STUDIA PHILOSOPHICA WRATISLAVIENSIA

Supplementary Volume, English Edition 2013

Edited by
Adam Chmielewski, Roman Konik
Damian Leszczyński, Artur Pacewicz

WROCŁAW 2013

Studia Philosophica Wratislaviensia

Supplementary Volume, English Edition 2013

Edited by
Adam Chmielewski, Roman Konik,
Damian Leszczyński, Artur Pacewicz

Publishers
Instytut Filozofii Uniwersytetu Wrocławskiego
Polskie Forum Filozoficzne

Niniejsza praca naukowa została sfinansowana w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „Narodowy Program Rozwoju Humanistyki" za pośrednictwem grantu nr 31H 11 0059 80 w latach 2012–2013.

The present issue of the quarterly journal Studia Philosophica Wratislaviensia has been financed from the grant no 31H 11 0059 80 of the Ministry of Science and Higher Education of the Republic of Poland, within the framework of the National Programme of the Development of Humanities in the years 2012–2013.

Editorial Office
Instytut Filozofii Uniwersytetu Wrocławskiego
ul. Koszarowa 3
51–149 Wrocław
e-mail: studiaphilosophica@poczta.fm
www.studiaphilosophica.pl

ISSN 1895–8001
ISBN 978–83–933495–04–1 (e-book)
ISBN 978–83–933495–05–8

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

Contents

Aesthetics

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

Editor's Foreword

The present issue of the Supplementary Volume is a continuation of the project of the staff of the Polish language quarterly *Studia Philosophica Wratislaviensia*. The English language edition of our journal has been made possible by the grant No 31H 11 0059 80 awarded to our team by the Poland's Ministry of Science and Higher Education within the framework of the National Programme of the Development of Humanities (M. P. No 86, 1014).

As the first volume, which was published precisely a year ago, it is aimed at promoting Polish philosophical thought abroad. Also, as in the first issue, most of the papers included in this volume were authored by the Polish philosophers who have contributed to our journal ever since its establishment in 2006. It also features contributions by philosophers from other countries. I would like to express my thanks to the authors of this volume and to the people involved in this project: Professors Damian Leszczyński and Roman Konik, Dr. Artur Pacewicz, and Urszula Lisowska M.A. I also thank Mr. Jacek Słupski who translated several papers, and Mr. Grant Hennessy, who edited some of the papers for the present issue. I would also like to thank Ms. Julita Mrzygłód for her assistance in managing the financial aspect of our project.

Adam Chmielewski
Professor, Editor-in-Chief

History of Philosophy

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

ARTUR PACEWICZ
University of Wrocław

Wisdom – Knowledge – Belief. The Problem of Demarcation in Plato's *Phaedo**

Abstract

The aim of the present paper is to show how Plato suggested demarcating between knowledge and other kinds of human intellectual activities. The article proposes to distinguish between two ways of such a demarcation. The first, called 'the external demarcation', takes place when one differentiates between knowledge and non-knowledge, the rational and non-rational or the reasonable and non-reasonable. The second, called 'internal', marks the difference within knowledge itself and could be illustrated by the difference between the so called hard and soft sciences. The analyses lead to the following conclusions. Plato refers to the whole of human intellectual activity as *doxa*, which is divided into two spheres. The first of them is knowledge proper whose criterion is *phronesis*. Three other kinds of *doxa* are derived from knowledge proper: 1) the traditional *peri phuseōs* investigation (called also *sophia*); 2) popular *doxai* concerning virtues; 3) wisdom of the antilogikoi. The difference represents the external demarcation. There may be, however, a difference in the scope of knowledge proper (the internal demarcation). If the *peri phuseōs* investigators were able to explain the field of values, the result of their investigation could be acknowledged as knowledge, although it would still be characterized as inferior due to its being based on senses. What is interesting about knowledge proper is that it is not firm and reliable but only hypothetical. It does not determine the skeptical reading of the *Phaedo* but it indicates that Plato has just begun his own philosophical project (which is still in progress) and the knowledge presented in the dialogue is his first positive suggestion how to solve the problem of demarcation.

When observing scientific life from a sociological perspective, one can clearly notice the important role played in modern society by various 'specialists', particularly if their area of expertise is exact sciences. We believe physicists when they

* This paper is a slightly revised and extended version of the text written in Polish and published in the volume: U. Wollner, M. Taliga (eds.), *Poznanie a demarkácia*, Tribun 2011, pp. 42–61, pp. 87–99.

claim that the universe came into being billions of years ago. It seems rational to assume that the Earth came into existence and initially was devoid of any life, which then appeared and evolved in such a way that it led to the emergence of *homo sapiens*. But it appears also irrational that man was created by God, that UFOs exist or that the fate of humans is written in the stars. The above examples indicate the first aspect of demarcation that I would like to discuss. I will refer to it as 'external demarcation'. This seems to be the most frequently represented aspect in modern epistemological reflection, as part of which a *criterion* is sought for differentiating science from that which is not science, and which can be contained in weak convictions or in irrational faith (both these areas may be contained in the concept of metaphysics, broadly understood[1]). Scientists tell us that there is no scientific proof for the existence of God or UFOs, or a scientific confirmation of the verifiability of claims made by astrologers. When seeking a criterion that, as mentioned above, would have a nature of external demarcation, one desires to designate the area of scientific cognition and to separate it from other, 'non-scientific' domains. This obviously translates into the need to formulate an appropriate concept (conditions) of scientific knowledge, which concept becomes in fact the very scientific criterion.

The other demarcation aspect I will call 'internal' and this might also be regarded as a weaker form of the previous aspect. It presupposes that there exists a group of convictions that, with a certain degree of probability, most people regard as science, and the problem of demarcation arises precisely in this area. An example of this type of differentiation is the criticism of psychology by psychiatry or the criticism of representatives of arts by representatives of exact sciences. In the first case (external demarcation), it is about the differentiation between science and non-science, between that which is rational and irrational (reasonable and unreasonable[2]), in the other case, it is rather about the dispute about the form of that which is scientific (science – pseudoscience or better – worse science), in which for instance attention is paid to accuracy or the method of investigation (deduction-induction) as criteria for hierarchization (hard science – soft science).

One may ask the question, however, whether such a differentiation between external and internal demarcation appeared in antiquity and, consequently, whether it can be applied to it. It is very difficult to settle this issue with reference to the currents in ancient philosophy that appeared before the sophists. With the pre-Platonists, Heraclitus' critical remarks about Homer and Hesiod[3] indicate that

[1] D. Gilles, *Philosophy of Science in 20th Century. Four Central Themes*, Oxford–Cambridge 1993, p. 155.

[2] L. Wittgenstein, *Tractatus Logico-Philosophicus*, 4.003, transl. D.F. Pears, B.F. McGuiness, London–New York 2001 [Polish translation: L. Wittgenstein, *Tractatus Logico-Philosophicus*, transl. B. Wolniewicz, Warszawa 1997].

[3] DK 22 B 56 (H. Diels, W. Kranz, *Die Fragmente der Vorsokratiker*, Berlin 1966) = fr. 21 Marcovich (M. Marcovich, *Heraclitus. Editio maior*, Merida 1967) = fr. XXII Kahn (Ch.H. Kahn, *The Art and Thought of Heraclitus. An Edition of the Fragments with Translation and Commentary*, Cambridge–London–Melbourne 1979): "Men are deceived in the recognition of what is obvious, like Homer who was wisest of all Greeks. For he was deceived by boys killing lice, who said: what we see and catch we leave behind; what we neither see nor catch we carry away" (all quotes from Heraclitus are from Kahn's translation) [Polish translation: Heraklit,

the two poets, while possessing a certain knowledge about the sensorily cognizable reality, did not complement it with anything that could be called 'knowledge of the essence'. They failed to recognize that apart from that which can be accessed with the senses, there exists a reality in the form of *Logos*, which causes, among other things, that opposites can and do constitute a unity. In the case of Heraclitus', the statements about Pythagoras,[4] emphasis is put on the latter's enormous dedication to science (*historia*), which only resulted in the creation – on the basis of other people's reflections – of knowledge (*sophia*), which combines many domains (*polumatheia*), and so knowledge referring to many truths. Such a multitude of truths ignores the uniqueness of *logos* and cannot be applied in order to become a sage-philosopher, and so as such constitutes bad art (*kakotechnia*). In all the three cases we can read the fragments in a context that is 'favourable' to Homer, Hesiod and Pythagoras, i.e. conclude that according to Heraclitus, they possess some knowledge, which is characterized above all by a certain multitude, because it refers to a multitude, but they failed to recognize the *essence* of reality – its ontic unity. Then we would have to do with an interpretation in which the Ephesian's fragments would talk about internal demarcation – within a broadly-understood knowledge. One can assume, however, that Heraclitus was much more critical in relation to the intellectual context that he found (and this is basically how his position is described in studies) and refused to designate the above-mentioned three authors as 'wise men'. Consequently, the fragments referred to above would present a thought about external demarcation. We have a similar situation in the case of the philosophy developed by Parmenides, who – as is universally known – differentiates between the way of truth and the way of opinion. The differentiation expressed in such a general way fulfils the demarcation function both in the external and internal aspects, and this is so because in the *doksa* area there appears a difference between the conviction of the poem's author and the convictions of other mortals. Thus, the juxtaposition knowledge-opinion, in which the author of the latter is Parmenides (or more generally – the poem's lyrical subject), indicates internal demarcation, while the same juxtaposition in which *doksai* are an expression of convictions of other thinkers has a nature of external demarcation, refusing to attribute any cognitive value to other convictions[5].

Such attempts to demarcate, i.e., separate philosophy/science from other domains of intellectual activity, were challenged by the sophists. Above all, they pointed out at the common basis of human knowledge, which – according to them

Fragmenty: nowy przekład i komentarz, transl. K. Mrówka, Warszawa 2004]. DK 22 B 57 (= fr. 43 Marcovich = fr. XIX Kahn): "The teacher of most is Hesiod. It is him they know as knowing most, who did not recognize day and night: they are one."

[4] DK 22 B 129 (= fr. 17 Marcovich = fr. XXV Kahn): "Pythagoras son of Mnesarchus pursued inquiry further than all other men and, choosing what he liked from these compositions, made a wisdom of his own: much learning, artful knavery".

[5] DK 28 B 8, 60–61: "I tell you all the likely arrangement in order that the wisdom of mortals may never oustrip you", translated by L. Tarán [*Parmenides. A Text with Translation, and Critical Essays*, Princeton 1965) [Polish translation: Parmenides, 'Fragmenty poematu o naturze', transl. M. Wesoły,*Przegląd Filozoficzny – Nowa Seria* 10 (2001), pp. 71–85].

– was sense perception.[6] This constitutes the criterion for the truthfulness of judgments and, being accessible to all people, does not allow for differentiating knowledge. The difference occurs within subjective human convictions originating from sense perception, but it is pre-arranged and relative in nature,[7] as basically all human judgments can be designated as equipollent.[8] A differentiation that is in fact ostensible can only be made by appropriate argumentation in favour of a given opinion, which is not its justification in the sense of provision of an objective or absolute principle-reason, but rather a specification of a larger number or stronger reasons whose strength of influence would tip the balance towards one of them through subjective persuasion.

