary Articles that lead to perpetual peace. More precisely, I will examine which articles are useful for the European Union; I will critically review each article, and propose an argument explaining why I would not consider it necessary for achieving peace. The second part will be focused on the 3 Definitive Articles that are important to follow to obtain perpetual peace. I will compare them with the ideas of the European Union and its function, and explain whether they are already implemented; if not, it will be elaborated whether they

[1] Some of these articles are: Wonicki, R. (2008). "Cosmopolitanism and liberalism: Kant and Contemporary liberal cosmopolitanism"; Cavallar, G. (2012). "Ethics & Global Politics".

can become an idea of the institution. Furthermore, I will examine the possibility of shortcomings that the European Union can have towards Kant's ideas, and see if it is possible to overcome them.

2 Preliminary articles of perpetual peace for modern Europe

It is rather impossible to establish perpetual peace at this point in humanity's evolution. However, it is very interesting to consider it as an idea and try to apply some of the articles of perpetual peace to the modern political status.[2] The impossibility of establishing perpetual peace can be proven in various ways, and although many might agree on this point, there will always be those who stand firmly behind the idea that peace can be ensured. If we apply Kant's idea to the modern world, the lack of unity among countries would cause a problem in the establishment of perpetual peace, besides, that is obvious in the Preliminary Articles of *Perpetual Peace*. Namely, this idea of unity is constantly mentioned in various ways. The reasons for that are the histories of various nations, especially those world leaders that have many worries despite their great power. For example, we can take the case of France after World War I. France was not in a position to maintain world peace according to the ideas of Kant's perpetual peace, but it was rational to keep the superiority in Europe that the French had before the War. Thus, the oppression of the German industry in order to settle the debts of the Great War and the 'ban' on befriending Russians made the French appear as 'egotistic nationalists' of that age who were not capable of withholding peace with their demands. And that is the problem: what kind of a country is not focused on its own interests? Perhaps it would be ideal if the leading forces of Europe were ultimately altruistic; however, it is not rational to be that way. This shows how, along with many other historical events, we can see how some of these postulates are relatively applicable, and they can lead us, in a way, to the right direction and good international relations, and perhaps to a better European Union in the future. Still, other postulates are unrealizable because of the development of Europe since Kant's age or because of the dubiousness of the sole meaning of the postulate.

Regarding the Preliminary articles of perpetual peace, I will isolate only a few applicable ones that Kant mentions in the first part of his work *Towards Perpetual Peace*. These 6 Preliminary Articles of perpetual peace are not as important for the modern construction of the European Union, as some of the articles are outdated, and some of them are already applied, and their most common topic is the war conflict between countries. When I mention outdated articles, I do not consider them completely unusable, but I think that the European Union is overdeveloped to deal with such old-fashioned ideas

[2] There are not a lot of anachronisms, but some of them are sometimes unescapable, so I will try to avoid them as much as I can.

and update its decisions in accordance with these articles. The first and the third article is problematic for peace-oriented politics in my opinion, and I have decided to declare the second and the sixth article old-fashioned for this theme and period because of the constant development of politics since Kant's time. In what follows, I will present two articles, the fourth and the fifth, which I consider to be important and constantly kept, for they can be fatal to the order and necessity for peace in Europe. Certainly, I will try to name other reasons why I consider this valid.

The first preliminary article of perpetual peace:

> No secret treaty of peace shall be held valid in which there is tacitly reserved matter for a future war. (Kant, 1964: 115)

Kant considers that not revealing old pretensions because of tiredness of conflict leads to malevolent intentions to use future opportunities to continue a war or to attack a previous enemy. For Kant, taking advantage of these opportunities is beneath the dignity of a leader, as well as the ministers who are inclined to make such decisions (Kant, 1964: 115–116). This law is restrictive and is one of those that needs to be carried out immediately (Kant, 1964: 119). My problem with this postulate refers to the following question: how to realize that and correct it as a third party? How can I tell someone to reject all pretensions that can lead to a future war? The situation is rather relative; I would perhaps act as Kant; however, I do not have to, which leads to the conclusion that, despite the moral pressure, there are no oppressive laws that would ban us from making peace that would be a motive for a future war. Therefore, if anyone makes peace with a future conflict in mind, there is no one who can prevent that person, except for his or her moral ideals, which don't have to be conflicted with this act. The problem is that there are no 'policemen' who would prevent this before such peace is made, and we are left to the mercy of those who decide to make peace with such pretensions. However, that is not the main counterargument that I wanted to mention. Namely, the problem arises when we include situations with various moral weights into the idea. I believe that Kant would agree that we could not compare a minor war that took a rather small number of victims and a genocide that affected the course of civilization or the state. If we presume that both hypothetical situations ended with a peace treaty, then we must agree that pretensions in the first situation would not be as 'heavy' as those in the case of a genocide would be. It is understandable that it would be immoral, in the first example, to use the situation that Kant describes as undesirable in the first postulate, especially if such a country had not endured such a great oppression, that is, that situation would be more inclined to the other one. In the case of genocide, it would be reasonable to act contrary to Kant: if the country cannot defend its

population, victimized by terrible crimes, for any good reason, not allowing such a crime or the taking of its territory and the oppression of its people, then it is perfectly reasonable to make peace with pretensions to take back their prior territories or to have revenge in the name of the people. I chose two extremities in order to explain the idea and to show that situations cannot be interpreted as 'black or white', which is how Kant's postulate interprets them.

The third preliminary article of perpetual peace:

> Standing armies shall in time be totally abolished. (Kant, 1964: 116)
>
> They constantly threaten with war to other countries, because they are always ready to show themselves as armed for such an undertaking. They force countries to compete between each other in the number of armed men, and since it takes great expenses, peace finally becomes more agonizing than a short war, and those armies become their own cause for aggressive wars that want to achieve a relief. (Kant, 1964: 116–117)

Kant considers that the people that are paid to fight and kill are being used as a means or a tool by someone else. This is a true example of an idea by Kant that is inapplicable in practice. I would agree with this idea when all the countries of the world would put down their arms at the same time, but considering the situation at that time (which can be applied to the modern age as well), it is impossible to find how countries which oppose the Western kind of life would agree with such a decision. If we consider imperialistic tendencies of Western powers in Kant's time, we will realize that non–European countries tried to resist to those powers, which was at that time possible only by using weapons.[3] It is important to mention that it is only possible to make countries give up their weapons by using weapons (since not all countries would welcome this idea, even though their interests would have less moral values than those countries who wish to achieve that goal), which is a paradoxical claim. Furthermore, there is a high probability it would be in conflict with the first preliminary article of perpetual peace.

The second preliminary article of perpetual peace says that independent states cannot come under the dominion of another state by exchange, purchase, inheritance, or donation (Kant, 1964: 116). The sixth preliminary article says that during war, no state will permit acts of hostility that make mutual

[3] Similar things happen today, but such a topic would demand exploring other spheres of research that go further away from Kant's ideas.

confidence in establishing peace impossible. That includes the employment of assassins, poisoners, breach of capitulation, and incitement to treason in the opposing state (Kant 1964, 118).

Some instances of a heavy breach of contract reported to various courts are examples of, according to Kant, dishonourable war tricks and oppressions described in the second and sixth preliminary article, and the United Nations, along with the European Union, made them impossible through their development. Hiring hitmen and poisoners are the kind of actions that are not often used and are rather strange to modern Europe. For other articles of this kind, I will emphasize that their ideas are impossible to apply because they are outdated. Besides, preliminary articles of perpetual piece are not the main idea of this work, but the 3 definitive articles that will follow later.

The fourth preliminary article of perpetual peace:

> National debts shall not be contracted with a view to the external friction of states. (Kant, 1964: 117)

If the purpose of the auxiliary source is to be a mechanism in the internecine wars of power, then this source becomes dangerous because, according to Kant, the national treasury for war costs becomes larger than the treasury of all other countries. If there is a need for economic aid within the country or abroad, then this auxiliary source is not suspicious. An inclination of a country towards such a warfare would question the realization of perpetual peace (Kant 1964, 117). Borrowing money for warfare was a trend during the Great War,[4] and here I must agree with the argument proposed by Kant. Namely, this kind of going into debt is not practical for a country that leads a war because the question arises whether it can settle those debts, especially if it concerns a war of large proportions. There is also the question of interest, which makes repaying the debt more difficult, and, furthermore, the question of winners. If a country that owns money loses a war (that with time only becomes more expensive and more destructive), it is unsure whether it will be able to repay the debt. I took the option of the outcome of war into account because it increases the costs, meaning it includes paying off the reparations.

The fifth preliminary article of perpetual peace:

> No state shall by force interfere with the constitution or government of another state. (Kant, 1964: 118)

[4] International Encyclopaedia of the First World War, *War Bonds*, http://encyclopedia.1914-1918-online.net/article/war_bonds (accessed: 12.02.2017).

A country has no right to do so, and Kant considers this a warning: "/…/ for it is an example of the enormous troubles that a nation brings on itself by lawlessness." (Kant, 1964: 118) Here we can add an example in which a country is divided into two separate parts after inner disturbances and riots. If a country gives aid to any of those two parts, Kant would not consider it interfering into a system of government, however if outer forces interfere with the conflict while it is not resolved yet, then they violate the rights of the people. (Kant, 1964: 118)

3 The three definitive articles of perpetual peace and the European Union

This part of the paper will demonstrate where the European Union is in Kant's spectrum of ideas of perpetual peace and examine where things can be altered or improved in the postulates of the EU, and whether there is an opportunity of development of the EU in direction of the cosmopolitan entity that Immanuel Kant imagines. Lastly, there will be a part about the problematic politics of the European Union, why the EU could not fulfil some of Kant's ideas, and the discussion of whether if it is possible to change the course.

The first definitive article of perpetual peace is that the civil constitution of every country should be a republican one. The constitution is republican when it is based on principles of liberty of all members of the society, and then based on principles of dependence on one legislation, which will stand above the community, and based on principles of equality of all (Kant, 1964: 120). This republican constitution, according to Kant, gives hope to bringing a desirable consequence and perpetual peace, and in its origin contains a pure notion of justice. Kant takes an example of a 'referendum' in which the citizens decide whether there should be a war or not, and depending on the decision of the people, the country goes to war or not. Only a republican system enables people to choose between these two options, and to think about whether it is a smart decision; non-republican constitutions often do not give their citizens the luxury of deciding. According to other constitutions, the chief of state becomes the owner of the country; while in a republican government, the chief of state serves its country with their full power, similarly to despotism. The only difference between these two kinds of government is that republicanism demands that the executive power is detached from the legislative power, and that that kind of a country needs a representative system. (Kant, 1964: 121–123). The republican constitution is based on the social agreement in which citizens become free and equal and create a country in which they can coexist with others in freedom (Wonicki, 2008: 272).

The European Union agrees with this rule. The organization of all members of the EU is republican, although not all member countries have a republican government. According to Kant, a country can have many forms of sovereignty including monarchy, democracy, etc., and it can be republican, which is especially valid for modern, traditional monarchies that can be considered republican countries because they are governed by that principle. Moreover, the country must have a constitution, which is why modern monarchies of the EU are different from various types of autocracies, since they all have a constitution. Thus, Kant's idea matches with the practice of the institution. It should be noted that the legislative and the executive power are separated in modern countries, as well as in Kant's country. (Salikov, 2015: 72–23)

The second definitive article of perpetual peace states that the right of the people should be based on the federalism of free countries. In a natural order, countries are considered as individual, but this state for Kant is a state of independence on foreign law. This way, all countries act to secure their own interests, but they can demand from another country to be included in an order similar to a civil one. Thus, these two countries can secure their own rights, and make an alliance of nations. Kant demands that we are careful in this matter, because this idea does not lead towards a country of the nations, this order can lead to one nation being subjugated by the other, or, in Kant's words:

> /…/ every country, namely, includes a relation of the higher (the lawmaker) towards the lower (The obedient, the people), so multiple nations in one country would constitute only one nation. (Kant, 1964: 124)

Kant thinks that countries must ask their right by war and a fortunate ending of it, but the problem is that a peace treaty does not completely end a war – a war status remains between the countries. When a nation would manage to form a republic that would incline towards perpetual peace, it would become a centre of a federation, the kind of administration that would include other countries, and thus ensuring its freedom. Kant's idea is that this union expands with time. (Kant, 1964: 125–126)

When we talk about the second article of Kant's perpetual peace, we talk about cosmopolitanism. That concept can be considered as an ideological manifestation of globalization, because it allows us the plurality of values, migration, etc. The idea is that all human beings belong to a community towards which they have moral obligations, the moral obligations that they have now to fellow members of their countries. Unlike a country, in this situation the people would have duties towards all the citizens of such a cosmopolitan institution. (Wonicki, 2008: 272)

It seems that the EU agrees with Kant's second condition, because it is an international structure that, together with other member states, establishes the institution. On the other hand, the EU is not a federation, many people consider it an incomplete federal country; however, in its essence it is of a federative nature. Although its legal and political structure cannot be literally described as federative, they contain elements of federation and tendencies to become one (Salikov, 2015: 73). The EU follows certain elements proposed in Kant's perpetual peace, and if we are to consider a possibility of achieving perpetual peace, we must separate three phases that lead us away from war towards a stable peace; negative peace, positive peace and integration. Negative peace excludes the existence of war in that moment. Positive peace is what Kant considers stability, what humanity should aspire to. Integration, one of the principle aims of the European Union, represents a compound of countries that are in such a peaceful state that there is no possibility of war between them. Say that the European Union is among them, although it is still trying to advance in this sense. Despite that, the European Union is more advanced than Kant's idea in this sense, because he did not believe in such a close relation between countries; therefore, if the EU becomes one federal country, it would be above Kant's idea of 'unity of peace'. We have enough reason to believe that a stable peace can be achieved via federation like the one the EU could become, meaning any military threat would be eliminated from that area. However, what is obvious from the EU's development is that the transfer of sovereignty towards federal government does not result in a disappearance of a country, which is contrary to what Kant said, and what I mentioned earlier while introducing Kant's theory of this condition for perpetual peace. Indeed, this transfer makes war seem almost impossible to achieve between countries in such a federation, especially because of the high level of cooperation (Salikov, 2015: 74).

The third definitive article of perpetual peace claims that the rights of the citizens of the world should be limited to common hospitality. Hospitality is the right of every man to step on foreign ground and not feel like an enemy. In peaceful governments, there should be no hostility towards such a person, but if he is considered dangerous and a threat with destructive capabilities, Kant writes that his right of hospitality can be withheld. Namely, no one has more right to be at one place than another (Kant. 1964: 127–128). There are several elements that are contained in the concept of universal hospitality, and those are accepting the autonomy of others, accepting the outer and inner freedom of an individual and respecting the universal law. If we take this into consideration, it would seem irrational for people to start wars, primarily for egotistic tendencies and moral reasons. People can lose their lives and all material belongings, and on the other hand, they do not want to break the universal law (Wonicki,

2008: 274). If they survive they would then suffer the consequences, depending on the outcome of war.

While Kant's theory and EU's practice are somewhat coherent regarding the domestic and international law, the idea of a cosmopolitan right is problematic. In other words, even though the First and the Second definitive article are followed by the European Union, the third one, which includes Kant's laws on hospitality, is not. It could be said that the EU is an internally consistent cosmopolitan federation, but not an external one (Brown, 2014: 687–688).

To understand what are the internal consistent factors of a cosmopolitan federation in Kant's writing, we have to turn to the first two definitive articles mentioned in the earlier part of the paper. As we can see, the First and the Second definitive article concern only the state. The first condition, as stated in the earlier part of the essay, is that the governmental structure should be a republican one. That condition puts emphasis on the external freedom, democratic governance and the rule of law in a state, which, as has already been stated, the Europen Union puts in practice. As it is suggested in Brown's article, the Second definitive article serves to restrict the institutional complexion of the federation. Kant is sceptical of a coercive world state and ultimately rejected such a model for a milder federation of independent states (Brown, 2014: 679). A coercive world government would be useful for practical reasons – to bring a permanent end to war and the expanding association of free and independent states. On the other hand, Kant argues that a state based on these principles would lead to despotism and would endanger freedom. Therefore, there should be an order like a lawful federation under a commonly accepted international right, a permanent congress of states. It is obvious that the EU follows those principles and that every state that wants to join the union has to exercise freedom, democracy, human rights, and the rule of law (Brown, 2014: 680). As it was mentioned in the beginning, it is obvious that the first two articles only concern the governmental structure itself, without turning to the external relations with other governments. However, the last article covers what the first two missed, the external governmental relations, and reveals some problems the EU did not avoid and which are different from Kant's ideas.

The Third definitive article is considered as a cosmopolitan right and is, as mentioned, externally oriented. It provides a legal obligation on individuals and states, regardless of which state the individual belongs to. That is the reason why it is considered cosmopolitan; it considers people as citizens of the world and it proclaims that people have a right to universal hospitality. Surely, a person can be turned away; the reasons are mentioned in the previous paragraphs (Brown, 2014: 683). The

cosmopolitan right is considered as a right of humanity, and it makes it possible for individuals to enter into relations with other people from other countries under a mutual friendly agreement, and that is why it is necessary (Brown, 2014: 684). Now, the laws of hospitality should not be applied only amongst the federated members, but beyond them and the borders encompassing them. We come now across a few problems in the political practice of the EU and the reason why it does not agree with the Third Definitive Article of Immanuel Kant. The European Union has not consistently applied Kant's ideas towards people that enter the territory of the federation. As an example, Brown mentions that:

> /.../ the EU remains inconsistent in its treatment and protection of people who are not EU citizens or who are not considered as being 'European' by various member states. This is often evident in the EU's inconsistent and negligent policies on asylum seekers /.../ and in its blatantly poor record on protecting the Roma people from the discriminatory policies of certain member states. (Brown, 2014: 686)

Because of the inconsistent positions to promote human rights and democracy, there are some irregularities that manifest themselves in many ways, and one of them is the treatment of foreigners. Thus, despite their cosmopolitical elements, the EU is, at the moment, not fulfilling Kant's Third Definitive Article, even though there are glimpses of accomplishment here and there. Because of some successful moments, there may be a bright future for the EU in becoming a manifestation of Kant's ideas, but for now we can only say that there are some similarities amongst the two.

4 Conclusion

Based on the argumentation and with the help of previously mentioned philosophers and their work, we can conclude that the European Union is on the right path to achieve Kant's ideal perpetual peace. Furthermore, in some instances, it is clear that the 'cosmopolitan version' of Europe has not surpassed the author's idea; by further development it could be brought to a contemporaneously adjusted level of a cosmopolitan idea similar to the one of Kant. The former development of the EU is noticeable, according to the preliminary articles of perpetual peace, the analysis of which could assume anachronistic criticism; however, it is necessary to perceive the whole idea of peace as Kant proposed it. Moreover, there is a tendency towards progress, and one of the fundamental criticisms of the Western moral competition is avoided in the three articles of perpetual peace, to which I dedicated most of this paper. It seems that Kant fenced himself from many omissions and errors, which is rare

for the immense ideas of this kind. I believe that a modern, adapted version of this theory will have problems in its further development if it remains close to the principal idea of the author – it is problematic to consider the whole theory to be practiced in reality. Furthermore, the common premise that constitutes perpetual peace can have a solution for modern and future problems of Europe, such as the immigration crisis, although it would have to be put together with the modern conservative ideas of today's Western democracy.

Reference list

Brown, G. (2014). The European Union and Kant's idea of cosmopolitan right: Why the EU is not cosmopolitan. E: Beate, J. *European Journal of International Relations* 20 (3). Retrieved (2017, Oct 5) from http://journals.sagepub.com/doi/10.1177/1354066113482991.

Cavallar, G. (2012). Cosmopolitanisms in Kant's Philosophy. In: Erman, E. ed. *Ethics and Global Politics.* Stockholm: Sweden Research Council.

Kant, I. (1964). *Pravno politički spisi*. Zagreb: Politička kultura.

Salikov, A. (2015). Kant's Perpetual Peace Project and the Project of the European Union. In: Kalinnikov, L.A. ed. *Kant Studies Journal*. Kaliningrad: Imanuil Kant Baltic Federal University Press, pp. 70-78.

Wonicki, R. (2008). Cosmopolitanism and Liberalism: Kant and Contemporary Liberal Cosmopolitanism. In: Čović, A. ed. *Synthesis Philosophica*. Zagreb: Croatian Philosophical Society, pp. 271-280.

Domestic Analogy: A Link between *Spheres of Justice* and *Just and Unjust Wars* by Michael Walzer

Miljan Vasić, University of Belgrade

Abstract

In his *Spheres of Justice*, Michael Walzer argues that society can be understood as a compound of a number of different segments, where each of them has its own distinct meanings and values of social goods. He calls these segments 'spheres of distributive justice', and, as distinct meanings and values point to different procedures of distribution, spheres are separated by criteria and arrangements which are considered appropriate for distribution of certain goods. We should seek to ensure that inequalities in one sphere do not spread into other spheres. Different goods should be distributed for different reasons and according to different procedures, and that is the consequence of cultural relativism and pluralism. Walzer also states that distribution can be just or unjust and that is relative to the social meanings of the goods. Thus, we can define injustice as applying the distributive principle from one sphere of social life to another – such a definition follows from the idea of spheres of distributive justice.

In *Just and Unjust Wars*, Walzer claims that states possess rights more or less as individuals do; therefore, it is possible to imagine a society among states more or less like the society of individuals. This relies on what is called a 'domestic analogy'. The primary form of the theory of aggression, which rests upon a domestic analogy, is called the 'legalist paradigm' by Walzer, and it can be summed up in a number of statements where it is claimed, among other things, that aggression justifies the war of law enforcement against the aggressor state, which can be punished once it has been militarily repulsed. According to Walzer, the implication of the paradigm is clear: if states are members of international society and thus the subjects of rights, they must also be the objects of punishment.

In this paper, however, I will try to show that if we accept Walzer's idea of spheres of distributive justice, maybe we should not accept domestic analogy at the same time. Namely, if the aggressor state is the object of punishment justified by legalist paradigm, then is it the case that distribution of punishment, which is a "negative" distributable good, is an example of applying a distributive principle from one sphere of justice (interpersonal relations) to another (international relations)? If that is the case, then the punishment of the aggressor state is unjust, according to definition of injustice which

follows from *Spheres of Justice*. As the punishment is justified by legalist paradigm, which is, on the other hand, based on domestic analogy, I conclude that a domestic analogy is not compatible with Walzer's idea of spheres of distributive justice in social life.

Key words: Walzer, distributive justice, domestic analogy, legalist paradigm

1 Introduction

First edition of Michael Walzer's *Just and Unjust Wars* was published in 1977 and, as the name suggests, this book deals with a just war problem, while *Spheres of Justice* was published in 1983, and in this book Walzer advocates a pluralistic theory of social justice. These two books deal with different issues, and "justice", which is present in both titles, doesn't necessarily need to be understood in the same way, but I think it is possible to bring them together. In this paper I will not be considering Walzer's arguments regarding theory of just war itself; instead, I will look upon his conception of states' rights, as that will be essential for my argument. More precisely, I will try to show that certain ideas supported by author in *Just and Unjust Wars* could be considered as incompatible with some other ideas he supports in the second book. These ideas rest upon a conception of 'domestic analogy', and I will try to show that its acceptance is in conflict with some of the general notions of *Spheres of Justice*.