One may ask why our deliberations are to focus on the *Phaedo*, which in antiquity was classified as an ethical dialogue, which meant that the deliberations presented there were, according to the ancients, to have a practical application.[9] However, quite often in antiquity, it was looked upon as a dialogue about the soul,[10] whose essential activity has, above all, an intellectual dimension, although surely the ethical aspect cannot be regarded as less important.[11] Thus, one can assert that deliberations on demarcation, although they do not constitute the principal focus of the work, constitute a significant part of it. There is also another argument. I am not an adherent to the evolutionary or developmental interpretation of the thought of the founder of the Academy's,[12] but if one were to agree with the findings of the studies into relative chronology, the *Phaedo* is to have been created after Plato founded his own school of philosophy.[13] Thus, it can be assumed that in such a work one can find important deliberations on the issue of demarcation, because Plato might have desired to define his own field of examination and education, which, on the one hand, would attest to the Academy's originality and would differentiate it from other Athenian schools, and, on the other hand, would encourage potential students to enrol.[14]

[6] DK 80 B 7; DK 82 B 11a.

[7] A. Pacewicz, 'Relatywizacja dobra w filozofii sofistáw?', *Przegląd Filozoficzny – Nowa Seria* 14 (2005), pp. 7–22; S. Consigny, *Gorgias. Sophist and Artist*, Columbia 2001, p. 40. Sextus Empiricus (*Adversus Mathematicos* VII 48, 60, 65) counts Protagoras and Gorgias among philosophers who rejected the criterion for truth. This, however, should be understood as a rejection of the criterion for the absolute nature.

[8] DK 80 A 6a, A 20; V. Brochard, *Les sceptiques grecs*, Paris 1932, p. 16.

[9] Diogenes Laertius, *Vitae philosophorum*, III, 49–50.

[10] *Ibidem*, III 58; Cicero, *Tusculanae disputationes*, I 11, 24; on the immortality of the soul see Gellius, *Noctes Atticae*, II 18.

[11] P. Stern (*Socratic Rationalism and Political Philosophy. An Interpretation of Plato's "Phaedo"*, New York 1993, pp. 6–7) draws attention to the two-dimensionality of the dialogue in another aspect. He discerns in it a tension between the reference to another world that appears in the arguments for the immortality of the human soul and in the description of the land in which the soul resides after the body dies, and the situation in which Socrates finds himself – being in this world. The tension is equivalent to the disproportion of the two aspects of teaching present in the dialogue.

[12] A. Pacewicz, *O ewolucyjnym charakterze filozofii Platona*, [in:] A. Pacewicz, A. Olejarczyk, J. Jaskóła (red.), *Philosophiae Itinera. Studia i rozprawy ofiarowane Janinie Gajdzie-Krynickiej*, Wrocław 2009, pp. 373–390.

[13] W. Stróżewski, *Wykłady o Platonie. Ontologia*, Kraków 1992, p. 26

[14] W. Lutosławski (*The Origin and Growth of Plato's Logic with an Account of Plato's Style*

The analyses to be done in this paper focuses on the following four terms: *doksa, sophia, epistēmē* and *phronēsis*.

Doksa and the derivatives of *dokeō*[15] are probably, in terms of range, the broadest epistemic term used by Plato in his dialogues. In the *Phaedo*, it principally seems to be a neutral notion, expressing a certain possibility, but receives a negative or positive connotation in a given context. Such a neutral expression becomes apparent for instance just before the first argument for the immortality of the soul. He formulates three conditions that are supposed to do away with the unbelief (*apistia*) that people have towards this issue. It should be demonstrated that: (1) the soul is after death, (2) it has some strength (*dunamis*), (3) it has *phronēsis*.[16]

Kebes wants to learn Socrates' opinion (*doksa*)[17] about it, but the manner of the discourse is determined by Socrates/Plato as *diamuthologein*,[18] with a limitation of the claim of truthfulness only to probability (*eite eikos houtōs echei eite mē*[19]).

In its negative connotation, *doksa* describes such an epistemic state that does not correspond to any state of things and, consequently, may or should be changed. This happens in the appearing threat about the suicide ban, as to which Socrates observes the possibility of holding the opinion (*dokseien*) that it is nonsensical (*alogos*).[20]. Meanwhile, *some* justification for it can be found, although in this particular case it is some enigmatic *aporretoi logoi*,[21] according to which it is

and of the Chronology of His Writings, London–New York–Bombay 1897) discusses the *Phaedo* in the chapter *Origin of the Theory of Ideas*. One can also mention the opinion formulated by K. Dorter (*Plato's "Phaedo": An Interpretation*, Toronto–Buffalo–London 1982, p. 134) and J. Dalfen (*Kenneth Dorter's Interpretation of the "Phaedo"*, [in:] C.L. Griswold (ed.), *Platonic Writings/Platonic Readings*, University Park 2002, p. 215) that one of the permanent issues present in Plato's dialogues is the problem of the limits of knowledge. R.K. Sprague interprets the *Phaedo* as a protreptic (*Plato's "Phaedo" as Protreptic*, [in:] S. Stern-Gillet, K. Corrigan (eds.), *Reading Ancient Texts. Vol. I: Presocratics and Plato. Essays in Honour of Denis O'Brien*, Leiden–Boston 2008, pp. 125–133).

[15] It should be pointed out that in ancient Greek there are at least two philosophically important groups of notions with a similar, if not the same, meaning. The first one comprises the verb *dokeō*, the adjective *doksastikos* and the nouns *doksa, doksis*, *dokēma, doksasia*, *doksasma*, while the other – the verb *oiomai*, the adjective *oiētikos* and the nouns *oiēsis*, *oiēma*. To date it has not been examined whether the groups differ in meaning or whether they can be used as synonyms. Studies into the term *doksa* concentrate primarily on the late dialogues; *cf.* for instance J. Sprüte, *Der Begriff der Doxa in der platonischen Philosophie*, Göttingen 1962; E. Tielsch, *Die Platonischen Versionen der griechischen Doxalehre. Ein hilosophisches Lexicon mit Kommentar*, Meisenheim am Glan 1970.

[16] Plato, *Phaedo*, 70b 2–4.

[17] *Ibidem*, 70b 9.

[18] Apart from the *Phaedo*, the verb is to be found in *Apology of Socrates* (39e 5), where Socrates uses it to explain and evaluate what happened on the day of the defence. As is known, the explanation also contains two hypotheses regarding what happens to humans after death. On the other hand, in the *Laws* (632e 3–5), it indicates a certain alleviation of the discipline (strictness) of the dispute, in this case – on virtues.

[19] Plato, *Phaedo*, 70b 6–7.

[20] *Ibidem*, 61b 1–2.

[21] *Ibidem*, 62b 3. K. Dorter (*Plato's "Phaedo"...*, p. 19) identifies their sources as Orphic, while M. Miles ('Plato on Suicide (*Phaedo* 60c–63c)', *Phoenix* 55 (2001), p. 244) talks about the

against the law (divine law – *ou* [...] *themiton*),[22] i.e. it is not pious (*ou hosion*).[23] This is also how the reference by Socrates/Plato and Simmias to the figure of Evenus can be interpreted.[24] Simmias asked by Xanthippe's husband whether Evenus is a philosopher, answers that he is convinved about it (*dokei*).[25] Within the context of the discussion on the sense of the philosopher's dying and the concept of philosophy presented in the *Phaedo*, this conviction of Simmias should change, as Simmias himself admits that Evenus is not inclined (*hekōn*) to follow Socrates' solutions and advice. *Doksa* is treated similarly in the case of the body-soul relation, which relation may have an appropriate or an inappropriate nature. The other form consists in making the soul similar to or even in equating it with the body. This is because the body has at its disposal stimuli in the form of pleasure and pain (*hēdonē kai lupē*), which are so strong that they may make the soul to be convinced (*doksadzousa*) that what the body is saying is true. Equation at the level of convictions (*homodoksein*)[26] has consequences in making similar in terms of action (*homotropos*) and in terms of eating (*homotrophos*), i.e. – generally speaking – ceases to be a pure being.[27] Conviction proper, however, treats this issue entirely the other way round – a (full) truthfulness is connected with that which is divine and what is *adoksaston*.[28] Possession of *doksa* in the negative sense is also reserved for those who are deprived of upbringing (*apaideutos*).[29] Such a person does not deliberate, but disputes (*amphisbetein*), does not use *phronēsis*, but his only goal is to win a dispute (*philonikōs*) and to convince the public.[30]

A positive connotation of the term *doksa* appears when there is a reference to the genuine philosophizing (*gnēsiōs philosophoi*) (or those who are *orthōs philo-*

Pythagorean tradition. Taking into account the emphasis on the role of piety and divine law in Plato's *Apology of Socrates* (30c 9–d 1), the Socratean tradition may be added or one may say that this was Plato's own idea.

[22] Plato, *Phaedo*, 61c 10.

[23] *Ibidem*, 62a 6.

[24] W. Nestle (*Vom Mythos zum Logos, die Selbstentfaltung des griechischen Denkens von Homer bis auf die Sophistik und Sokrates*, Stuttgart 1975, p. 420) thinks Evenus is a sophist, and so does C. Rowe ('Contre Platon: Philosophie et littérature dans le *Phédon*', [in:] M. Dixsaut (ed.), *Contre Platon.* vol. II: *Le Platonisme renversé*, Paris 1995, p. 278) interprets the mention of Evenus as a philosopher as irony, which is regarded as groundless by T. Ebert (Platon, *Phaidon*, T. Ebert (Übers. & Komm.), Göttingen 2004, p. 113).

[25] Plato, *Phaedo*, 61c 6–7.

[26] The significance and strength of such an equation is attested by several other fragments of *Corpus Platonicum*. In the*Republic* (433c 6), *homodoksia* of those in power is one of the elements considered when evaluating *polis* and a characteristic of a prudent human soul, i.e. a soul in which two inferior soul powers are subordinated to the supreme power (*ibidem*, 442c 10–d 3). In the *Statesman* (310e), one of the tasks of the royal art is not to allow the separation of prudent characters (*sōphrona*) from brave characters (*andreioi*), but to bind them with, among other things, similar opinions (*homodoksiai*,*doksai*).

[27] Plato, *Phaedo*, 84d 4–e 3.

[28] *Ibidem,* 84a 8.

[29] The lack of upbringing is also connected with the above-mentioned impurity, which can be seen in Plato's *Sophist* (320d 6–e 3), which results from a failure to succumb to the elenctic procedure. In the *Timaios* (86e 1–2), the lack of upbringing food (*apaideutos trophē*) is directly identified as the reason for becoming a bad person.