2 *Spheres of Justice*

In *Spheres of Justice*, Michael Walzer argues that it is impossible to form a universal theory of justice[1] – such a theory would certainly ignore the differences among histories and cultures of various communities, because there cannot be a perspective external to the community. The only way to estimate whether something is just or not is to see how each particular community understands its value of social goods. Therefore, a society is just if it acts in accordance with the shared understandings of its members, which are shown through characteristic practices and institutions (Kymlicka, 2002: 211; Walzer, 1983: 3–5). Walzer states that justice is a human construction, and thus cannot be made in only one way. His theory is, therefore, pluralistic: different goods should be distributed for different

[1] He explicitly criticizes Rawls's theory: if the conceived ideally rational individuals were forced to choose impartially, knowing nothing of their own situation, they would probably choose one and only one distributive system; however, it is doubtful whether those same individuals, if they were transformed into ordinary people with a sense of their own identity, with their own goods in their hands and caught up in everyday troubles, would repeat that hypothetical choice or even recognize it as their own (Walzer, 1983: 5).

reasons and according to different procedures, and that is the consequence of cultural relativism and pluralism. To identify the principles of justice would be more a matter of cultural interpretation rather than a philosophical argument (Kymlicka, 2002: 211; Walzer, 1983: 5–6).

Walzer argues that social life can be conceived as compounded of a number of different segments, where each of them has its own distinct meanings and values of social goods and thus corresponds with a different distributive principle (Walzer, 1983: 7–8). He calls these segments *spheres of distributive justice*, and, as distinct meanings and values point to different procedures of distribution, spheres are separated by criteria and arrangements which are considered as appropriate for distribution of certain goods (Walzer, 1983: 10).

We cannot achieve justice by trying to create a system which will eliminate inequalities between goods and individuals; instead, we should seek to ensure that inequalities in one sphere do not spread into other spheres. For example, wealth, which plays a major role in a sphere of market, should not play any role in spheres of health care, education, political power, etc. (Kymlicka, 2002: 211; Walzer, 1983: 97–103). Walzer calls this system of distribution a *complex equality*.[2] He also states that distribution can be just or unjust, and that is relative to the social meanings of the goods (Walzer, 1983: 9). Thus, we can define injustice as applying the distributive principle from one sphere of social life to another – such a definition follows from the idea of spheres of distributive justice, and will be of high importance for the latter argument.

3 Domestic Analogy

In *Just and Unjust Wars*, Walzer writes that two fundamental rights which belong to states are political sovereignty and territorial integrity. Those rights, however – although they belong to states – derive ultimately from the rights of individuals, from whom they take their force; therefore, states' rights are a collective form of individual rights (Walzer, 2006: 54). Let us assume that every man has a right to defend his own home. From this point, we can come to their right to defend a territory, because territory is owned collectively as the homes are owned privately. The right to territorial integrity is thus derived from the individual right to property.[3] Similar to this, there is also a resemblance between individual

[2] E.g. person X may hold an important political office, while person Y holds no political office; thus, they would not be equal in terms of political power. However, as long as X's office doesn't give him any advantage over Y in any other sphere, they would not be unequal generally. Therefore, they are considered equal, and that is the example of complex equality (Walzer, 1983: 19–20).

[3] Walzer suggests another way to derive a right to territorial integrity from a right to defend a home, which doesn't include a right to property. If a state has ownership over a vast reaches of land, it could be problematic to derive a right to defend it from a right to defend a home, unless its large territory can be somehow tied to the requirements of national survival and

liberty and political sovereignty, and therefore comes the right to defend the latter, as men already have the right to defend the former (Walzer, 2006: 54).

Now, if the states possess rights more or less as individuals do, then it is possible to think of a *society among states* more or less *like the society of individuals*. Pointing to an aggression as the equivalent of robbery or murder, comparing home to a territory, or individual liberty to political sovereignty, relies on what is called a 'domestic analogy'.[4] Our understandings and judgments of aggression are, therefore, products of reasoning by analogy which, when made explicit, makes the world of states take on the shape of political society; its character is then accessible through such notions as crime and punishment, self-defense, law enforcement, etc. (Walzer, 2006: 58).

However, the question arises whether such analogy is acceptable. Walzer admits that the language of international law is strangely poor – every violation of the political sovereignty or territorial integrity of an independent state is called aggression: the equivalents of assault, armed robbery, extortion, murder in all degrees, etc., have but one name. Thus, it is impossible to distinguish between various levels of aggression by classifying them into different categories. In this sense, the seizure of a piece of land and the complete destruction of a state's independence are regarded as the same crime (Walzer, 2006: 51–52). Walzer also explains this: what all aggressive acts have in common is that they include the use of force, but force cannot be used between states as it often can be between persons, without putting life at risk. "Aggression is a singular and undifferentiated crime because, in all its forms, it challenges rights that are worth dying for." (Walzer, 2006: 53)[5]

4 Legalist Paradigm and Locke's theory

political independence. But these two by themselves generate territorial rights which don't need to be connected with a right to property, just as a man who lives in a home which he doesn't own still has a right to defend it, if his life or liberty is endangered (Walzer, 2006: 55).

[4] The conception of domestic analogy originates from interpretations of Hobbes, who, in *Leviathan*, after introducing his notions of a state of nature, argues from war of all against all among individuals to such war among different states (or, more precisely, their rulers) (Hobbes, 1998: 85). However, such interpretations, which imply that international relations are the same as interpersonal ones, could be inconsistent with Hobbes's notion of equality among all men, as states are apparently not equal in a way that even the weakest one could (alone or united with another one) destroy the strongest one. It would also be absurd to say that a state could sacrifice its liberty (or rather sovereignty) in order to gain more security – for a good survey on the subject see: Williams (1996), Grewal (2016).

[5] However, it is exactly for the fact that not every crime, which one person might commit on another, puts one's life in danger, we can say that the analogy between aggression and everyday crimes is not applicable. For example, armed robbery puts one's life in danger, while burglary doesn't have to, but they are both violation of right to property. As the state's right to territorial integrity is understood by analogy with the right to property, it would follow that we can divide violations of territorial integrity into those which put life at risk and those which don't, but according to Walzer this is not the case. Therefore, this Walzer's explanation seems to point to unfitness of the analogy, rather than defending its usage.

The theory of aggression first takes shape in accordance with domestic analogy. Walzer calls this primary form a legalist paradigm, since it reflects the conventions of law and order, and it can be summed up in six statements (Walzer, 2006: 61–63):

1) There exists an international society of independent states.
2) This international society has a law that establishes the rights of its members – above all, the rights of territorial integrity and political sovereignty.
3) Any use of force or imminent threat of force by one state against the political sovereignty or territorial integrity of another constitutes aggression and is a criminal act.
4) Aggression justifies two kinds of violent response: a war of self-defense by the victim and a *war of law enforcement* by the victim and *any other member of international society*.
5) Nothing but aggression can justify war.
6) Once the aggressor state has been militarily repulsed, *it can also be punished*.

In Locke's political theory, a man who violates the law of nature can be punished by any other member of the community. As men are often incompetent to carry out justice (they are not adequate to judge their own acts, nor those of their friends and enemies), Locke says that in order to be just, punishment must be administered by a universal magistrate, who can objectively judge any violation of the law of nature (Locke, 1988: 271–275).

It seems that the legalist paradigm rests upon the analogy to Locke's conception of just society[6] – Walzer himself also states this (Walzer, 2006: 209, 221). The fourth and sixth statement are the best examples. In the fourth one, it is claimed that war by law enforcement is just; it also allows other members of international society to join the war on the victim's side. Walzer says this is best understood by analogy to the decision of an individual to help someone who is attacked on the street (Walzer, 2006: 62).[7] Once the aggressor state has been militarily repulsed, it can also be *punished*. There could be various reasons for the punishment, but most common ones are to deter other states

[6] Locke's law of nature forbids man, unless it is to do justice on an offender, to take away or impair life, or what tends to preservation of the life, liberty, health, body or goods of another (Locke, 1988: 271). In this sense, violation of political sovereignty or territorial integrity of a state could be compared to violation of liberty or body. It could also be possible to form an independent institution to play a role of Locke's magistrate in international society (e.g. United Nations Security Council).

[7] Locke's law of nature, aside from forbidding one man to assault another, also obliges him to preserve others' lives (Locke, 1988: 271). This too points to analogy between legalist paradigm and Locke's theory.

from aggression or to reform the defeated one.[8] The question remains whether the state as a whole or only particular persons are the objects of punishment, but the implication of the paradigm is clear: if states are members of an international society and thus the *subjects of rights*, they must also be the *objects of punishment* (Walzer, 2006: 63).

5 Take on Domestic Analogy

If the aggressor state is an object of punishment justified by legalist paradigm, then is it the case that the distribution of punishment, which is also a distributable good, is an example of applying a distributive principle from one sphere of justice (interpersonal relations) to another (international relations), i.e. the interference of different distributive spheres? [9]

If that is the case, then the punishment of the aggressor state is unjust, according to the definition of injustice that follows from *Spheres of Justice*. As the punishment is justified by the legalist paradigm, which is, on the other hand, based on domestic analogy, we can say that a domestic analogy is not compatible with Walzer's idea of spheres of distributive justice.

I will try to clarify this in the following argument:

(i) Interpersonal and international relations are two separate spheres of social life.

(ii) If domestic analogy is acceptable, then international relations are relevantly similar to interpersonal ones.

(iii) Applying a distributive principle from one sphere of social life to another is unjust.

(iv) Punishment is a distributable good in interpersonal relations.

(v) If international relations are relevantly similar to interpersonal relations, then a distribution of punishment in international relations is just.

(vi) If punishment is a distributable good in interpersonal relations, then its distribution in sphere of international relations would be unjust (from (i) and (iii)).

(vii) Distribution of punishment in international relations is unjust (from (iv) and (vi)).

(viii) International relations are not relevantly similar to interpersonal ones

[8] And here is another link between this and Locke's theory, since Locke argues that the main purposes of punishment are restraint and reparation (Locke, 1988: 273).

[9] It could be objected here that punishment is governed by principle of not distributive but retributive justice. However, Walzer explicitly states that punishment is a distributive good and that it should be distributed according to the need of the injured party (Walzer, 1983: 9, 65–6). Distribution of punishment is therefore not substantially different from distribution of health (Walzer, 1983: 66), although they belong to different spheres. While Walzer doesn't mention retributive justice, it seems that it should be, at best, considered as a subtype of distributive justice.

(from (v) and (vii)).

(ix) Therefore, domestic analogy is not acceptable (from (ii) and (viii)).

Now I will consider the validity of premises, starting with those less problematic. In premise (ii), it is explicated what domestic analogy is. Premise (iii) is a definition of injustice, which follows from *Spheres of Justice*. Regarding premise (iv), Walzer classifies punishment as a distributable good, although he places it into a group of so-called "negative" goods (Walzer, 1983: 9, 85). I think that the validity of (v) rests upon the legalist paradigm (especially statements 4) and 6)) and its analogy to Locke's conception of society. Also, in premise (iv), punishment is considered a distributable *good*, whilst (iii) is about interfering distributive *principles*. Walzer explicitly states that different goods should be distributed for different reasons and according to different procedures (Walzer, 1983: 6). Although this does not have to imply that the same goods must be distributed for same reasons and according to same procedures, it seems that this is what Walzer has in mind.

I have not considered the validity of the first premise above, since I regard it as the most problematic one. Could interpersonal and international relations really be taken for separate spheres? It should be obvious that, for example, education and health care represent two different spheres, but the given problem, I think, is much more complicated. Perhaps a possible answer could be to assume that society of states is the *society* itself; in that case, a comparison between interpersonal and international relations would actually be the comparison between relations among members of one society and among members of another society, the latter being a society of states. But society itself *is not a sphere*; it is instead composed of a number of them. And when we talk about multiple societies, that is exactly when Walzer allows a plurality of principles, and that could be one of the possible objections to the argument.

6 Conclusion

Regardless of whether the argument is tenable or whether it has fulfilled its negative goal (showing inconsistency in Walzer's work), I think it has marked out a positive one, namely, to point out that Walzer's conception of justice from his different works can be brought together into something that can be regarded as Walzer's comprehensive theory of justice. Also, the argument itself could have taken the opposite direction: by making Walzer's theory of just war a primary one, it is possible, by using the similar reasoning, to conclude that his theory of distributive justice is untenable. Another possible direction is to reject domestic analogy by, for example, stating that the punishment of the state might include the punishment of innocent individuals and, by keeping premise (ii), conclude that interpersonal and international relations are not relevantly similar and therefore should be considered as separate spheres. However, I don't think this strategy would be the preferred one, since it would require, as a premise, the rejection of the fundamental postulation of *Just and Unjust Wars* – then the result that Walzer's position is untenable one would not be surprising. In the end, I will point out another way of how the two books could be linked by domestic analogy: In *Just and Unjust Wars*, Walzer criticizes the idea of preventive war (Walzer, 2006: 74–85), while in *Spheres of Justice*, he criticizes preventive detention (Walzer, 1983: 271–272). If these two are somehow related, they are, at any rate, related by domestic analogy, which, in this case, does not have to lead to inconsistencies.

Reference list

Grewal, D. (2016). "The Domestic Analogy Revisited: Hobbes on International Order**".** In: *The Yale Law Journal*, Vol. 125, No. 3, pp. 560–795.

Hobbes, T. (1998). *Leviathan*. Oxford: Oxford University Press.

Kymlicka, W. (2002). *Contemporary Political Philosophy: An Introduction*. Oxford: Oxford University Press.

Locke, J. (1988). *Two Treatises of Government*. Cambridge: Cambridge University Press.

Walzer, M. (2006). *Just and Unjust Wars*. New York: Basic Books.

Walzer, M. (1983). *Spheres of Justice*. New York: Basic Books.

Williams, M. C. (1996). "Hobbes and International Relations: A Reconsideration". In: *International Organization*, Vol. 50, No. 2, pp. 213–236.

Is It Really Vitally Important for Us to Accept Most Significant Propositions as True? Our Disagreements Seem to Say No

Tim Koprivnik, University of Maribor

Abstract

Intuitively, most people (let alone philosophers) would, if somebody asked them, "Is accepting most significant propositions as true (rather than false) vitally important for us?", reply with a clear affirmative answer; but in this article, it will be argued that this is not the right answer – that is to say, it will be argued that *accepting most significant propositions as true is not vitally important for us*.

The course of the article is as follows: In the first part of the article, (i) main concepts, with which we will operate, are defined – that is, 'truth' is defined as "the quality of those propositions that accord with reality, specifying what is in fact the case"; 'proposition' as "a statement in which something is affirmed or denied, so that it can therefore be significantly characterized as either true or false"; 'significant proposition' as a proposition which seems to play a special role in our lives, that is to say, a proposition which is of scientific, philosophic, religious, ethical, social, or political nature; and 'vital importance' is defined in the following way: X is of vital importance for us if we cannot survive/live without it.

In the second part of the article, the argument for the above-mentioned claim is presented:

1. If accepting some significant proposition as true is vitally important for us, then we should expect that this significant proposition would also be (almost) universally accepted as true.
2. Most significant propositions are not universally (nor almost universally) accepted as true.
3. Therefore: Accepting most significant propositions as true is not vitally important for us.

Next, both premises in the argument are given justification. The first premise is justified with the following *reasoning by analogy*: it could be said that because we all eat, sleep, reproduce (have sex), live in a society, tell lies, gossip, show kindness and compassion, have friends and/or partners, and strive for reputation (social status/influence/respect), accepting/perceiving these things/propositions (for example, the proposition "Human beings need water in order to survive.") as true must be vitally important for us. But following the same logic, we should expect the same will also be the case with

accepting all other significant propositions – if accepting them as true is vitally important for us, we should expect that (almost) all of us also accept them as true.

The second premise is justified with a simple empirical finding: We disagree about almost all significant propositions. In other words, there is no universal (nor almost universal) agreement about them.

Finally, in the third part of the article, some general thoughts about this topic are presented.

Key words: truth, significant propositions, vital importance, evolution, psychology, epistemology

1 Introduction

Suppose that there are infinitely many true propositions (in a more or less traditional sense of corresponding to the matters of fact). For most of them, we all seem to agree that they are insignificant for our lives – some examples of them would be the following: "Today I ate eggs for breakfast.", "Michael Jordan was born in 1963 on 17th of February.", "The TV series 'Lost' consists of 6 seasons.", "My phone number is 051 xxx xxx.", etc. – however, we think that some other propositions are significant – that is to say, they seem to play a special role in our lives. Some examples of them would be: "The universe is 13.8 billion years old.", "God does not exist.", "Humans were not created by God but evolved from earlier species.", "Climate change is real and human-induced.", "Vaccinations are safe to use.", "Astrology is a load of rubbish.", "There is no soul that is separate from the body.", "There is no afterlife, heaven and hell.", "We don't have free will, but we should still be punished for our offences.", "People will mostly act selfishly if they have a chance to do so.", "The best educational system is that of Finland.", "Societies function better if the gap between the rich and the poor is not (too) big.", "All people do not have the same chances to succeed in life.", etc.

Obviously, not everybody accepts these significant propositions as *true*. In fact, when it comes to matters of science, philosophy, art, politics, or everyday life, we can hardly find any universal agreement among us at all (more about that in section 2.2.2.2). And here comes the catch – if that is really the case, then it seems that accepting/perceiving significant propositions as true (rather than false) is *not* vitally important for us at all. At least that's what I will try to argue.

But first of all, let me start with explaining the main concepts that will be used in the argumentation.

2 Body

2.1 Conceptualization

What I had in mind when I first mentioned ‘true propositions’ was the following definition of ‘proposition’:

> a statement in which something is affirmed or denied, so that it can therefore be significantly characterized as either true or false. (Dictionary.com, 2017)

And the following definition of ‘truth’:

> the quality of those propositions that accord with reality, specifying what is in fact the case. (Audi, 1999: 929)

When I am using the word ‘significant proposition’, I am adhering to the following definition:

> a proposition which seems to play a special role in our lives, that is to say, a proposition which is of scientific, philosophic, religious, ethical, social, or political nature.

Lastly, my definition of the concept ‘vital importance’ is defined as:

> X is of vital importance if we cannot survive/live without X. (Almost) universal acceptance is a necessary condition for vital importance.

You may be wondering what vital importance has to do with survival/living and universality, but the relation among them does exist if we are familiar with the basic *evolutionary logic*, which goes like this:

> (1) If X is *(almost) universal*, then it *doesn’t* necessarily mean/follow that X is also *vitally important*; but on the other hand;
>
> (2) if X is *vitally important*, then it *does* necessarily mean/follow that X is also *(almost) universal*. In other words, if X is *not (almost) universal*, then it *does* necessarily mean/follow that X is also *not vitally important*.

There are many examples that prove this kind of evolutionary logic and, therefore, the above definition of vital importance: drinking, eating, sleeping, reproducing (having sex), living in groups/societies, lying, gossiping, being self-interested, showing kindness and compassion to others, having friends

and/or partners, striving for good reputation (social status, influence, respect). Why? Because these things are *universal* – we all do, have, or at least strive for them, and none of us can really *survive/live* without them. Hence, they are vitally important for us.

2.2 Argumentation

2.2.1 General argument

The argument for the claim "Accepting most significant propositions as true is not vitally important for us." is as follows:

1. If accepting some significant proposition as true is vitally important for us, then we should expect that this significant proposition would also be (almost) universally accepted as true.
2. Most significant propositions are not universally (nor almost universally) accepted as true.
3. Therefore: Accepting most significant propositions as true is not vitally important for us.

Logically speaking, this a valid argument; but because formal logic is only concerned with (valid) *forms*, the *content* (i.e., both premises) in this argument can still be false or unjustified. The next reasonable step is thus to try to justify both premises.

2.2.2 Justification of the argument

2.2.2.1 Justification of the first premise

How can we justify the first premise? With the following *reasoning by analogy*:

> Because we all drink, eat, sleep, reproduce (have sex), live in a society, tell lies, gossip, are self-interested, show kindness and compassion to others, have friends and/or partners, and have some kind of a reputation (and strive for even bigger reputation); these things – and we could also say, *accepting/perceiving* these things/propositions (for example, the proposition that "Human beings need water in order to survive") as true – must be vitally important for us. But, following the same logic, we should expect – and here comes the analogy – the same will also be the case when accepting all *other* significant things/propositions – if they are vitally important for us, *we should expect that (almost) all of us also accept them as true*.

But is that really the case – are most significant propositions really (almost) universally accepted as true among people? Let's see.

2.2.2.2 Justification of the second premise

As we could have seen from the argument, the second premise says that most significant propositions are *not* (almost) universally accepted as true. But how can we justify this premise?

Quite simply: by showing that *we disagree about almost everything* when it comes to matters of science, philosophy, art, ethics, politics, or everyday life. Which means that people accept very different significant propositions as true.

Let me give some examples of matters in science, philosophy, art, ethics, politics, or in everyday life, that we don't agree on:

> How old is the universe? Does God exist? Are we created or evolved? Is global warming real and human-induced? Are vaccines safe to use? Does the soul exist? Is there an afterlife, heaven and hell? Do we have free will? Is free will compatible with determinism? Are we morally responsible (and/or can be punished) for what we do, if we don't have free will? Can we trust our intuitions about the world and ourselves? Is consciousness an illusion? What is 'consciousness'? Are animals conscious? What, if anything, is the meaning of life? Are people selfish, altruistic, or both? Is morality just a figment of our imagination? What is the best educational system? What is the best economic system? How should we organize societies? Are elites ('bourgeoisie') conspiring against us ('proletariat')? Can we see the world as it is? What is the nature of reality? Is math discovered or invented? Is there such a thing as analytic-synthetic distinction? What counts as 'art'? Is beauty in the eye of the beholder or is it objective?

I do have answers to all of these questions, and I also think that everyone who disagrees with me is quite probably wrong. But does it matter whether (I think) others are wrong? If it would, then there shouldn't be a lot of (if any) people who think wrongly. Well, let us look at some numbers: 62% of people in the world consider themselves as religious (and only 9% as atheists), 74% think they have soul that is separate from the body, 71% believe in God, 56% believe in heaven, 54% in afterlife, and 49% in hell (*WIN/Gallup International*, 2017); 42% of Americans think the theory of evolution is false (*Gallup*, 2014); 66% of Americans think creationism is (probably) true (*Gallup*, n.d.); 25% of Americans don't know whether the Earth revolves around the Sun or whether the Sun revolves around the Earth, and 52% don't know that human beings evolved from earlier species of animals (*National Science Foundation*, 2014); 41% of Americans think antibiotics can effectively treat viral infections (*YouGov*, 2015); 51% of people think free will is compatible with determinism, and (at least) 15% of

people believe we don't have free will, 34% of people believe beauty is objective (in reality) and 41% that it is in the eye of the beholder, 58% of people think there is an analytic-synthetic distinction and 29% don't (The *PhilPapers Survey*, 2009).

Quite big disagreements, would you not agree? But things do not stop here, because other kinds of disagreement among people are also possible – that is, we can agree about the facts, but disagree about the *interpretation* of these same facts. In other words, we (can) see the same facts/things differently, that is, through different perspectives (belief systems, points of view, worldviews, standpoints, approaches, theories) – which basically means that we don't agree on what we're seeing.

Let me give you some examples: Metaphysical realists, mind representationalists, moral realists, liberals, Marxists (socialists/communists), feminists, and life optimists see the world quite differently than metaphysical idealists, mind anti-representationalists, moral anti-realists (e.g., nihilists, political realists), conservatives, proponents-of-capitalism, anti- or non-feminists, and life-pessimists.