[30] Plato, *Phaedo*, 91a.

mathes[31]). In a rhetorically very well thought out disquisition,[32] there is presented a conviction which such people necessarily should present, which talks about equating the body with evil, and the soul with that which is good, about attaining the truth only after separating with the body.[33] The really philosophizing (the really loving *phronēsis*) are convinced that *phronēsis* will only be achieved in Hades.[34]

At last one can mention two more characteristics which may be attributed to convictions and being convinced. The first characteristic is variability, which can be noticed in the disquisition devoted to misology, and which is not really about conviction, but about the subject condition of being convinced. According to Socrates, sometimes man happens to believe (*pisteusēi*)[35] some true *logos*, but because of the lack of the art of argumentation (*technē peri tous logous*), over time one can become convinced (*doksēi*) that it is false, regardless of whether it is indeed so.[36] Such a perspective is certainly connected with the above introduced division into two aspects of conviction that can now be at least partly equated with the truthfulness and untruthfulness of *doksa*. The other characteristic is a certain gradeability of conviction. Socrates/Plato does not exclusively care about being convinced about something – such a goal motivates those only interested in winning disputes. The objective is to have a conviction in the highest degree (*malista dokein*) that one has something somewhat. The way to increase the degree of conviction is reasoning (*logidzesthai*) of a hypothetical nature: if X happens to be true then it is a beautiful/good conviction, and if not – then one stays ignorant (*anoia*), which in turn is evil.[37]

Originally, the term *sophia* was used to refer to poets, clairvoyants – generally to those who disclose knowledge that is inaccessible to mortals and cannot be disclosed in any other way. Its subject matter is not technology as such, but gods, humans, society. Wise men (*sophistes*) include Homer and Hesiod, musicians, pre-sophist philosophers – some of whom, e.g. Xenophanes and Heraclitus, equate *sophia* with *aretē*[38] – and heroes of stories – Prometheus or Odysseus.[39] In Plato's

[31] In the *Republic* (376b 8–c 2; 581b 9), a science lover is equated with a wisdom lover, who desires the truth from the earliest years (*ibidem*, 485d 3–4), by nature insistently aspires to being (*to on*), to encompass its nature (*phusis*) by becoming such as what is really real (*to on ontōs*), and does not stop at individual beings, which are the subject of beliefs (*ibidem*, 490a 8–b 7).

[32] T. Ebert (*Phaidon*, p. 140) talks about a peculiar *confessio Pythagorica*, and as a parallel, quotes a fragment from Archytas, which is cited in Cicero's *Cato the Elder On Old Age.*

[33] Plato, *Phaedo*, 66b–67b.

[34] *Ibidem,* 68a–b.

[35] This perspective shows a difference between the *Phaedo* and the *Republic*. In the latter dialogue, in the famous metaphor of the divided line (*Republic* 509d–511e), although belief constitutes a type of conviction, it does not relate to the sphere of *logos*, but the objects of the sensory world.

[36] Plato, *Phaedo,* 90b 4–d 7.

[37] *Ibidem*, 90d 9–91b 3.

[38] DK 21 B 2; DK 22 B 112 (= fr. 23f Marcovich = fr. XXXII Kahn). S.D. Sullivan, *Psychological and Ethical Ideas: What Early Greeks Says*, Leiden–New York–Köln 1995, pp. 170–171; *cf*. C.J. Vamvacas, *The Founders of Western Thought – The Presocratics. A Diachronic Parallelism between Presocratic Thought and Philosophy and the Natural Science*, Dordrecht 2009, p. 115; P. Hadot, *Czym jest filozofia starożytna?*, tłum. P. Domański, Warszawa 1992, pp. 42–47 [English translation: *What is Ancient Philosophy*, transl. P. Chase, Harvard 2002].

[39] G.B. Kerferd, *The Sophistic Movement*, p. 24.

dialogue *Phaedo*, it only appears on its own twice, but nearly 40 times in the compound *philosophia*. However, these two meanings appear to have a considerable significance for the issue in question.

It occurs for the first time in the fragment presenting the intellectual path of Socrates/Plato, who admits that when he was young, he desired wisdom referred to as *historia peri phuseos*.[40] The reason for his desire was the knowledge (*eidenai*) about the causes (*aitiai*) that were supposed to be answers to the questions why (*dia ti*) every individual thing came into existence, perished and was. The scope of that knowledge, as is known, was not limited to learning the rules, but other issues were also considered within it. Socrates provides their examples, e.g. the coming into being of living creatures or epistemological issues (perception, thinking). In the explanation of *peri phuseōs* there appears a position that may be described as physicalistic, and which tends to present relations between individual beings in the quantitative aspect. However, from the perspective of the issue of demarcation, the most important of Socrates'/Plato's confessions is the one that after he started such deliberations he became convinced that by nature he was incapable (*aphues*) of conducting them, because they led him to scepticism (he ceased to know what he had thought he knew [*ōimēn eidenai*]). As is known, even the recognition by Anaxagoras of the proper cause of everything, i.e. the mind, did not help, because apart from indicating the cause, the philosopher of Clazomenae did not use it in explaining all spheres and aspects of reality. Socrates/Plato regarded the absence of valuation deliberations and criteria on the basis of which those might appear as a major deficiency. In my opinion, this is an attempt at resolving the issue of demarcation; the question is whether in the external or internal aspect. I believe that *historia peri phuseōs*, in which all data are based on the senses and in which physicalistic explanation predominates, is excluded from the scope of knowledge. If such philosophizing encompassed the axiological aspects and assumptions resulting not only from sense perception, this would be knowledge *sensu stricto* – a true philosophy of nature. According to Socrates/Plato, this does not apply, however, to the concepts that arose before his deliberations.[41]

For the second time the term *sophia* appears in the deliberations on the concept of hypotheses.[42] As is known, Sokrates/Plato, discouraged by deliberations like *peri phuseōs historia*, began the so-called 'second flowing' (*deuteros plous*),[43] which is to consist in presenting the truth of beings (*alētheia tōn ontōn*), with

[40] Plato, *Phaedo*, 96a–99d.

[41] R. Bolton (*Plato's Discovery of Metaphysics. The New Methodos of the "Phaedo"*, [in:] J. Gentzler (ed.), *Method in Ancient Philosophy*, Oxford 1998, pp. 91–111) argues that Socrates rejects outright the possibility of physics existing as a science, because the only possible science is metaphysics; *cf*. H. Wagner, *Platos "Phaedo" und der Beginn der Metaphysik als Wissenschaft ("Phaedo" 99d–107b)*, [in:] F. Kaulbach, J. Ritter (hrsg.), *Kritik und Metaphysik. Studien. Heinz Heimsoeth zum achtzigsten Geburstag*, Berlin 1966, pp. 363–382.

[42] Plato, *Phaedo*, 99d.

[43] For more on this *cf*. e.g. D.L. Ross, 'The *deuteros plous*, Simmias' Speech, and Socrates' Answer to Cebes', *Hermes* 110 (1982), pp. 19–25; S.M. Tempesta, *Sul significato di deutros plous nel "Fedone" di Platone*, [in:] M. Bonazzi, F. Trabattoni (cur.), *Platone e la tradizione platonica. Studi di filosofia antica*, Milano 2003, pp. 89–125.

the use of arguments (*logoi*).[44] This is how he found his own method, which – generally speaking – consists in:

1) assuming *logos* regarding a problem which is the strongest and recognizing something as to which one is convinced that it agrees with such logos (*sumphonein*) as true, and that which does not agree – as untrue;[45]

2) checking whether the theorems resulting from such a hypothesis mutually agree or disagree;[46]

3) in order to substantiate such a hypothesis later, another one, which is higher (*anōthen*) and the best (*beltistē*) is proposed, and so on, until there is obtained a hypothesis that is the most general and sufficient (*hikanos*)[47] for substantiating all the other ones.[48]

At the same time, Socrates/Plato warns against dialectic consideration of the very principle-hypothesis together with its consequences. This procedure is not followed by antilogicians, and this is because of their own wisdom; however, it should be complied with by philosophers.[49] Such a behaviour of antilogicians is justified by their views on things and *logos*. Earlier,[50] Socrates demonstrated that according to antilogicians there is nothing logical or certain in either sphere (*ouden hugies oude bebaion*), because everything is subject to constant change. Thus, for the second time, we have the issue of demarcation expressed – this time, between antilogic and philosophy; the question arises, however, whether in the external or internal aspect. Of key importance to the resolution of this issue is – in my opinion – the designation by the *Phaedo's* author of the ontological-and-gnoseological concept which lies at the base of antilogicians' intellectual position, with the term *atechnōs*, i.e. 'against art'. I believe that the designation indicates external demarcation, ruling out antilogical wisdom from the sphere of knowledge. This happens not only because of the ontological position referred to above, but probably also because they use the juxtaposition truth-falsehood, however treating

[44] Socrates adds the reservation that a study of beings in *logoi* does not consist more in studying them in images than in action (*erga*). In this case, most probably we do not have to do with a traditional juxtaposition of theory and practice, but action refers to the operation of the senses; *cf*. P. Thanassas, 'Logos and Forms in *Phaedo* 96a–102a', *Bochumer Philosophisches Jahrbuch für Antike und Mittelalter* 8 (2003), p. 9.

[45] Plato, *Phaedo*, 100a 3–7.

[46] *Ibidem*, 101d 3–5.

[47] Obviously, the question might arise whether objectively or subjectively, especially in the context of the later mention of the method of hypotheses (*Phaedo*, 107b): "You are not only right to say this, Simmias, Socrates said, but our first hypotheses require clearer examination, even though we find them convincing. An if you analyze them adequately, you will, I think, follow the argument as far as a man can and if conclusion is clear, you will look no further" (translated by G.M.A. Grube, [in:] Plato, *Complete Works*, J.M. Cooper (ed.), Indianapolis–Cambridge 1997). It seems that the (inter)subjective aspect of sufficiency is emphasized here. For more on the method of hypotheses in the *Phaedo cf*. e.g. J.T. Bedu-Addo, 'The Role of the Hypothetical Method in the *Phaedo'*, *Phronesis* 24 (1979), pp. 111–132; Y. Kanayama, 'The Methodology of the Second Voyage and the Proof of the Soul's Indestructibility in Plato's *Phaedo'*, *Oxford Studies in Ancient Philosophy* 18 (2000), pp. 41–100.

[48] Plato, *Phaedo*, 101d 6–e 1.

[49] *Ibidem*, 101d–e.

[50] *Ibidem*, 90b–c.

these notions interchangeably – the same can be true in one case and false in another case.[51] This does not mean, however, that antilogic must be entirely devoid of values, because although Plato "has a low opinion of antilogic as a style of philosophical debate, he does not suppose that its practice establishes that its practitioner is therefore a sophist. It is not in itself dishonest or directed to deceive".[52]

As indicated by the above deliberations, both considerations of *peri phuseos* type and antilogical deliberations are excluded from the scope of wisdom or knowledge. It remains to be examined what falls within the scope of the term *epistēmē*.

For the first time, the term appears in the *Phaedo* when the argument from anamnesis is presented.[53] Knowledge together with *orthos logos* are located (*enousa*) in man, thanks to which man can answer properly asked questions (e.g. about geometrical figures[54]). Thus, knowledge is created by reminding and as such can be equated with reminding. Generally speaking, Socrates'/Plato's deliberations enable us to determine that such knowledge is internally differentiated and the criterion for the differentiation may constitute the object to which the knowledge relates ('a different knowledge about man and about a lyre'), and/or time (a reminder means perception of something earlier). The scopes of different domains of knowledge can be independent (man-lyre) or co-dependent on one another (Simmias-drawn Simmias). Sense perception is not a source of knowledge, because together with it only recognition (*gnōsis*) is created, and Socrates/Plato regards the general object (e.g. man, equality) as the proper object of knowledge; such general object can be compared with a sensory object and also reveal some deficiency in the latter.

However, it seems that apart from this type of knowledge, some other, inferior type of knowledge is allowed to exist, if one assumes that apart from a reminder *sensu stricto*, there is also anamnesis *sensu largo*. This is how one can interpret at least one example given in the *Phaedo*, which is said not to meet the requirements imposed on a correctly proceeding anamnetic process.[55] It is about Simmias' reminder from his image. In this case, there occurs at a certain time interval first a sensory recognition of the sensory characteristics whose combination is referred to as 'Simmias', and then the recognition of a similar combination of characteristics reproduced in the image. Because even then the similarity has no character of identity, and so in perception one notices the difference between that which is perceived and that which is reminded (the difference between the image and the painted object). As a result of the other recognition there occurs a reminder,

[51] It is interesting that *pseudēs* appears in the *Phaedo* only once (90b 8).

[52] G.B. Kerferd, *The Sophistic Movement*, Cambridge 1981, p. 61.

[53] Plato, *Phaedo*, 73a–d.

[54] It is in the area of geometry that the concept of anamnesis is presented in Plato's *Meno* (80d–86c); for more on this *cf.* e.g. R.E. Allen, 'Anamnesis in Plato's *Meno* and *Phaedo*', *Review of Metaphysics* 13 (1959–1960), pp. 165–174; S.-I. Lee, *Anamnesis im "Menon" Platons Überlegungen zu Möglichkeit und Methode eines den Ideen gemaßen Wissenerwerbes*, Frankfut am Main–Berlin–Bern–Bruxelles–New York–Oxford–Wien 2001.

[55] J.L. Ackrill, *Anamnesis in the "Phaedo": Remarks on 73c–75c*, [in:] J.L. Ackrill, *Essays on Plato and Aristotle*, Oxford 1997, p. 22.

which according to the general postulate has the nature of knowledge, although it does not meet the condition of anamnesis *sensu stricto*, i.e. the falling within another scope of knowledge (*heteron ennoēsē hou mē hē autē epistēmē all' allē*[56]). Perhaps this type of knowledge includes for instance the ability to give poison to those sentenced to death (Socrates refers to that who is to give him poison as *epistemon*[57]). Another condition imposed on knowledge, namely its substantiation (*logon dounai*),[58] would apply to both types of knowledge, and obviously such substantiation will be of an entirely different character. In the case of anamnesis *sensu stricto*, the substantiation consists in the existence in the soul/memory of general concepts that are to be present in it all the time, and were forgotten upon incarnation and have to be brought back from oblivion. In the case of anamnesis *sensu largo*, however, the substantiation only consists in a number of earlier recognitions of the objects of sense perceptions and in an 'external' similarity between the object and the reproduction. If the above analysis is correct, then we have to do with internal demarcation, according to which knowledge is divided into that which is based on the general and that which is based on sensory experience.

With the knowledge of the first type the last concept I would like to consider, is connected. The concept is *phronēsis*. It is usually translated as 'wisdom'[59] or 'thinking'.[60] Before the *Phaedo*, it seems to function above all as a concept from the realm of ethics, which has not become well grounded in the ontological concept. It consists in knowing that which is good and which drives man's actions in such a way that he achieves happiness. This is obstructed by mistaken convictions, which should be overcome by using the elenctic approach.[61] In the *Phaedo*, its slightly different status is clearly seen. On the one hand, it is present in the soul in the form of general concepts, which is proved by an argument from anamnesis,[62] and, on the other hand, already after another incarnation, the soul has to acquire – or rather recover – it together with the truth, which it does by such activities as *dianoesthai* or *logismos*.[63] Full recovery is only possible if the soul is entirely separated from the corporeal factor – then the soul itself encompasses in thought that among beings which is in itself (*hoti an noēsēi autē kath' hautēn auto kath' hauto tōn ontōn*).[64]

[56] Plato, *Phaedo*, 73c 8.
[57] *Ibidem,* 117a.
[58] *Ibidem*, 76d.
[59] So e.g. R. Legutko (Platon, *Fedon*, Kraków 1995) and G.M.A. Grube (Plato, *Phaedo*, [in:] Plato, *Complete Works*).
[60] Platon, *Phédon*, trad. L. Robin, [in:] Platon, *Ouvres complètes*, t. IV 1, Paris 1965. W. Witwicki (Platon, *Fedon*, [in:] Platon, *Dialogi*, Warszawa 1993) translates *phronēsis* as *poznanie*, while T. Ebert (Platon, *Phaidon*) as *Einsicht*.
[61] B. Rosenstock, *From Counter-Rhetoric to Askesis: How the "Phaedo" Rewrites the "Gorgias"*, [in:] B.D. Schildgen (ed.), *The Rhetoric Canon*, Detroit 1997, p. 85.
[62] Plato, *Phaedo,* 76c.
[63] *Ibidem*, 65e 6–66a 8.
[64] *Ibidem*, 83b 1–2.

Another aspect of *phronēsis* noticeable in the *Phaedo* is its ethical dimension. It comes to the fore above all in the monetary exchange metaphor.[65] It is the very *phronēsis* that is the right means/measure (*orthos nomisma*) of a rightly made change (*orthē allagē*) aimed at achieving virtue (*aretē*). As a means/measure, it becomes a criterion in obtaining or realizing (buying) or giving, in the realizing in somebody else (selling) of individual true virtues (*alēthēs aretē*): valour, moderation, justice. Within the context of the above-mentioned deliberations concerning the way of conduct of antilogicians (although in the very dialogue they are mentioned later), it seems important that *phronēsis* constitutes a criterion external in relation to obtained or given virtues, because if somebody wants to realize a virtue without having such an external measure at their disposal, and only possessing an immanent criterion, he cannot justify rationally why he for instance prefers a given fit of passion to another one or a fit of passion with a higher intensity to another one with a lower intensity. And even if such a person seems to practise one of virtues thanks to their conduct, such conduct and such a virtue have a nature of a certain fiction (*skiagraphia*)[66] – something slavish, unhealthy and untrue. Plato recognizes then that there are two types of virtues, and the virtues of non-philosophers do not really deserve the denomination, as they are not substantiated in *phronēsis*. What is more, they do not meet another important condition that is imposed on the virtues and the very *phronēsis* – namely, the being of that which purifies (*katharsis, katharmos*).

In the *Phaedo* there also comes to the fore the peculiar attitude that is present in man when he becomes a philosopher, and which will be so strongly emphasized in the *Symposium*. It is about a philosopher as a lover, although not a lover of wisdom (*sophia*), because such wisdom has turned out not to be it, but a lover of (*erastēs*) *phronēsis*. A lover is above all aware that he will fully achieve the object of his love only after he dies.[67]

I would venture a thesis that *phronēsis* perceived as a certain type of wisdom replaces – in Plato's philosophical construction – the traditional *sophia*, both as regards the theoretical aspect and the practical-and-moral aspect. As such, it is the objective of human cognitive aspirations, and together with *epistēmē sensu stricto* it constitutes the gnoseological postulate, according to which there exists knowledge/wisdom, although it is not fully achievable during our intellectual activity as long as we are alive. Perhaps it is the best hypothesis because of which other hypotheses are worth considering, e.g. that whether the soul is immortal.

[65] *Ibidem,* 89a 6–89e 5. The fragment with the metaphor of the exchanging has been subject to numerous interpretations, the older ones of which are presented e.g. by J.V. Luce ('A Discussion of *Phaedo* 69a 6–c 2', *Classical Quarterly* 38 (1944), pp. 60–64; *cf.* also R. Weiss, 'The Right Exchange: *Phaedo* 69a 6–c 3', *Ancient Philosophy* 7 (1987), pp. 57–66; R. Legutko, *Komentarz*, [w:] Platon, *Fedon*, pp. 83–85; T. Ebert, *Kommentar*, [in:] Platon, *Phaidon*, pp. 148–149.

[66] *Skiagraphia* refers to the chiaroscuro or perspective method of painting, which is associated with the metaphoric meaning of 'illusion'. In this last meaning, it also occurs in the*Republic* (*skiagraphia aretēs* – 365c 4; *skiagraphia epithemenē goēteias* – 602d 2), while already in the *Theaetetus* it reflects a certain distance, thanks to which we can see something more clearly (208e 7–10).

[67] Plato, *Phaedo*, 66e 2–4; 67e 6–68a 3.

Apart from determining that it is the highest value, it is difficult to establish whether it constitutes it on its own or together with other virtues, and what its relation to pleasure is, if one assumes that it is only juxtaposed with carnal pleasures.[68]

One can conclude, then, that most probably during the writing of the *Phaedo*, knowledge in the scrict sense, based on a being in itself, is only a postulate, a kind of project, because Socrates/Plato seems to present it in the conditional: if there really is some true, reliable and cognizable *logos* (i.e. justification) (*ei ontos dē tinos alēthous kai bebaiou logou kai dunatou katanoēsai*), then we have access to the knowledge and the truth about the beings.[69] Similarly, the disquisition on the immortality of the human soul is also accepted conditionally (*eikos*). The status of the very hypothetical method is also not entirely clear. We learn when following the autobiographical disquisition in the dialogue that having become fascinated with *peri phuseōs*-type study, Socrates/Plato *was convinced* that it is wonderful to know the causes of all things.[70] The disappointment with this type of explanation may have been alleviated by Anaxagoras' concept, because once again Socrates/Plato *came to believe*[71] that this philosophy would provide answers to his doubts. Again, he was disappointed. He took more effort then, turned his attention to another sphere and again *came to believe*[72] that he should make use of *logos*. All three attempts are described as something that took place, but are no longer (*aoristus* is used there). Thus, it is not known for certain what philosophical position Socrates/Plato represented while recounting his philosophical way – whether he had rejected the hypothetical method or whether he was using it to continue his philosophical search. The hypothetical nature of the deliberations is also attested to by the frequency of use of terms expressing doubt, above all probability (*eikos*), which is not always translated correctly. Knowledge is then a distinguished area of *doksa*, which can also encompass an improved (if such a version existed) version of knowledge *peri phuseōs*, which would anyway be a knowledge with a status inferior to that based on *logos*. Within *doksa*, and beyond *epistēmē*, a number of various views would function, such as the traditional explanation of *peri phuseōs*, untrue views of the virtue or the concept of antilogicians.