To be even more specific: I, for one, cannot see how the theory of evolution is compatible with the idea of God, but 31% of Americans think it is, since they interpret the same fact (i.e. that we have evolved) through a creationist's point of view – they believe we have evolved but due to God's planning (*Gallup*, 2014). But maybe the best and easiest way for one to understand how people with different political (e.g. liberal and conservative) worldviews see the same historical, economical, and/or political facts differently, one should *see* Haidt (2014) and Lakoff (2002).

To continue, *we also disagree about the methodology*, i.e. the ways/methods of deciding whether some significant proposition is true or false. For instance: *empiricists* believe that the right way of deciding whether some significant proposition is true or false is through our senses, whereas *rationalists* believe it is through pure reason, *fideists* (or *mystics*) through faith, *scientismists* through scientific method; and there are also *sceptics* and *postmodernists*, who believe that there is no good/right method. In philosophy, there are generally two methodological 'schools' – *analytic philosophers*, who believe that the right way of deciding whether some significant proposition is true or false (or at least justified or unjustified) is through logical reasoning, conceptual analysis, and scientific facts,[1] and *continental philosophers*, who dislike the methodology of analytic philosophers and, therefore mostly use non-scientific methods, such as hermeneutics, phenomenology, and psychoanalysis. A similar division

[1] Even among analytical philosophers there is disagreement about the methodology. For example, *see* Šuster (2016: 22–24).

exists in social sciences (e.g., in sociology and psychology) where you have *positivists*, who believe the right way of seeking psychological and/or sociological truths is through scientific method (i.e., through surveys, observations, experiments, statistics), and *interpretivists* (or *anti-positivists*), who believe that we cannot adequately explain the inner nature of social phenomena through the scientific method, and that the right way of deciding whether some significant proposition is true or false is through some kind of an in-depth understanding of the phenomena.[2]

Finally, *we are all dogmatic and self-serving biased*. Whether we admit it or not, we all think we are right (that we accept significant propositions as true) and others, who don't think the same as we do, are wrong (and therefore do not accept significant propositions as true). Greene (2002: 94) puts it nicely: "We all believe we are masters in the realm of opinions and reasoning." The best example of this is once again in politics, where you have liberals and conservatives, who both think that the other side is biased, but the reality is, as Ditto et al. (2017) have shown, that both sides are biased. But dogmatism and confirmation biases do not exist only in the political domain of life, but in every domain of life, because, if for no other reason, it is in principle impossible to argue for or against something without at least some faith (Bentley Hart, 2015), and/or because it is in principle impossible to avoid confirmation bias where you have your own opinion on the matter in question. In addition, results from Stanovich's and West's (2008) have showed that no matter how smart you are (i.e. how high is your 'cognitive ability'), you are still going to biased in your favour. Finally, a recent empirical work has also shown that if you have more unrealistically positive self-appraisals and beliefs, you are more likely to be optimally mentally healthy (Taylor & Brown, 1994; *see* also McKay & Dennett, 2009). To oversimplify: self-deception is useful and common.

3 Conclusion

What does this all mean?

Firstly, it seems that this finding – that accepting most significant propositions as true (rather than false) is not vitally important for us – gives a simple explanation as to why a lot of people don't know a lot of things (and/or why there are such big disagreements among people) when it comes to matters

[2] Please note that this presentation of different methodologies is somehow simplified. Additionally, I might have presented some methodologies unfairly. Like all, I do not know everything and am biased too. If you want to learn more about the differences between these methodologies, *see* Jones (2009) and "Positivism and Interpretivism in Social Research" (2015). Furthermore, if you want to learn more about different kinds of disagreement among people, see Elgin (2012).

of science, philosophy, religion, ethics, or politics. The explanation is that, from an evolutionary perspective, it just does not make any big difference if we know these things or not.

Secondly, I think that this finding also supports the idea – which a lot of people still have problems accepting it – that we are not rational, passive information processing machines (simply put, computers) that make cold-blooded decisions based on facts, but rather active, dynamic, mad, sentient creatures who are constantly being pushed by our environments, our changing perceptions, our desires, our passions, and our instincts. "It's not facts that move us, it's our desires." (Asacker, 2014)

Finally, this finding also seems to fit very well with Mercier's and Sperber's 'argumentative theory', which states that we shouldn't see human reason(ing) from an 'intellectualist' perspective, but from a social 'interactionist' one, which means that we should not think of reason as a tool that helps us to arrive at truths, to "solve abstract, logical problems or /…/ helps us draw conclusions from unfamiliar data" (Kolbert, 2017), but rather as a tool that helps us *justify/rationalize* our beliefs and actions to others and, more importantly, *convince/persuade* others – through argumentation – that we are right and that others, who disagree with us, are wrong (Mercier & Sperber, 2011: 57).

Reference list

Asacker, T. (30. 6. 2014). *Why TED Talks don't change people's behaviors* [Video]. Retrieved (2017, March 16) from <https://www.youtube.com/watch?v=W0jTZ-GP0N4>.

Audi, R. (1999). *The Cambridge University of Philosophy, Second Edition*. New York: Cambridge University Press.

Bentley Hart, D. (March 2017). "Reason's Faith". *First Things.* Retrieved (2017, April 21) from <https://www.firstthings.com/article/2015/03/reasons-faith>.

Ditto, P. H. et al. (2017). "At Least Bias Is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives". *SSRN*. Retrieved (2017, April 21) from <https://ssrn.com/abstract=2952510>.

Elgin, C. Z. (2012). "Begging to Differ". *The Philosopher's Magazine,* 59(4), pp. 77–82.

Gallup. (n.d.). "Evolution, Creationism, Intelligent Design." Retrieved (2017, April 15) from <http://www.gallup.com/poll/21814/evolution-creationism-intelligent-design.aspx>.

Gallup. (2014). "In U.S., 42% Believe Creationist View of Human Origins". Retrieved (2017 April 15) from <http://www.gallup.com/poll/170822/believe-creationist-view-human-origins.aspx>.

Greene, R. (2000). *The 48 Laws of Power.* London: Profile Book Ltd. Retrieved (2017, March 13) from

<http://cnqzu.com/library/Anarchy%20Folder/Political%20and%20Government/The%2048%20Laws%20of%20Power%20-%20Robert%20Greene.pdf>.

Haidt, J. (2014). *Three Stories about Capitalism* [Video]. Retrieved (2017, April 3) from <https://www.youtube.com/watch?v=iOu_8yoqZoQ>.

Jones, K. (2009). "Analytic versus Continental Philosophy". *Philosophy Now*. Retrieved [2017, June 3] from <https://philosophynow.org/issues/74/Analytic_versus_Continental_Philosophy>.

Kolbert, E. (27. 2. 2017). "Why Facts Don't Change Our Minds". *The New York Times*. Retrieved (2017, March 12) from < http://www.newyorker.com/magazine/2017/02/27/why-facts-dont-change-our-minds>.

Lakoff, G. (2002). *Moral Politics: How Liberals and Conservatives Think*. Chicago: University of Chicago Press.

McKay, R. & Dennett, D. (2009). "The Evolution of Misbelief." *Behavioral and Brain Sciences*, 32, pp. 493–561.

Mercier, H. & Sperber, D. (2011). "Why do humans reason? Arguments for argumentative theory". *Behavioral and Brain Sciences*, 34, pp. 57–111.

National Science Foundation. (2014). "Chapter 7. Science and Technology: Public Attitudes and Understanding". (2014, February 14). Retrieved (2017 April 15) from < https://www.nsf.gov/statistics/seind14/index.cfm/chapter-7/c7h.htm#s1>.

Positivism and Interpretivism in Social Research. (2015, May 18). Retrieved (2017, June 3) from <https://revisesociology.com/2015/05/18/positivism-interpretivism-sociology/>.

proposition. (2017). In *Dictionary.com*. Retrieved (2017, September 30) from <http://www.dictionary.com/browse/proposition>.

Stanovich, K., and West, R. F. (2008). "On the failure of cognitive ability to predict myside and one-sided thinking biases". *THINKING & REASONING*, 14(2), pp. 129–167. Retrieved (2017, April 21) from <http://keithstanovich.com/Site/Research_on_Reasoning_files/SWTandR08.pdf>.

Šuster, D. (2016). "Odločnost in spravljivost" ["Steadfastness and Conciliation"]. *Analiza*, 20(1), pp. 11–26.

Taylor, S. E. & Brown, J. D. (1994). "Positive illusions and well-being revisited: Separating fact from fiction". *Psychological Bulletin*, 116(1), pp. 21–27.

The Philpapers Surveys (2009). Retrieved (2017, October 8) from <https://philpapers.org/surveys/results.pl?affil=All+respondents&areas0=0&areas_max=1&grain=coarse>.

WIN/Gallup International (2017). “Religion prevails in the world.” Retrieved (2017 April 15) from <http://www.gallup-international.bg/en/Publications/2017/373-Religion-prevails-in-the-world>.

YouGov. (2015). “A third of Americans think antibiotics cure the flu.” Retrieved (2017 April 15) from <https://today.yougov.com/news/2015/01/20/young-people-overestimate-usefulness-antibiotics/>.

Khalidi: The Problems of Crosscutting Categories in Biology

Urška Martinc, University of Maribor

Abstract

In this article, we will analyse the problems of crosscutting between categories. These problems will be analysed using examples from biology. We will help ourselves mainly with the works of Muhammad Ali Khalidi and Ian Hacking.

The question that Muhammad Ali Khalidi poses in his work, *Natural Kinds and Crosscutting Categories* (1998), is the following: Do the systems' categories crosscut? A sub-question is this: If they do crosscut, under what conditions is the crosscutting done (Khalidi, 1998, p. 33)? Philosophers Hacking and Thomason argue that such systems, i.e. systems that crosscut each other, cannot exist as natural systems or natural kinds (Khalidi, 1998: 33–35). In the first part, we will therefore analyse the following question: Can crosscutting categories count as natural kinds?

In the second part, we will try to find an answer to the following question: Are natural kinds arranged in a hierarchy? The main problem is: Can the higher taxonomic categories cross lower categories? We will start with Khalidi's arguments. Categories that crosscut in sciences are common and they cannot be eliminated as non-real kinds, says Khalidi (Khalidi, 1998: 40–41). Here we will show that the higher taxa cannot crosscut lower taxa.

In the last part, we argue for the thesis that crosscutting categories are not an indicator that there are no natural kinds in biology. We will help ourselves mainly with the works of Muhammad Ali Khalidi.

Key words: crosscutting categories, natural kinds, taxonomy, hierarchical systems, Muhammad Ali Khalidi

1 Khalidi on Crosscutting

One of the problems that Muhammad Ali Khalidi tackles in his work *Natural Kinds and Crosscutting Categories* (1998) is the following: Do the systems' categories crosscut? The sub-problem is this: If they do crosscut, under what conditions is the crosscutting done (Khalidi, 1998: 33)? Philosophers

Hacking and Thomason argue that such systems, i.e. systems that crosscut each other, cannot exist as natural systems or natural/real kinds (Khalidi, 1998: 33–35).

Khalidi says that Thomason's position is based on the idea that natural kinds are arranged hierarchically. This means that higher taxonomic categories cannot cross lower categories (Khalidi, 1998: 33–34; Thomason, 1969). Thomason explains it explicitly: »No natural kinds a and b of a taxonomic system overlap unless a < b or b < a« (Thomason, 1969: 98). In biology, this means that the taxa kingdom cannot crosscut taxa species, that the taxa phylum can't cross the taxa genus, and so on.

Hacking replaces the term natural kinds with scientific kinds. As he says, we are interested in kinds that are associated with instruments, apparatus, and artificial phenomena (Hacking, 1993: 278).