[68] J.C.B. Gosling, C.C.W. Taylor, 'The Hedonic Calculus in the *Protagoras* and the *Phaedo*: a Reply', *Journal of the History of Philosophy* 28 (1990), pp. 115–116.
[69] Plato, *Phaedo*, 90c 8–d 7.
[70] *Ibidem*, 96a 8–9.
[71] *Ibidem*, 97c 3.
[72] *Ibidem*, 99e 4.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

MIKOŁAJ DOMARADZKI
Adam Mickiewicz University, Poznań

Lucius Annaeus Cornutus' Ethnographic Investigations Into Mythology*

Abstract

The purpose of the present article is to show that the hermeneutical activity of Lucius Annaeus Cornutus is best characterized as "ethnographic" rather than merely "allegorical" or "etymological". Without denying the presence of both these dimensions in the philosopher's exegeses, the paper suggests that Cornutus' analyses aimed first and foremost to excavate the ancient world picture that the philosopher believed to underlie the theology transmitted by Homer and Hesiod. Thus, the philosopher regarded conventional mythology and traditional religion as sources of information about the primeval accounts of the cosmos: his analyses of various etymologies discovered not merely the origin of the word in question but also the origin of the ancient cosmological conceptions. Consequently, interpreting myths was for Cornutus tantamount to gaining profound insights into the pristine theology that was skillfully developed by the wise men of antiquity and poorly transmitted by the poets. Cornutus' hermeneutics built on the assumption that interpreting mythology provided the interpreter with a better understanding of not only the ancient world but also the present one.

Etymological analyses of the gods' names and epithets belong undoubtedly to one of the most interesting and, at the same time, controversial developments within Stoic philosophy. While the purpose of these investigations was to extract the ancient world view that according to the Stoics underlay the theology transmitted by the poets, such analyses formed an integral part of the Stoic physics. By viewing traditional mythology as a prefiguration of their own cosmological doctrines, the Stoics came to treat conventional myths as important sources of information on the primordial beliefs about the gods and the cosmos. Thus, examining a given etymology would provide the philosophers not only with information about the origin of the word in question (in this regard their analyses were frequently

* This paper is a substantially revised version of a text published in the *Studia Philosophica Wratislaviensia* VII [2] (2012), pp. 7–25.

naïve and fantastic), but also with information about the genesis of the pristine world picture that was mirrored in the particular etymologies.

Whilst such ethnographic interests can already be found in some of the extant testimonies on the hermeneutical activity of the early Stoics,[1] the present paper will focus exclusively on the hermeneutics of Lucius Annaeus Cornutus. This Stoic philosopher, who lived in the first century of our era, wrote a very interesting work, whose title ΕΠΙΔΡΟΜΗ ΤΩΝ ΚΑΤΑ ΤΗΝ ΕΛΛΕΝΙΚΗΝ ΘΕΟΛΟΓΙΑΝ ΠΑΡΑΔΕΔΟΜΕΝΩΝ is customarily translated into English as *Compendium of (the Traditions of) Greek Theology.*[2]

Cornutus' work provides us with a unique insight into the *ethnographic* nature of Stoic *etymologizing*. At the outset, however, it needs to be emphasized that characterizing Cornutus' etymological interpretations as "ethnographic" is a certain simplification.[3] It goes without saying that when interpreting ancient thinkers, one should refrain from imposing contemporary categories and concepts on their intellectual work. Although at first sight such "reconstructions" might seem quite "natural", they, nevertheless, inevitably distort the objects of interpretation. Thus, we need to emphasize that employing the term "ethnography"

[1] It needs to be emphasized that there is a heated controversy as to how Stoic approach to mythology ought to be classified. For scholars who have some reservations regarding the allegorical dimension of Stoic hermeneutics see especially A.A. Long, *Stoic Readings of Homer*, [in:] A.A. Long (ed.), *Stoic Studies*, New York 1996, pp. 58–84; *cf.* also P. Steinmetz, 'Allegorische Deutung und allegorische Dichtung in der alten Stoa', *Rheinische Museum für Philologie* 129 (1986), pp. 18–30 and D. Dawson, *Allegorical Readers and Cultural Revision in Ancient Alexandria*, Berkeley 1992, pp. 23–38. While the majority of scholars are inclined (rightly, in my opinion) to characterize the Stoics' hermeneutics as in one way or another "allegorical", it would be virtually impossible to enumerate all the relevant studies. See, however, the following works: J. Tate, 'Cornutus and the Poets', *Classical Quarterly* 23 (1929), pp. 41–45; F. Buffière, *Les Mythes d'Homère et la pensèe grecque*, Paris 1956, pp. 137–154; J. Pépin, *Mythe et allégorie: Les origines grecques et les contestations judéo-chrétiennes*, Paris 1976, pp. 125–167; J. Whitman, *Allegory. The Dynamics of an Ancient and Medieval Technique*, Cambridge 1987, pp. 31–47; G. Most, *Cornutus and Stoic Allegoresis: A Preliminary Report*, [in:] W. Haase (hrsg.), *Aufstieg und Niedergang der römischen Welt*, Bd. II 36.3, Berlin–New York 1989, pp. 2014–2065; C. Blönnigen, *Der griechische Ursprung der jüdisch-hellenistischen Allegorese und ihre Rezeption in der alexandrischen Patristik*, Frankfurt am Main 1992, pp. 22–42; L. Brisson, *Introduction la philosophie du mythe*, vol. 1:*Sauver les mythes*, Paris 1996, pp. 61–72; G.R. Boys-Stones, *The Stoics' Two Types of Allegory*, [in:] G.R. Boys-Stones (ed.),*Metaphor, Allegory and the Classical Tradition: Ancient Thought and Modern Revisions*, Oxford 2003, pp. 189–216 and P.T. Struck, *Birth of the Symbol: Ancient Readers at the Limits of Their Texts*, Princeton 2004, pp. 111–151. I have argued that 1) the Stoics' hermeneutical activity comprises an allegorical as well as ethnological dimension and 2) that Cornutus' exegetical activity continues the hermeneutical efforts of the early Stoics in: M. Domaradzki, 'From Etymology to Ethnology. On the Development of Stoic Allegorism', *Archiwum historii filozofii i myśli społecznej* 56 (2011), pp. 81–100.

[2] In the present paper, the text is cited after: C. Lang, *Cornuti theologiae Graece compendium*, Leipzig 1881.

[3] For the sake of our considerations, Cornutus' hermeneutical activity could also be described as "ethnological" or 'anthropological". Long is clearly right when he classifies Cornutus as an "ethnographer" and "cultural anthropologist", A.A. Long, *Stoic Readings...*, p. 73. The view put forward in this paper is nicely expressed by the scholar's following diagnosis: "the Stoics treated early Greek poetry as ethnographical material and not as literature", A.A. Long, *Stoic Readings...*, p. 82. I cannot, however, agree with Long's denial of the allegorical dimension of Stoic hermeneutics. *Cf. infra* note 34.

with reference to Cornutus' hermeneutics is, in fact, tantamount to cramming the thinker into modern and, thereby, alien framework.[4] Notwithstanding this, such a simplification appears to be justified by the fact that an uncontroversial classification of Cornutus' hermeneutical activity is far from easy. As will be argued below, Cornutus aims to etymologically "excavate" the primeval world picture that has been preserved in the gods' names and epithets. That is why his approach invites the label of "ethnography".

The present considerations will be structured in the following way: firstly, I will briefly examine the possibility of Aristotle's influence on Stoic ethnography; then, I will move on to discussing the relation between the Stoics' theory of language and their recourse to etymology as a basic interpretative tool; finally, I will show in what sense Cornutus' investigations can be characterized as "ethnographic". An assessment of Cornutus' cultural relevance will conclude my considerations. The purpose of this paper will be to show that Cornutus' etymological analyses serve the purpose of eliciting the profound ancient wisdom that lies beneath the veneer of the naïve and primitive language of mythology: through his etymologizing, Cornutus wants to demonstrate that anthropomorphic and often fatuous myths *allegorically* express a valuable cosmology that frequently anticipates the physical and theological views of the Stoics.

1. Aristotle and the emergence of Stoic ethnography

When trying to make sense of Stoic hermeneutics, Aristotle's account of the cyclical recurrences of human civilizations is a good place to begin.[5] According to this account "one must acknowledge that not once, not twice, but countless times the same beliefs come to us" (οὐ γὰρ ἅπαξ οὐδὲ δὶς ἀλλ᾽ ἀπειράκις δεῖ νομίζειν τὰς αὐτὰς ἀφικνεῖσθαι δόξας εἰς ἡμᾶς).[6] This shows the Stagirite to have believed

[4] Especially in light of the fact Cornutus modestly stresses (76.6–7) that he merely confines himself to "recapitulating" (ἐπιτετμημένως) the views of "the older philosophers" (τοῖς πρεσβυτέροις φιλοσόφοις).

[5] Although the influence of Aristotle on the development of Stoic allegoresis has been frequently discussed, the scholars are far from reaching any consensus as to where exactly the influence should be located. For example, Wehrli finds the traces of Aristotle's influence already in Chrysippus, *cf.* F. Wehrli, *Zur Geschichte der allegorischen Deutung Homers im Altertum*, Borna–Leipzig 1928, pp. 56–57. Tate, on the other hand, suggests that the Stagirite's influence is limited to Cornutus only, *cf.* J. Tate, *Cornutus...*, pp. 43–44. Lastly, Struck expresses his doubts as to the importance of Aristotle's civilization theory for the formation of any Stoic's approach to mythology, *cf.* P.T. Struck, *Birth of the Symbol...*, p. 150 n. 19. While clearly Aristotle's influence on Stoic ethnography "darf [...] nicht überschätzt werden" (F. Wehrli, *Zur Geschichte...*, p. 57), total skepticism in this regard does not appear to me particularly attractive. A well-balanced discussion of this issue is to be found in: J. Pépin, *Mythe et allégorie...*, pp. 121–124; L. Brisson, *Introduction...*, pp. 58–60 and G.R. Boys-Stones, *The Stoics' Two Types...*, pp. 191–192. I wholeheartedly agree with the following opinion: "A l'inverse de Platon, Aristote ne voit donc pas dans le mythe d'Homère et d'Hésiode une fiction purement arbitraire et dépourvue de toute portée didactique; [...] le mythe est pour lui l'expression allégorique d'un enseignement *rationnel*, qualité sur laquelle il insiste", J. Pépin, *Mythe et allégorie...*, pp. 123–124. In a very similar vein, Brisson emphasizes: "A la différence de Platon, Aristote n'adopte pas à l'égard du mythe une attitude de rupture radicale", L. Brisson, *Introduction...*, p. 59.

[6] Aristotle, *De caelo*, 270b 19–20. This view is also expressed in the *Meteorology* (339b 27–30), where the same beliefs "return cyclically" (ἀνακυκλεῖν) many times and in the *Politics* (1329b 25–27), where various things are repeatedly "discovered" (εὑρῆσθαι).

that knowledge is gradually and cumulatively obtained at various stages of the development of human civilization. While certain views appear cyclically (i.e., they are repeatedly discovered), the cyclicality points to an important affinity between philosophy and mythology: for Aristotle, myth should be taken as a valuable prefiguration of philosophical knowledge.[7]

From the Stagirite's perspective, then, the process of gaining knowledge can, at least to some extent, be regarded as a rediscovery and reworking of ideas that were articulated in the days of old. The idea of cyclically recurring views builds on the assumption that at the earliest stages of human civilization there lived certain wise men who acquired fairly reliable knowledge of the world and its mechanisms. This conviction is most clearly expressed in the *Metaphysics*, where Aristotle declares that: "whilst in all probability every art and philosophy has repeatedly reached its peak ability, upon which it perished again, the [particular] beliefs have been preserved to the present day as remnants of those" (κατὰ τὸ εἰκὸς πολλάκις εὑρημένης εἰς τὸ δυνατὸν ἑκάστης καὶ τέχνης καὶ φιλοσοφίας καὶ πάλιν φθειρομένων καὶ ταύτας τὰς δόξας ἐκείνων οἷον λείψανα περισεσῶσθαι μέχρι τοῦ νῦν).[8] Let us note that the particular beliefs can resurface in different periods of time, since they are contained in the society's mythology that has been transmitted by the poets. Hence, the ancient wisdom can be reconstructed by examining myths that conceal profound truths and precious intuitions articulated by men of antiquity.

Cornutus bases his investigations on a very similar assumption: in the symbolic and enigmatic works of the poets the profound wisdom of the ancients[9] has been handed down[10] to posterity. Thus, with regard to Hesiod's genealogy, Cornutus makes the following comment: "some parts of it were taken by him from the ancients, whereas other parts were added by him in a more mythical way; and in this way most of the ancient theology was corrupted" (τὰ μέν τινα [...] παρὰ τῶν ἀρχαιοτέρων αὐτοῦ παρειληφότος, τὰ δὲ μυθικώτερον ἀφ᾽ αὑτοῦ προσθέντος, ᾧ τρόπῳ καὶ πλεῖστα τῆς παλαιᾶς θεολογίας διεφθάρη).[11] Cornutus diagnoses here that Hesiod distorted the original theology and that his distortions must have resulted from his inability to fathom the depths of the ancients' physics and cosmology. While Cornutus believes that his task consists precisely in excavating this profound wisdom, the philosopher also assumes that at least to some extent philosophical accounts of reality can be regarded as "rationalized translations" of ancient myths. Naturally, the translations are always more accurate as they are gradually distilled from the various irrational and anthropomorphic concepts that were unnecessarily added by the poets. Yet, there is a direct link between philosophy and mythology so that ultimately the former is but a refinement of the latter.

[7] This is spectacularly attested by the philosopher's famous remark (*Metaphysica*, 982b 18) that there is a certain intellectual affinity between a "lover of myth" (φιλόμυθος) and a "lover of wisdom" (φιλόσοφος), i.e., a philosopher.

[8] Aristoteles, *Metaphysica*, 1074b 10–13.

[9] While οἱ ἀρχαῖοι are mentioned already at the very beginning of the work (2.18), their authority is continually cited throughout the whole book.

[10] Cornutus' favorite verb is παραδίδωμι, appearing already in the very title of work.

[11] Cornutus, *Compendium*, 31.14–17.

From Aristotle's perspective, too, philosophical knowledge originates from mythology. The Stagirite observes, then, that the ancients handed down to posterity a valuable tradition "in the form of a myth" (ἐν μύθου σχήματι), stressing also the fact that a great deal of this tradition had to be given its mythical form only "to persuade the many and to be useful for the laws and for the general good" (πρὸς τὴν πειθὼ τῶν πολλῶν καὶ πρὸς τὴν εἰς τοὺς νόμους καὶ τὸ συμφέρον χρῆσιν).[12] While according to Aristotle the precious and valuable intuitions articulated by men of antiquity were *purposefully* camouflaged in the various mythical formulations, Cornutus believes them to have become *inadvertently* contaminated by the poets who transmitted them. That is why Cornutus stresses the necessity of approaching myths in the right way. On the one hand, the philosopher admonishes that one should not "conflate the myths" (συγχεῖν τοὺς μύθους), "transfer the names from one [myth] to another" (ἐξ ἑτέρου τὰ ὀνόματα ἐφ᾽ ἕτερον μεταφέρειν) or rashly "consider [them] irrational" (ἀλόγως τίθεσθαι).[13] One the other hand, Cornutus emphasizes that "something has been added to the genealogies that have been handed down [to us] by those who failed to understand what [the myths] hint at enigmatically" (τι προσεπλάσθη ταῖς παραδεδομέναις κατ᾽ αὐτοὺς γενεαλογίαις ὑπὸ τῶν μὴ συνιέντων ἃ αἰνίττονται).[14] This passage shows that for Cornutus (as for Aristotle) there is an important continuity between the mythical and philosophical description of the world. The continuity is, nevertheless, frequently obscured by the various contaminations that come from the poets.

Cornutus firmly believes that he can extract the "ancient theology" (παλαιὰ θεολογία) from the distorted transmission of the poets. In this context, one should pay particular attention to Cornutus' diagnosis that the poets' contaminations are due to their incapacity to comprehend the symbols and enigmas that have been used for conveying the ancient wisdom. The original αἰνίττονται suggests that the ancient mythmakers *spoke enigmatically* in the sense that they *hinted at* something that needs to be appropriately interpreted. Cornutus uses the word, as he seems to be deeply convinced that *speaking through enigmas* is characteristic of everyone who possesses profound knowledge and thorough understanding of things that can actually only be expressed in such symbols and riddles. That is why in the final part of his work the philosopher asserts that "the ancients were no common men but able to understand the nature of the world and inclined to philosophize about it through symbols and enigmas" (οὐχ οἱ τυχόντες ἐγένοντο οἱ παλαιοί, ἀλλὰ καὶ συνιέναι τὴν τοῦ κόσμου φύσιν ἱκανοὶ καὶ πρὸς τὸ διὰ συμβόλων καὶ αἰνιγμάτων φιλοσοφῆσαι περὶ αὐτῆς εὐεπίφοροι).[15]

This reveals what Cornutus perceives as his task: to properly interpret the "symbols" and "enigmas" that obfuscate the ancient theology. It is worth noting that Cornutus does not interpret the (evidently fallible) poets in accord with their presumed intentions. This is understandable in light of the fact that (according to his view) the poets do not fully understand what they actually convey. Thus,

[12] Aristoteles, *Metaphysica*, 1074a 38–1074b 5.
[13] Cornutus, *Compendium*, 27.19–28.2
[14] *Ibidem*, 27.20–28.1.
[15] *Ibidem*, 76.2–5.

the philosopher believes that the ancient wisdom can be recovered not *owing to* but rather *in spite of* the poets. That is why Cornutus' tool for unravelling this wisdom is etymology: as using language is automatic and unconscious, etymological analyses are the best way to excavate the world picture that underlies the particular vocabulary.

2. Etymology as an interpretative tool

Cornutus employs etymology to explore the relation between the form of the word and its meaning (which results from the underlying world view). For example, the philosopher derives the name "Prometheus" from "the foresight of the world's soul" (ἡ προμήθεια τῆς ἐν τοῖς ὅλοις ψυχῆς) that is also equated with "the providence" (ἡ πρόνοια).[16] Analyses of this sort show that for Cornutus the gods' names and epithets are not contingent and arbitrary. When seeking to uncover the ancient theology, Cornutus embraces the Stoics' view of language according to which the relation between the names and their referents is *natural* and not purely conventional.[17] This view is of paramount importance for understanding Cornutus' ethnography, for it provides a *direct* link between studying the words of a language and studying the world picture preserved in the vocabulary of that language. Let us, therefore, briefly examine this.

When referring the Stoics' view on the origin of names, Origen reports the philosophers to have maintained that "the first sounds imitate the things of which the names are said" (μιμουμένων τῶν πρώτων φωνῶν τὰ πράγματα, καθ' ὧν τὰ ὀνόματα), upon which he explains that this view entails recourse to etymology.[18] Cornutus continues the early Stoics' etymological investigations into why words have the form they do.[19] The philosopher believes that language provides us with

[16] *Ibidem*, 32.1–3.

[17] That is why Stoic use of such terms as "symbol" or "allegory" must not be equated with modern understanding of these concepts. In this respect see M. Domaradzki, 'Symbol i alegoria w filozoficznej egzegezie stoików', *Filo-Sofija* 13–14 (2011), pp. 719–736. In what follows, I draw on some of the findings presented there. The close connection between the Stoics' view of language and the philosophers' hermeneutics has been thoroughly discussed by C. Blönnigen, *Der griechische Ursprung...*, pp. 23–27; D. Dawson, *Allegorical Readers...*, pp. 28–35 and P.T. Struck, *Birth of the Symbol...*, pp. 123–141.

[18] Origenes, *Contra Celsum*, I 24 (= *SVF* II 146 [J. von Arnim (ed.), *Stoicorum Veterum Fragmenta*, vol. I–III, Stuttgart 1968]).

[19] Buffière was clearly right in characterizing such etymology as "moyen d'exégèse", F. Buffière, *Les Mythes d'Homère...*, p. 60. While Cornutus speaks (2.4) of "analyzing the origin" (ἐτυμολογοῦσι) of a given god's name, the value of etymological exegeses is thoroughly discussed in Plato's *Cratylus*, which is commonly regarded as "[t]he first work that deals with etymology, and uses it systematically", H. Peraki-Kyriakidou, 'Aspects of Ancient Etymologizing', *Classical Quarterly* 52 (2002), p. 478. Two points need to be stressed here. First of all, in antiquity, etymology was considered to be a reliable source of information about the cosmos and its mechanisms. Thus, in connection with Plato's etymological analyses Sedley aptly diagnoses that "*no one* in antiquity ever thought Plato was joking", D. Sedley, *Plato's Cratylus*, Cambridge 2003, p. 37. Naturally, it remains highly debatable whether and, if so, to what extent Plato himself was inclined to take seriously such etymological analyses. Yet, even if one refuses to agree with Sedley that the *Cratylus* be read as "a serious exploration of etymology and its lessons" (*ibidem*, p. 172), it cannot be disputed that such a characterization fits perfectly the Stoics. This is closely connected with another matter than needs to be noted here. Cornutus shares with Plato

the best insight into the way in which the users of a given language understood the world, since the names of things reflect how those things were comprehended. Hence, when putting forward his (often fantastic) etymologies, Cornutus aims to show that etymological investigations provide us with an access to the ancients' world picture (i.e., the profound wisdom of their theology), for establishing the relevant etymological connections makes it, subsequently, possible to link language with various cultural practices, rituals, rites.