> I'll replace the phrase 'natural kinds' by 'scientific kinds', for we're as much concerned with kinds of instruments, apparatus and artificial phenomena as we are with kinds found in nature. (Hacking, 1993, p. 278)

Khalidi says that Linné's hierarchical classification system in biology is one of the most recognizable systems where the categories are non-overlapping (Khalidi, 1998: 35–36).

Hacking provides a definition of taxonomies, of taxonomy classes, and categories, and says that a specific taxonomy is determined by the class of entities and previous asymmetric relations (Hacking, 1993: 286):

> A taxonomy is determined by a class of entities C and a transitive asymmetric relation K, {C, K} is a taxonomy if and only if (1) it has a head, a member of C that does not stand in the relation K to any member of C but such that every other member of C stands in the relation K to it; (2) every member of C except the head stands in the relation K to some member of C. (Hacking, 1993: 286)

Khalidi notices a difference between his and Hacking's terminology and says that in Hacking's scientific terminology, categories cannot overlap, but in his terminology, the scientific category cannot crosscut (Khalidi, 1998, p. 36). Therefore, we should also pay attention to terminology. We will use the term crosscutting.

> The argument for untranslatability has two parts. First, it is claimed that the kinds investigated in a branch of science can be arranged in taxonomic trees. Scientific kinds in a taxonomy never overlap; either one is properly contained in the other, or they are mutually exclusive. (Hacking, 1993: 278)

As Khalidi explains further, the application of taxonomy is equivalent to the idea that a scientific category in taxonomy does not cross the boundaries between categories of lower taxa (Khalidi, 1998: 36). Let's see what Hacking is saying about categories.

> If K is a 'kind of' relation, the head of each C is a category. When K is given or assumed, for short I will say that the class itself is taxonomic. I may also take each individual taxonomy in C as a category, named by its head-the category of colors or experiments or mammals, for example, in some suitably chosen C. (Hacking, 1993: 286)

As Khalidi says, Hacking admits that there is no evidence that real kinds belong to the taxonomic hierarchy or that taxonomic hierarchies include real kinds (Khalidi, 1998: 40; Hacking, 1993: 300–303).

> Classes can be taxonomic with respect to an asymmetric and transitive relation for all sorts of reasons. A class of mutually exclusive sets is trivially taxonomic with respect to the subset relation; every set is a trivial head. (Hacking, 1993: 287)

Categories that crosscut in sciences are common, and they cannot be eliminated as non-real kinds, says Khalidi (Khalidi, 1998: 40). Let us take Khalidi's example of the parasites category. This category is the one category that crosscuts Linné's tree (Khalidi, 1998, pp. 40-41). It belongs to the real kind because it has the characteristics that are associated with other parasites, and these characteristics were detected before they were placed in any classification system, says Khalidi (Khalidi, 1998: 40–41). Parasitology and zoology can normally coexist, which means that they cannot be incomparable theories, because those kinds of theories must be rivals (Khalidi, 1998: 41). A similar conclusion can be drawn for the following example. We will use Khalidi's example with the following three categories: larvae, pupa, and adults (Khalidi, 2013).

Linné's taxonomy (hierarchical)	Classification scheme by development (non-hierarchical)
Insecta (Class)	larvae
Lepidoptera (Order)	pupa
Danaus plexippus (monarch butterfly) (Species)	adult

Table 1: Comparison: Linné's taxonomy (hierarchical) and the classification scheme by development (non-hierarchical).

For some organisms, on the other hand, the classification on the basis of developmental stages is also useful for explaining the characteristics of these organisms (Khalidi 2013: 70). Let us look at Khalidi's example. Many insects go through the larvae stage and the developmental biology explains aspects of their morphology and behaviour by interpreting the development of the larvae, says Khalidi (Khalidi, 2013: 70).

> Many larvae are specialized for feeding, by contrast with mature adults who are more specialized for reproduction, and generalizations can be made about larvae that are related to this and similar properties. (Khalidi, 2013: 70)

Therefore, argues Khalidi, the classification by developmental stage can crosscut the classification by species or higher taxa (Khalidi, 2013: 70–71). In biology, there are a lot of similar cases; therefore, we may conclude that most categories can crosscut. We agree with Khalidi on that.

> Similarly, if an individual organism classified as a monarch butterfly (Danaus plexippus) can belong to the more inclusive kinds insect and Lepidoptera, why not allow it to belong to the kind larva, all members of which share certain important behavioral properties such as efficiency at locating food sources? In addition to the hierarchy of nested categories to which these individuals belong, they also belong to a set of categories that crosscut those others, overlapping with them only partially. If belonging to a multiplicity of kinds that is hierarchically arranged is allowable, then it is not clear why belonging to a nonhierarchical multiplicity of kinds should be disallowed. (Khalidi, 2013: 72)

Here we have helped ourselves with Khalidi's work. As we have already said, there are many similar examples in biology where the categories crosscut. Depending on which 'classification schemes' we look at, we can differentiate between the crossing categories. Let us look at Khalidi's example, the

monarch butterfly, which belongs to insects, butterflies, and larvae – it is important to distinguish between a certain level and a single category of an individual species. The monarch butterfly (*Danaus plexippus*), according to Linne's taxonomy, falls within the range of species which is in a lower taxa than a category of insects, which is of taxa class, according to this classification. The larvae category is not Linne's taxonomy, but one of the developmental stages. Let us look at what Khalidi says:

> In biology, classification by species is the paramount mode of classifying organisms, but for many organisms classification by developmental stage is also efficacious in explaining features of those organisms. Many insects and amphibians undergo a larval phase, and developmental biology explains aspects of their morphology and behavior with reference to their being larvae. Many larvae are specialized for feeding, by contrast with mature adults who are more specialized for reproduction, and generalizations can be made about larvae that are related to this and similar properties. Hence, classification by developmental phase crosscuts classification by species (and higher taxa). In these cases and many others, two or more classification schemes sort individuals into systems of categories that are not nested wholly within one another but partially overlap. (Khalidi, 1998: 70–71)

Crosscutting categories are also present in other sciences. Khalidi provides an example from chemistry. Solid state, liquid state, and gas are categories that can crosscut with categories of the periodic system (Khalidi, 1998: 41). Chemical categories acids and bases crosscut with the categories of organic and inorganic (Khalidi, 1998: 41-42). Thus, all these categories are real kinds (Khalidi, 1998: 41-42).

Therefore, it is correct to claim that categories that crosscut cannot translate into one another (Khalidi, 1998: 42). Khalidi illustrates this position with the following example of the translatability of genetics to cosmology:

> We should not expect crosscutting categories to be intertranslatable, any more than we should expect genetics to be translatable to cosmology. But, it may be protested, there is a difference in two cases: the relationship of genetics to cosmology in not the same as that of parasitology to zoology. There is still something of a problem in determination whether two theories are in competition with one another and can be translated into one another, or whether they merely coexist comfortably, though they share a subject matter. (Khalidi, 1998: 42)

According to our interpretation of Khalidi's position, this means that we cannot transpose the language and content used in genetics to cosmology. For example, there is no DNA in cosmology, and, *vice versa*, there are no stars in genetics. Therefore, these two sciences do not crosscut. Therefore, as Khalidi warns us, we need to understand that the relationship between genetics and cosmology is not the same as the relationship between parasitology and zoology (Khalidi, 1998: 42).

"The idea that there are crosscutting taxonomies is closely related to the view that scientific categories classifications is interest relative", says Khalidi (Khalidi, 1998: 42). In the case of zoology and parasitology, both are dealing with the same or similar things – with organisms. "A parasite is an organism that lives on or in a host and gets its food from or at the expense of its host" (*CDC*, 2017) and zoology is "the scientific study of the behaviour, structure, physiology, classification, and distribution of animals" (*Oxford Dictionaries*, 2017).

2 Scientific categories and crosscutting

Let us closely examine what Khalidi says about scientific categories and crosscutting. Khalidi wonders whether all natural kinds are scientific categories, and whether some natural kinds could result from human endeavour and not from science (Khalidi 2013: 55). He also suggests that we should avoid talking about the raw similarities between natural kinds (Khalidi, 2013: 57).

Khalidi and Dupré agree that different classification schemes reflect various interests (Khalidi 2013: 63; Dupré, 1999).

> I concur with him in thinking that different classification schemes reflect different interests and that there is no "uniquely best system of classification for all purposes or, which comes to the same thing, independent of any particular purpose" (1999, 473). However, unlike Dupré, I privilege epistemic purposes over other purposes and I therefore accord a special status to those classifications that are introduced primarily to serve those purposes. (Khalidi 2013: 63)

The traditional view on natural kinds is that each individual belongs to one and only one kind, says Khalidi (Khalidi 2013: 69).

> According to the weaker thesis, call it the hierarchy thesis, an individual can belong to more than one kind provided that the kinds it belongs to are arranged in a nested hierarchy. Any individual belonging to kind K1 can also belong to other kinds on condition that those kinds constitute a series of increasingly inclusive kinds. Another way of putting this is by saying that

> natural kinds cannot crosscut or that there can be no partial overlap between natural kinds. (Khalidi 2013: 69)

The next problem is connected to the thesis of mutual exclusion.

> The weakened hierarchy thesis allows an individual to belong to a series of kinds as long as those kinds are all contained within one another. But this weakening of the thesis already drains away much of its apparent appeal. If each individual is allowed to belong to a series of kinds, then we can no longer say that every individual in the universe belongs to some unique kind that supplies its distinctive essence. On the weakened version of the thesis, each individual belongs to a series of more inclusive kinds. But if an individual atom is allowed to belong both to the kind lithium-8 and to the more inclusive kind lithium, then why not allow it also, if it happens to be ionized, to be a member of the kind cation? (Khalidi, 2013: 71–72)

The same applies to examples from biology. Khalidi asks the following: if a single organism, monarch butterfly (*Danas plexippus*), belongs to the kinds *insect* and *butterfly* (*Lepidoptera*), then why is it not allowed to belong to the kind *larva* (Khalidi, 2013: 72)? These individuals, therefore, belong to a set of categories that crosscut each other (Khalidi, 2013: 72).

Thomason says that higher taxa cannot crosscut lower taxa (Thomason, 1969). This is certainly true for hierarchical systems, but the question is whether this is valid in both directions. This means that the lower categories cannot switch borders of a higher category. We will analyse Khalidi's example of the monarch butterfly. If we have a category of the monarch butterfly kind, then we necessarily have all the categories that are higher than this specified species. The higher categories are kingdom, phylum, class, subclass, order. Let us look at the classification of the monarch butterfly (Khalidi's example).

Kingdom: *Animalia* (animals)
Phylum: *Arthropoda* (arthropods)
Class: *Insecta* (insects)
Subclass: *Pterygota* (winged insects)
Order: *Lepidoptera* (butterflies), Linnaeus, 1758
Superfamily: *Papilionoidea*
Family: *Nymphalidae*, Rafinesque, 1815 (Nymphalidae)
Genus: *Danaus*

Species: *Danaus Plexippus* (monarch butterfly), (Linnaeus, 1758)

Therefore, every monarch butterfly must be from the genus *Danaus*, of the family *Nymphalidae*, and so on up the taxa. On the other hand, not every animal has to be of this specific species. Therefore, it is true that the lower categories cannot crosscut higher taxa.

Khalidi states that such systems are not rivals but co-exist (Khalidi 2013: 71). This would be the case when the monarch butterfly is considered as an insect, and at the same time as a larva or an adult (depending on the stage of development).

As we have already said before, there are systems of taxonomies that crosscut quite commonly in different sciences (Khalidi, 2013). Khalidi gives us the following example of crosscutting of scientific categories: classification of stars when observed as a specific category of physical science (Khalidi 2013: 115).