To understand the specificity of the Stoics' (and Cornutus') approach we need to briefly consider Augustine's *De dialectica*, since this testimony offers the clearest exposition of Stoic original use of etymology[20]. According to the testimony, the Stoics were convinced that it is possible to explain the origin of every single word.[21] Augustine relates further that in their etymological investigations the Stoics would look for a point where "the thing corresponds with the sound of the word in some similarity" (*res cum sono verbi aliqua similitudine concinat*), beginning, thereby, with such onomatopoeias as "clanging" (*tinnitus*), "neighing" (*hinnitus*), "bleating" (*balatus*) etc.[22] Naturally, the Stoics were aware of the fact that the richness of natural languages does not exhaust itself in onomatopoeias. When explaining that "these words sound like the things themselves which are signified by these words" (*haec verba ita sonare, ut ipsae res quae his verbis significantur*), Augustine reports the Stoics to have realized that "there are things that do not sound [in any particular way]" (*sunt res quae non sonant*), upon which they posited "the similarity of touch to apply to them" (*in his similitudinem tactus valere*).[23]

While the notion of similarity of "touch" refers to the *direct* effect of things on our senses, the Stoics took the effect to be *reflected* in the particular words. Augustine expands upon this idea, stressing that according to the Stoics' position the things "smoothly or roughly touch the sense, as the smoothness or roughness of the letters touches the hearing" (*leniter vel aspere sensum tangunt, lenitas vel asperitas litterarum ut tangit auditum*).[24] Hence, when explaining that for the Stoics "the things themselves affect us in the same way as the words are experienced" (*res ipsae afficiunt, ut verba sentiuntur*), Augustine illustrates this argument with the example of "honey" (*mel*), which itself affects the taste "pleasantly" (*suaviter*), as it "smoothly touches the hearing with its name" (*leniter nomine tangit auditum*).[25] In conclusion, Augustine makes it clear that the Stoics regarded these cases as "the cradle of words" (*cunabula verborum*), arguing that if "perception

the belief that etymological investigations have a didactic as well as pedagogical value, *cf.* in this respect H. Peraki-Kyriakidou, *Aspects*..., p. 481. As Cornutus embraces the view of etymology that emerges from Plato's *Cratylus*, he repeats some of Plato's etymologies. *Cf. infra* notes 39 and 42.

[20] See especially C. Blönnigen, *Der griechische Ursprung* ..., pp. 24–27 and P.T. Struck, *Birth of the Symbol*..., pp. 125–126. In what follows, I use the text from B.D. Jackson, *Augustine. De Dialectica*, Dordrecht 1975, albeit I frequently modify the translation.

[21] Augustine, *De dialectica*, VI 9: *Stoici autumant,* [...] *nullum esse verbum, cuius non certa explicari origo possit*.

[22] *Ibidem*, VI 10.

[23] *Ibidem*.

[24] *Ibidem*.

[25] *Ibidem*.

of things concords with perception of sounds" (*sensus rerum cum sonorum sensu concordarent*), then "the license of naming must proceed from this point to the similarity of things themselves to each other" (*[h]inc ad ipsarum inter se rerum similitudinem processisse licentiam nominandi*).[26]

The Stoics' assumption about the isomorphism between language and external reality entails that for the philosophers the world and words are in complete harmony: names reproduce reality, upon which there is a *natural* bond between words and their referents. The idea that words mimic the world is most clearly expressed in the Stoic assumption that sense perceptions translate *directly* to the names of things. While words are, thus, formed in conformity with how their referents are perceived, one could summarize the Stoics' position by saying that words express the world, because the world imprints itself in words.

Cornutus embraces the Stoics' theory of language: the philosopher also believes that the names were formed in accord with the above discussed isomorphism between language and external reality. Consequently, Cornutus, too, believes that investigating the origin of the words is tantamount to investigating their underlying perceptions. At this point it needs to be noted that the Stoics' position makes etymology a natural tool for examining not only the underlying perceptions, but also the underlying conceptions.

We know that the Stoics believed the meanings of words (and the corresponding concepts) to be related to one another. Augustine makes it clear that the afore-analyzed "similarity of things and sounds" (*similitudo rerum et sonorum*), was not the only motivation for the origin of the words that the Stoics identified. He relates that the philosophers distinguished also between "the similarity of things themselves" (*similitudo rerum ipsarum*), "contiguity" (*vicinitas*), and "opposition" (*contrarium*).[27] A parallel testimony is provided by Diogenes Laertius, who apart from "experience" (περίπτωσις), mentions such concept-forming mechanisms as "similarity" (ὁμοιότης), "analogy" (ἀναλογία), "transposition" (μετάθεσις), "composition" (σύνθεσις), "opposition" (ἐναντίωσις) "transition" (μετάβασις) and – lastly – "privation" (στέρησις).[28]

As the Stoics accounted for the emergence of words (and the corresponding concepts) in terms of such mechanisms, it is hardly surprising that they should play an important role in Cornutus' etymological exegeses. Indeed, David Dawson has shown that Cornutus "draws on some of the modes of concept formation that Diogenes outlines".[29] The scholar cited the example of Ares, whose name Cornutus derives from the conceptions of "seizing" (αἱρεῖν), "killing" (ἀναιρεῖν) and "bane" (ἀρή), referring in his etymological interpretation to the mechanism of "opposition" (ἐναντίωσις).[30] Plenty of other examples could be given. Consider for instance

[26] *Ibidem*.
[27] *Ibidem*, VI 11.
[28] Diogenes Laertius, *Vitae philosophorum*, VII 52–53 (= *SVF* II 87).
[29] D. Dawson, *Allegorical Readers...*, p. 29.
[30] Cornutus, *Compendium*, 40.19–41.3. D. Dawson, *Allegorical Readers...*, p. 29. Although Dawson does not mention it, this etymology can be found in Chrysippus. Plutarch reports the philosopher to have derived the name Ares from the verb ἀναιρεῖν so that the god could stand for our aggressive instincts (*Amatorius*, 757b = *SVF* II 1094). *Cf.* also J. Pépin, *Mythe et*

Cornutus' analysis of Hermes. Having established a connection between Hermes' being called a "patron of public assemblies" (ἀγοραῖος) and his being a "guardian of those who speak in public" (ἐπίσκοπος [...] τῶν ἀγορευόντων), Cornutus diagnoses that "this was extended from the market onto those who buy or sell anything" (ἀπὸ τῆς ἀγορᾶς διατείνει καὶ εἰς τοὺς ἀγοράζοντάς τι ἢ πιπράσκοντας), since "everything should be done with reason" (πάντα μετὰ λόγου ποιεῖν δέοντος), upon which Hermes became a "custodian of merchandise" (τῶν ἐμποριῶν ἐπιστάτης) and was named "Commerce" (ἐμπολαῖος).[31]

While the associative character of such analyses is clear, it is worth stressing here that Cornutus' etymological analyses continue the Stoics' project of discovering the fundamental concepts that are reflected in the particular etymologies. It is precisely the assumption that words imitate the world that makes this approach possible: just as words mirror the various qualities of their referents, so do the gods' names and epithets reflect the ancient conceptions of the cosmos. The above cited interpretations (including the derivation of Prometheus from προμήθεια) show that for the Stoics there is nothing contingent about our language: analyzing words, names and epithets invariably reveals that the relation between *signifiants* and *signifiés* is not arbitrary. That is precisely why Cornutus assumes that investigating language of a given community provides valuable insights into how the vocabulary of that community reflects its archaic world picture.

3. Ethnographic dimension of Cornutus' etymological analyses

While Cornutus decodes the names and epithets of the gods so as to arrive at their underlying primeval world view, his hermeneutical activity frequently transmogrifies into an allegorical interpretation, since the ancient wisdom that is discovered in the course of his analyses often transpires to anticipate the cosmological views of the Stoics. For example, if Hesiod asserts that "at first (πρώτιστα) Chaos came into being",[32] then Cornutus suggests the passage be understood in such a way that "once fire was everything and will become [that] again in the [recurring] world cycle" (ἦν δέ ποτε [...] πῦρ τὸ πᾶν καὶ γενήσεται πάλιν ἐν περιόδῳ).[33] Thus, it is difficult to agree with those scholars who altogether deny the allegorical dimension of Cornutus' exegeses.[34] It seems best to say that Cornutus' approach to

allégorie..., pp. 129. A parallel explanation of Ares' name is given by Heraclitus the Allegorist (31.1), whereas somewhat different interpretations appear in Plato (*Cratylus*, 407d) and Lydus (*De mensibus* IV 34). Buffière offers an exhaustive discussion of the various etymological and allegorical interpretations of Ares that were put forward in antiquity, F. Buffière, *Les Mythes d'Homère...*, pp. 297–301.

[31] Cornutus, *Compendium*, 25.2–7. The epithet ἐμπολαῖος appears for example in Aristophanes (*Plutus*, 1155). In what follows, I make use of some of the findings presented in: M. Domaradzki, *From Etymology...*, pp. 95–99.

[32] Hesiodus, *Theogonia*, 116.

[33] Cornutus, *Compendium*, 28.10–12.

[34] The view that the Stoics' hermeneutics should not be characterized as allegoresis has been most forcefully put forward by A.A. Long, *Stoic Readings...*, pp. 59–60 and 71–82. With regard to Cornutus, Long stresses for example that the philosopher never refers to allegory and that, consequently, he is "an etymologist, not an allegorist" (p. 71). A similar opinion is expressed by Blönnigen, who likewise observes that Cornutus "bewusst keine eigene Allegorese betreibt", C. Blönnigen, *Der griechische Ursprung...*, p. 37. Still, in the passage cited above, Cornutus

mythology combines etymology, allegoresis and ethnography. In what follows, the uniqueness of this hermeneutics will be illustrated with Cornutus' interpretations of Rhea and Hermes.

When proceeding to reconstruct the genealogy of Rhea, Cornutus notes first that the Greek goddess seems to be "the same" (αὕτη) as the Syrian Atargatis.[35] Subsequently, he explains that the goddess is also called "the Phrygian" on account of how she is worshipped among the Phrygians, suggesting that there is a relation between the Galli practice of self-castration and "the Greeks' myth of the castration of Ouranos" (παρὰ τοῖς Ἕλλησι περὶ τῆς τοῦ Οὐρανοῦ ἐκτομῆς μεμύθευται).[36] What is noteworthy about these observations is that when searching for the ancient wisdom, Cornutus, on the one hand, does not confine himself to Greek mythology only and, on the other, posits the existence of various correlations between certain ancient beliefs and specific cult practices. Cornutus' interest in religious syncretism and his attentiveness to the cultural background of religious beliefs seem, thus, to justify characterizing his analyses as ethnographic: the philosopher makes references to numerous popular convictions and beliefs held by the ancient Greeks, Magi (i.e., Persians), Phrygians, Egyptians, Celts, Libyans and others,[37] for he believes that comparing *different* worldviews facilitates reconstructing the world picture of the ancient Greeks.[38]

Cornutus' approach seems to be fairly holistic: the philosopher assumes that the object of his investigations (the ancient wisdom preserved in the vocabulary of traditional theology) can best be understood if it is placed within the widest cultural context possible. While Cornutus views culture as a network of interrelated concepts and beliefs, his holistic approach manifests itself clearly in his interpretations: when reconstructing the genealogy of a particular deity, Cornutus makes references not only to language but also uses the "discovered" etymological connections to establish further links between language and various folk beliefs, mythical formulations, rituals, rites etc. He combines, thereby, etymology with ethnography, on the one hand, and etymology with allegoresis, on the other.