> Ruphy (2010, 1111) observes that astrophysicists involved in stellar classification are interested in several features of stars, such as temperature, density, and mass loss. Since these features are independent, stars can be cross-classified with respect to these features, none of which are primary or basic. Two stars can be classified in the same category with respect to one dimension but not according to another. (Khalidi 2013: 115)

Khalidi says that if equally legitimate categories can crosscut, then there is no hope for a general taxonomic hierarchy that applies to the whole natural world (Khalidi, 1998: 50). The concept of natural kind is not always in accordance with the methods of classification, and thus has to be rejected, says Khalidi (Khalidi, 1998: 50). However, it is possible that the concept of natural kinds will be better articulated in the future, says Khalidi (Khalidi, 1998: 50).

> The world that does not change is a world of individuals. The world in and with which we work is a world of kinds. (Hacking, 1993: 306)

3 Conclusion

In the article, we analysed the problem of crosscutting between classification systems. We helped ourselves mainly with the works of Muhammad Ali Khalidi. The open-ended question here is the following: Are there any classification systems where natural kinds can crosscut? We concluded that

crosscutting categories are not an indicator which would prove the thesis that there are no natural kinds in biology.

In the first part, we tried to analyse the question whether the crosscutting categories can be natural kinds. We tried to show that natural kinds can be arranged hierarchically, and that the higher categories cannot crosscut lower taxa. We analysed Khalidi's view. We tried to explain why crosscutting categories are not an indicator that there are no natural kinds in biology.

The main question of the article was the following: Can the higher taxonomic categories cross the lower categories? We tried to show that crosscutting is possible with lower categories.

We helped ourselves with Khalidi's example of larva, pupa, and adult. As we understand, Khalidi says that larva, pupa, and adult are natural kinds. On the other hand, larva, pupa, and adult are not arranged hierarchically, but are stages of development. Our opinion is that in these examples natural kinds do not necessarily have to be in a hierarchical system. The categories in taxonomy can crosscut. The *Lepidotera* category can crosscut the monarch butterfly category, but not the other way around. This proves that crosscutting is possible with lower categories. We analysed Khalidi's example. Every monarch butterfly must belong to the genus *Danaus,* of the family *Nymphalidae,* and so on up the taxa. On the other hand, not all from the genus *Danaus* have to be the monarch butterfly. Therefore, it is true that the lower categories can crosscut the higher taxa, and that higher taxa cannot crosscut the lower taxa. We believe that crosscutting is not an indicator that there are no natural kinds. Even if the categories crosscut, they could still be natural kinds.

Reference list

Dupré, J. (1999). Are Whales Fish. D. Medin & S. Atran (ed.), *Folkbiology.* MIT Press: pp. 461–476.

Hacking, I. (1993). Working in a New World: The taxonomic solution. Paul Horwich (ed.), *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, Mass.: MIT Press: pp. 275–310.

Khalidi, A., M. (1998). Natural Kinds and Crosscutting Categories. *The Journal of Philosophy*, Vol. 95, No 1: pp. 33–50.

Khalidi, A., M. (2013). *Natural Categories and Human Kinds. Classification in the Natural and Social Sciences.* Cambridge: Cambridge University Press.

Thomason, R., H. (1969). Species, Determinates and Natural Kinds. *Noûs*, Vol 3 (1): pp. 95–101.

Center of Disease Control and Prevention. CDC. (2017). https://www.cdc.gov/parasites/.

Accessed 20. 3. 2017.

Oxford dictionaries (2017). https://en.oxforddictionaries.com/definition/zoology. Accessed 20. 3. 2017.

How to Distinguish between Different Sub-Genres of Utopia with Thought Experiments

Tadej Todorović, University of Maribor

Abstract

In this article, I will endeavour to show how we can use the recent literature on the relationship between thought experiments and fiction to differentiate between sub-genres of utopia. First I establish the relationship between thought experiments and fiction upon which I will build my argument. Next I introduce the classification of utopia into its sub-genres as envisaged by Sargent. In what follows, I aspire to show that we can substantiate the same classification if we use different kinds of thought experiments to categorize utopian works. By introducing Brown's taxonomy of thought experiments, dividing thought experiments into destructive and constructive, I suggest that (i) we can categorize sub-genres of utopia via thought experiments and substantiate Sargent's definitions, and (ii) that the proposed relationship between thought experiments and fiction may serve a practical purpose when it comes to classifying literary genres.

Key words: thought experiments, fiction, literary genres, utopia, dystopia

1 Introduction

When we think of something that is utterly essential to the methodology of analytic philosophy, thought experiments definitely come to mind. Used throughout the history of the discipline, they have become almost indispensable to philosophy, and to imagine philosophy without them would paint a dismal picture, indeed. We can think of thought experiments (TEs) as imagined scenarios that, after contemplation, allow us to learn something new, especially about abstract concepts. The fact that TEs are, by definition, imaginary, immediately brings them close to a literary works of fiction, which share the same characteristic of imaginativeness. This lead to the idea that we could explain the question of "How can a work of fiction which is not and is known not to be true provide any measure of epistemic access to the way things actually are?" (Elgin, 2014: 221), by somehow connecting or even equating fiction with thought experiments. Many philosophers have different ideas about how to connect fiction and thought experiments: Carrol (2002), for example, argues that there are literary thought experiments akin to philosophical thought experiments in fiction, and Elgin (2014) herself argues that works of fiction are *de facto* elaborate thought experiments, thus forming a sort of identity between TEs and

fiction. In this article, I will not argue which theory describes the relationship between TEs and fiction best; however, I will build upon a modest claim that there are TEs in works of fiction and that we can say that some works of fiction are built upon certain TEs. This claim seems fairly uncontroversial, especially in the light of the fact that some fiction is set on TEs – e.g. *The Matrix* is built upon Descartes' *Evil Demon* TE – and that some TEs are derived or at least inspired from works of fiction, as for example Jackson's famous *Knowledge argument*, which seems to draw upon H.G. Wells' *The Country of the Blind*.

To articulate the premise again, I accept that, at the very least, some works of fiction contain TEs and that some are built upon TEs. In what follows, I frame my argument that this line of thinking can have a practical purpose when it comes to the classification of literary genre. I exemplify this claim by using TEs to discern between sub-genres of a specific literary genre – utopia.

2 Utopia and its sub-genres

Laymen and perhaps general aficionados of fiction, and utopian and dystopian fiction in general, would most likely defend the assertion that utopia is "a place of ideal perfection especially in laws, government, and social conditions" (Merriam-Webster, "Utopia") and that dystopia is "an imaginary place where people lead dehumanized and often fearful lives" (Merriam-Webster, "Dystopia"). And while these seem to be the standard senses in which people use these words, the experts' usage differs. As Sargent writes,

> The word *utopia* or *outopia* simply means *no* or *not place*. *Topos* means place; "u" or "ou" means "no" or "not." Thomas More, inventor of the word, punned on *eutopia* or *good place*, and we have since added *dystopia* or bad place. Many, many variations exist on these three words; some, like *uchronia* or *not when*, as serious attempts to add to our understanding and others through sheer authorial self-indulgence. Thus, the primary characteristic of the Utopian place is its non-existence combined with a *topos* –a location in time and space–to give verisimilitude. In addition, the place must be recognizably good or bad or at least would be so recognized by a contemporary reader. (Sargent 5)

In line with this, I will use the word 'utopia' as the umbrella term for the entire utopian genre, i.e. utopia is "a non-existent society described in considerable detail and normally located in time and space" (Sargent, 1994: 9). Furthermore, I will use the word 'eutopia' in the sense that laymen usually

use the word utopia, and I will use the word 'dystopia' in the sense that most people use the word dystopia.

To recapture, utopia is the genre, where a non-existent society is described in considerable detail and is normally located in time and space, eutopia is a sub-genre of utopia, where the said society is somewhat akin to Merriam-Webster and laymen's sense of utopia, i.e. "a place of ideal perfection", and dystopia's sense remains the same.

Before we proceed to the connection with TEs, more detailed definitions of eutopia and dystopia are required.

Eutopia

The first trap that we run into when defining eutopia (i.e. the standard sense of utopia) is its sense of perfection, which does not seem to be grounded in literature. Indeed, as Sargent writes,

> perfect, perfection, and their variants are freely used by scholars in defining Utopias. They should not be. First, there are in fact very few eutopias that present societies that the author believes to be perfect. Perfection is the exception not the norm. (Sargent, 1994: 9)

Thus, we should not try defining eutopia in terms of perfection. However, if perfect society is not the benchmark for eutopian fiction, what is? Sargent suggests that we ought to look at eutopia as a society that seems to be somewhat better than the society in which the author and the intended reader lived in. Eutopia is thus, according to Sargent,

> a non-existent society described in considerable detail and normally located in time and space that the author intended a contemporaneous reader to view as considerably better than the society in which that reader lived. (Sargent, 1994: 9)[1]

Notice that Sargent emphasizes the role of the author's intention when he forms the definition; this seems to be unavoidable. Darko Suvin defines eutopia – albeit, he is defining the word 'utopia' (as mentioned, various authors use various definitions) – in a similar manner, he likewise stresses the author's intention and the community in which the author lived. He writes that eutopia is

[1] I am using Sargent's definitions of utopia, eutopia, and dystopia not only because I find them exceptionally suitable for my purpose, but also because The *Routledge Companion to Science Fiction* (Bould et al. 2009) lists Sargent's definitions for both eutopia and dystopia.

> a verbal construction of a particular quasi-human community where sociopolitical institutions, norms, and individual relationships are organized according to a more perfect principle than in the author's community. (Suvin, 1979: 49)

Omitting the author's intention and the society in which the author lived seems to be almost inevitable when defining either eutopias or, as I will show further on, dystopias, simply because of the notion that past eutopias might seem dystopian to a reader of our time and vice versa; we need not look further than Plato's *Republic* for that, albeit intended as the perfect state, it would presumably horrify any contemporaneous reader as an actual world order. As Sargent writes

> fashions change in Utopias; most sixteenth-century eutopias horrify today's reader even though the authors' intentions are clear. On the other hand, most twentieth-century eutopias would be considered dystopias by a sixteenth-century reader and many of them would in all likelihood be burnt as works of the devil. (Sargent, 1994: 5)

This implies that without the consideration of the author's intention or the society in which the author lived, the eutopia/dystopia disjunction becomes almost relative to who reads it, and this is surely not something we desire. On the other hand, a definition or a criterion that would differentiate between eutopia and dystopia without the inclusion of the author's intention is also something we should strive for, especially if we are advocates of the reader-response criticism, but also if we are not. Surely there must be something more substantial about the sense of eutopia and dystopia besides the intention of the author. Before I focus on this issue, I will briefly define the other sub-genre that I keep mentioning, dystopia.

Dystopia

There are not as many problems when it comes to the definition of dystopia as there are with the definition of eutopia. First of all, there is no ambiguity in word usage, as there are no other words to describe the sense of dystopia. And secondly, one does not have the notion of *the worst* possible world when he thinks of a dystopian world, only of an acutely worse world. In other words, there is no issue similar to the issue of the eutopian sense of perfection, as people do not think of dystopia as the

perfectly horrible world – just plain horrible seems sufficient. Thus, if we see dystopia as the opposite of eutopia, it indeed makes more sense to set the antithesis like Sargent does. He defines dystopia[2] as

> a non-existent society described in considerable detail and normally located in time and space that the author intended a contemporaneous reader to view as considerably worse than the society in which that reader lived. (Sargent, 1994: 9)

Defined in this manner (i.e. dystopia as a society that is considerably worse), which seems the definition that fits the common-sense view of dystopia, snuggly aligns with the notion that eutopia is the opposite of dystopia, i.e. a society that is considerably better. Indeed, here we can see how the common-sense meaning of eutopia (or utopia) fails; if dystopia is the opposite of eutopia, then, if eutopia is the *perfect* society, dystopia has to be *the worst*. As dystopian worlds are clearly not the worst societies, but we still think of eutopias as the opposite of dystopias, it makes sense to define eutopias as simply better and not perfect societies – this preserves the notion that dystopia and eutopia are opposite terms *and* the common-sense meaning of dystopia. If we were to preserve the common-sense notion of eutopia, we would have to sacrifice the common-sense notion of dystopia, as few dystopian worlds could qualify as absolutely *the worst*. Furthermore, as Sargent mentions, "there are in fact very few eutopias that present societies that the author believes to be perfect" (Sargent, 1994: 9), not to mention that even these few eutopias are most likely far from being *de facto* perfect.

With the literary definitions explained, a turn to thought experiments seems appropriate. As I hinted at earlier, when discussing the dependency of the definitions of eutopia and dystopia on the author's intention and the author's society, I wish to substantiate these definitions through thought experiments. In other words, I wish to define eutopia and dystopia with thought experiments. Here, I will take for granted, as elaborated in the introduction, that certain works of fiction are built upon thought experiments. I will argue that eutopias and dystopias are such works of fiction *and* that they correspond to two different kinds of TEs.

[2] Sargent further breaks down the category of utopia into categories like anti-utopia, critical utopia, utopian satire, etc. While I see how further dividing the genre into more and more categories can be beneficial, I do not think that adding these categories would be of any benefit here, as I think that they could be seen as sub-categories of higher categories – if we adopt Sargent's terminology, I would have to say that the higher categories are eutopia and anti-utopia with dystopia a sub-category of anti-utopia *or* eutopia and dystopia with anti-utopia as a sub-category, depending on what one focuses on. However, even if we see dystopia as a sub-category of anti-utopia, the argument remains the same, as everything in the argument that holds for dystopia also holds for anti-utopia.

3 Classification of thought experiments

To show this, I must first classify kinds of thought experiments. I will rely on Brown's taxonomy of TEs, or at his general division of TEs into two kinds. He proposes the following taxonomy of TEs:

> The taxonomy I propose is as follows. First, thought experiments break into two general kinds, which I'll call *destructive* and *constructive*, respectively. The latter kind break into three further kinds which I'll call *direct*, *conjectural*, and *mediative*. There is a small class of thought experiments which are simultaneously in the destructive and the constructive camps; these are the truly remarkable ones which I call *Platonic*. (Brown, 2011: 32)

The division that is pertinent for my cause is the one delineating TEs into *destructive* and *constructive*. Whether further demarcation of constructive TEs into *direct*, *conjectural*, and *mediative*, and the cross-sectional *Platonic* TEs is correct or even necessary is not important for my claim; furthermore, it is beyond the scope of this article. The entirety of Brown's taxonomy has its opponents (*see* Norton, 1993); however, I think that the general division of TEs into destructive and constructive is at the very least useful if not correct. A similar distinction is made by Karl Popper (*see* Popper, 1959) – he divides TEs into heuristic (the aim of which is to illustrate a theory), critical (which are aimed against a theory), and apologetic (which are in favour of a theory). As Brown writes, "Karl Popper refers to the critical and the heuristic uses of thought experiments, which corresponds roughly to my destructive and constructive types" (Brown, 2011: 32). Thus, equating destructive TEs with critical TEs, and constructive TEs with heuristic seems like a safe thing to do.

Constructive TEs

Constructive TEs, as the name suggests, all "aim at establishing a positive result" (Brown, 2011: 35). Again, I do not wish to further differentiate between different kinds of constructive TEs, thus I will first quote Brown in saying that "the contrast between direct and mediative may just be a matter of degree" (Brown, 2011: 41). What is important to know about all kinds of constructive TEs is that they either elucidate a certain theory (Brown calls these mediative) or they construct or develop a theory (direct TEs – according to Brown). Compared to Popper's terminology, they roughly correspond to apologetic and heuristic TEs. Thus, they are *constructing*, developing, or clarifying a certain theory, position, etc.

Destructive TEs

As expected, destructive TEs do the opposite; they are directed against a theory – they challenge or otherwise expose a serious problem of some theory. At this point, a remark on the similarity with Popper's critical TEs seems redundant.

As its name suggests, a destructive thought experiment is an argument directed against a theory. It is a picturesque *reductio ad absurdum*; it destroys or at least presents serious problems for a theory, usually by pointing out a shortcoming in its general framework. Such a problem may be anything from a minor tension with other well-entrenched theories to an outright contradiction within the theory itself. (Brown, 2011: 33)

4 Framing eutopias and dystopias as constructive and destructive TEs

Eutopias and constructive TEs

Comparing the above definitions of eutopia and constructive TEs, we can notice a considerable number of similarities between them. Eutopia is "a non-existent society described in considerable detail and normally located in time and space that the author intended a contemporaneous reader to view as considerably better than the society in which that reader lived" (Sargent, 1994: 9). To create such a society, one must speculate or create the rules (laws, government jurisdiction, ethical values, gender issues, freedom of speech, etc.) of an imaginary society that would be *better* than the current society. These rules, or premises, form the backbone of the eutopia upon which the story usually develops.[3] As the author creates this kind of society, she has to work out solutions to the problems that arise from them, eliminate the inconsistencies, clarify possible ambiguities, and, in general, predict and answer any, or as many as possible, other objections that might be raised against the idea that the society presented is a *better* society than the current society.

This bears uncanny similarities to constructive TEs, where the goal is to form or elucidate a theory; it seems that eutopia does just that – it either creates a theory of a society that is considerably *better* than

[3] In a lot of cases, the eutopian or dystopian society described is not the focus of the work, but is usually described in passing as we follow the plot development. In *Nineteen Eighty-Four,* for example, we follow Winston Smith's journey, while the dystopian society and its rules are described more or less opportunistically. Thus, when I am speaking about the backbone of the eutopia or dystopia, I am talking about the socio-political arrangements, rules, and conventions that form the eutopian or dystopian society – and I think that this is something that can be isolated from the novel, i.e. we can talk about the dystopian society of *Nineteen Eighty-Four* **without** mentioning the plot.

the current society, or it clarifies how a certain kind of society is better than the current society. Of course, the best eutopias ought to do both.

Thus, if we accept the premises that some works of fiction are built upon TEs and that utopian fictions are these kinds of fiction, we can think of eutopias as works of fiction that are built upon constructive TEs. The TEs that they are built upon are TEs of societies that are better than the current society. This might eliminate the need for the author's intention, as we can objectively compare the current society to the constructive TE,[4] or merely analyse the constructive TE to determine whether it, in fact, leads to a better society than the current one.[5]

Dystopias and destructive TEs

A similar point can be made, *mutatis mutandis*, in regard to destructive TEs and dystopias. Dystopia is "a non-existent society described in considerable detail and normally located in time and space that the author intended a contemporaneous reader to view as considerably worse than the society in which that reader lived" (Sargent, 1994: 9); whereas a destructive TEs "is an argument directed against a theory. It is a picturesque *reductio ad absurdum*; it destroys or at least presents serious problems for a theory, usually by pointing out a shortcoming in its general framework" (Brown, 2011: 33). Like eutopias, dystopias also build a society with certain rules and laws, with the expected exception that a dystopian society is considerably worse than the current one. Furthermore, if we accept that utopias are built upon a TE, then dystopias must also be built upon a TE. The question is only how to show that they are based on destructive TEs, as opposed to constructive ones. One way of approaching this is to claim that dystopias evidently show how a certain organisation of a society is *worse* than the current one, therefore they show how a certain premise or theory of a society, even if well-intended at the starts, leads, through a *reductio ad absurdum*, to a society that is considerably worse. However, perhaps better evidence for the claim might be the historical fact that many, in fact most famous, dystopias were written as a response to eutopian ideas. Wheen writes, in the "Introduction" to H. G. Wells' *A Modern Utopia*, that Aldous Huxley's *Brave New World* is "a straight parody of Wells"

[4] It goes without saying that not all eutopias are automatically better than the current societies. The constructive TE within the work of fiction should be subject to the same scrutiny as other constructive TEs – and just because something *is* a constructive TE does not mean that it is a *good* TE. It can still be inherently inconsistent, poorly thought out or incoherent. In short, even though the author's intention might have been to imagine a better society, he might have done exactly the opposite.

[5] Here I am saying that the constructive TE should be compared to the current society – there might be two ways to approach this. Either we are looking for an objectively better society – then we have to compare it to the current one (or at least to the best one that we have) – or we are appraising the TE according to the time in which it was written; then it would be completely viable to compare it to the society in which the work was written. In both cases, we avoid the author's intention.

(Wheen, xxiii). Furthermore, Rahv claims that *Nineteen Eighty-Four* might be "the best antidote to the totalitarian disease that any writer has so far produced" (Rahv 749), and Moylan wrote that "LeRoy sees Orwell's text as a brilliant example of Utopia 'in breakdown'" (Moylan, 2000: 123), and that *Brave New World* "exposes the terrible truth of the totally administered and commodity-driven social system" (Moylan, 2000: 122). Thus it seems, at least with the two most famous, and definitely paradigmatic, examples of dystopia, that they are a *reductio ad absurdum* of a certain eutopian idea or premise. Dystopias do not just describe the horrible conditions or the abhorrent behaviour of society, they show how a certain organisation of a society *leads* to such a horrible, or less dramatically, considerably worse society. In other words, they show how a certain imagined society, with admirable goals in sight (e.g. happiness for all in *Brave New World*) is not sufficient condition for a better society. More than that, it shows that a society that only has one such goal might end up worse overall, which is nothing less than, as Brown puts it, "a picturesque *reduction ad absurdum*" of a eutopian idea.

5 Conclusion

In this paper, I have operated upon the assumption that there are works of fiction built upon TEs and that utopian fictions are that kind of works. I have presented the classification of utopian fiction and highlighted the issue of dependency on the author's intention when classifying something as either a eutopian or dystopian work. Then I presented the taxonomy of TEs, as envisaged by Brown, and drew a parallel between constructive TEs and eutopian works of fiction on one hand, and destructive TEs and dystopian works of fiction on the other. I proposed the idea that we can look at and classify utopian fiction through the lens of TEs; namely, if a work is built upon a destructive TE, we classify it as dystopian, and if a work is built upon a constructive TE, we classify it as eutopian. This might show how TEs can help with the disambiguation of certain literary definitions *and* how they might eliminate the need for the incorporation of the author's intention into the definitions of eutopia and dystopia, respectively.

Reference list

Bould, M., Butler, A. M., Roberts, A., Vint, S. (2009). *The Routledge Companion to Science Fiction*. Routledge.

Brown, J. R. (2011). *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, 2nd edition. Routledge.

Carroll, N. (2002). "The Wheel of Virtue: Art, Literature, and Moral Knowledge." *The Journal of Aesthetics and Art Criticism*, vol. 60, no. 1, pp. 3–26.

"Dystopia." Merriam-Webster, 24 May. 2018,

<https://www.merriam-webster.com/dictionary/dystopia>.

Elgin, C. Z. (2014). "Fiction as Thought Experiment." *Perspectives on Science*, vol. 22, no. 2, pp. 221–241.

Moylan, T. (2000). *Scraps of the Untainted Sky: Science Fiction, Utopia, Dystopia*. Westview Press.

Norton, J. D. (1993). "Seeing the Laws of Nature" *Metascience*, vol.3, pp. 33–38.

Popper, Karl. (1959). "On the Use and Misuse of Imaginary Experiments, Especially in Quantum Theory." In *The Logic of Scientific Discovery*, Hutchinson & Co., pp. 442–456.

Rahv, P. (1949). "The Unfuture of Utopia." *Partisan Review*, vol. 16, no. 7, pp. 743–749.

Sargent, L. T. (1994). "The Three Faces of Utopianism Revisited." *Utopian Studies*, vol. 5, no. 1, pp. 1–37.

Suvin, D. (1979). *Metamorphoses of Science Fiction*. New Haven.

"Utopia." Merriam-Webster, 24 May. 2018,

<https://www.merriam-webster.com/dictionary/utopia>.

Wheen, F. (2005). "Introduction." In Wells, H. G. *A Modern Utopia*. Penguin Books.