Etymological interpretation always provides a point of departure. Thus, Rhea's name is associated with a "flow" (ῥύσις), upon which Cornutus cites the belief that the goddess is "the cause of rainstorms" (τῶν ὄμβρων αἰτία).[39] As rainstorms are

does read the famous Stoic idea (*cf.* *SVF* I 98, 497, II 528, 596–632) into Hesiod's theogony and this interpretation seems to be a classical example of allegoresis. For a critical assessment of Long's position see: P.T. Struck, *Birth of the Symbol* ..., p. 113 and T. Tieleman, *Galen and Chrysippus On the Soul: Argument and Refutation in the De Placitis, Books II and III*, Leiden 1996, pp. 221–223.

[35] Cornutus, *Compendium*, 6.11–12.

[36] *Ibidem*, 6.16–19. *Cf.* also Lucianus, *De syria dea*, 15 and Lucretius, *De rerum natura*, II 611–617.

[37] Cornutus, *Compendium*, 26.7–11.

[38] Boys-Stones aptly observes (*ad loc.*) that in Cornutus one can find "a proper science of comparative mythology", G.R. Boys-Stones, *The Stoics' Two Types*..., p. 202. See also D. Dawson, *Allegorical Readers*..., p. 38. *Cf. supra* note 3.

[39] Cornutus, *Compendium*, 5.10–11. It is worth noting that in the *Cratylus*, Socrates associates (402a 4–b 4) the name "Rhea" with the "flow" or "current" (ῥοή) of a river and with the "streams" (ῥεύματα).

typically accompanied by thunder and lighting, the philosopher elucidates that Rhea was pictured as "delighted by drums, cymbals, horns and torch processions" (τυμπάνοις καὶ κυμβάλοις καὶ κεραυλίαις καὶ λαμπαδηφορίαις χαίρουσαν),[40] for "these rains clatter down from above" (ἄνωθεν οἱ ὄμβροι καταράττουσι).[41] Various popular beliefs about the goddess enable Cornutus to establish a correlation between the sounds of the cult instruments and the noises of a storm, on the one hand, and between the flashes of thunder and lighting and the glow of a torch, on the other.

The same approach is used in the case of Hermes. Cornutus derives the god's name from "contriving tales" (ἐρεῖν μήσασθαι), which he, then, equates with "speaking" (λέγειν),[42] suggesting that the god can owe his name to the fact that he is our "fortress" (ἔρυμα) and "stronghold" (ὀχύρωμα).[43] Irrespective of how fantastic Cornutus' analyses might seem, the philosopher's goal is to establish a connection between speech and language, on the one hand, and reason, on the other. That is why Cornutus identifies Hermes with reason (λόγος) that "the gods have sent to us from Heaven" (ἀπέστειλαν πρὸς ἡμᾶς ἐξ οὐρανοῦ οἱ θεοί), making, thus, "man the only rational animal on the earth" (μόνον τὸν ἀνθρώπων τῶν ἐπὶ γῆς ζῴων λογικόν).[44] Subsequently, Cornutus proceeds to show that the view of rationality that underlies the god's name is reflected in various cultural practices. For example, Cornutus suggests that the practice of heaping up stones beside Herms could be motivated by the desire "to symbolize that the uttered word consists of small parts" (πρὸς σύμβολον τοῦ ἐκ μικρῶν μερῶν συνεστάναι τὸν προφορικὸν λόγον).[45]

In a similar way, various other epithets of Hermes are connected with Stoic view of rationality. As we have noted, Hermes is a "patron of public assemblies" (ἀγοραῖος), a "guardian of those who speak in public" (ἐπίσκοπος [...] τῶν ἀγορευόντων), and – lastly – a "custodian of merchandise" (τῶν ἐμποριῶν ἐπιστάτης), for "all these things need to be done with reason" (πάντα μετὰ λόγου ποιεῖν δέοντος).[46] Likewise, Hermes is called "herald" (κῆρυξ),[47] because "through a loud voice he presents to the listeners the things signified according to the *logos*" (διὰ φωνῆς γεγωνοῦ παριστᾷ τὰ κατὰ τὸν λόγον σημαινόμενα ταῖς ἀκοαῖς),[48] "messenger" (ἄγγελος),[49] because "we learn the will of the gods from the con-

[40] *Cf.* also Lucretius, *De rerum natura*, II 618–619 and Ovidius, *Fasti*, IV 181–186.

[41] Cornutus, *Compendium*, 5.12–16.

[42] This etymology is also to be found in Plato's *Cratylus*. Having suggested that the name Hermes has to do with "speech" (λόγος)" and signifies that the god is an "interpreter" (ἑρμηνεύς), Socrates posits (407e 5–408b 2) a connection between such words as εἴρειν, ἐμήσατο, λέγειν and μηχανήσασθαι so as to derive the name Εἰρέμης from the fact that the god "contrived speaking and speech" (τὸ λέγειν τε καὶ τὸν λόγον μησάμενον) as well as "tales" (τὸ εἴρειν ἐμήσατο). *Cf.* also M. Domaradzki, *From Etymology...*, p. 96.

[43] Cornutus, *Compendium*, 20.21–23.

[44] *Ibidem*, 20.18–21.

[45] *Ibidem*, 24.11–25.2. When referring to προφορικὸς λόγος, Cornutus alludes to the Stoics' account of language and rhetoric, *cf.* e.g. *SVF* II 223.

[46] Cornutus, *Compendium,* 25.2–7. *Cf. supra* note 31.

[47] *Cf.* e.g. *Hymni Homerici*, IV 331.

[48] Cornutus, *Compendium*, 21.20–22.1.

[49] *Cf.* e.g. *Hymni Homerici*, IV 3.

cepts which have been bestowed upon us according to the *logos*" (τὸ βούλημα τῶν θεῶν γινώσκομεν ἐκ τῶν ἐνδεδομένων ἡμῖν κατὰ τὸν λόγον ἐννοιῶν)[50] and – finally – a "leader" (διάκτορος)[51] either since he is "piercing" (διάτορος) and "clear" (τρανός), or since "he leads our thoughts into the souls of our fellow men" (διάγειν τὰ νοήματα ἡμῶν εἰς τὰς τῶν πλησίον ψυχάς), which is also why "they sacrifice the tongues to him" (τὰς γλώττας αὐτῳ καθιεροῦσιν).[52]

These analyses show how Cornutus combines etymology and ethnography: on the one hand, the philosopher tries to show that the ancient thinkers must have perceived reason as piercing and clear (if it were to effectively fulfill its communicative functions), and on the other, he supports his considerations with a particular ritual practice. In order to justify his interpretation, Cornutus posits a connection between Hermes' names and images. To give yet another example, the philosopher explains that the god is sculpted "in a square shape" (τετράγωνος τῷ σχήματι), since the god is "steadfast" (ἑδραῖος) and "secure" (ἀσφαλής).[53] Thus, Cornutus suggests that the motive behind presenting Hermes in a quadrangle shape was that the ancient thinkers conceptualized reason as stable, solid and infallible.

Cornutus' analyses of Rhea and Hermes show that the philosopher proceeds from a simple (if arbitrary) etymological analysis and then moves on to more complex interpretations that build on extensive cultural knowledge. His approach is, therefore, holistic inasmuch as the ancient vision of the world (that Cornutus retrieves) is placed within as broad a cultural context as possible: the philosopher buttresses his considerations with references not only to the language of the ancient community, but also to its rites, rituals, images, etc. In other words, Cornutus believes that it is possible to extract the archaic conception of a deity by examining how its etymology is connected with some cultural practice(s). In the case of Rhea, it is the use of drums, cymbals, horns and torches that is supposed to imitate the image of the goddess. In the case of Hermes, Cornutus cites such cultural practices as sacrificing the tongues to the god, heaping up stones beside Herms and sculpting the god in a square shape. In both cases (as in other analyses), Cornutus attempts to reconstruct the whole conceptual framework that he takes to motivate the name and conception of a particular deity by examining the specific cultural context.

Cornutus' holistic approach comprises also allegoresis. The purpose of his various allegorical interpretations is, naturally, to show that the ancient worldview in one way or another anticipates Stoic philosophy. Thus, traditional mythology transpires often to be an allegorical prefiguration of the wisdom proclaimed by the Stoics. Here, we should note that whilst Cornutus' interpretation of Hermes only alludes to the Stoic idea of προφορικὸς λόγος, the philosopher's interpretation of Rhea makes an explicit and elaborate reference to the Stoics' physics. As a matter of fact, Cornutus turns this traditional myth into a full-fledged narrative that proves to allegorically express Stoic cosmology.

Cornutus' interpretation of Rhea builds on the connection between Kronos, who "swallows" (καταπίνειν) his children with Rhea and "time" (χρόνος), whose

[50] Cornutus, *Compendium*, 22.1–22.3.
[51] *Cf.* e.g. *Hymni Homerici*, IV 392.
[52] Cornutus, *Compendium*, 21.1–4.
[53] *Ibidem*, 23.11–13.

similarity consists in that all things that come into being in time are "consumed" (δαπανᾶται) by it.[54] While according to traditional accounts[55] Rhea prevented Kronos from devouring Zeus by feeding him with a stone wrapped in swaddling clothes, Cornutus suggests that "the swallowing be grasped differently" (ἄλλως εἴληπται ἡ κατάποσις), since in reality "the myth was composed about the generation of the world" (συντέτακται γὰρ ὁ μῦθος περὶ τῆς τοῦ κόσμου γενέσεως).[56] In Cornutus' allegorical interpretation, the world could only emerge when "the nature governing it" (ἡ διοικοῦσα αὐτὸν φύσις), i.e., Zeus, "matured and prevailed" (ἀνετράφη [...] καὶ ἀπεκράτησεν), whereas "this stone" (λίθος οὗτος), i.e., the earth, "was swallowed" (καταποθείς), i.e., "fixed firmly" (ἐγκατεστηρίχθη) as a "foundation" (θεμέλιος) for all things that come into being.[57] Hence, the myth that has Zeus banish Kronos from his kingdom and hurl him down to Tartarus is interpreted by Cornutus as an enigmatic hint[58] at "the ordering of the world's becoming" (ἡ τῆς τῶν ὅλων γενέσεως τάξις).[59]

Cornutus' allegoresis is based on the supposed etymological connection between the god's name (Κρόνος) and his "accomplishing" (κραίνειν), i.e., restricting "the flow of the matter surrounding the earth" (ῥύσιν τοῦ περιέχοντος ἐπὶ τὴν γήν).[60] Whilst Kronos is, therefore, associated by Cornutus with the force that "makes the exhalations finer" (λεπτοτέρας ποιήσασα τὰς ἀναθυμιάσεις), Zeus is interpreted by the Stoic as "the nature of the world" (κόσμου φύσις) that is responsible for "curbing the excessive rush of the change, putting it in bonds, and, thus, giving a longer life to the world itself" (τὸ λίαν φερόμενον τῆς μεταβολῆς ἐπέσχε καὶ ἐπέδησε μακροτέραν διεξαγωγὴν δοὺς αὐτῷ τῷ κόσμῳ).[61]

Evidently, then, Cornutus' allegorical interpretation of the myth about Ouranos' castration builds on the Stoics' physics. Kronos is interpreted as an *allegory* of the force that rarifies the matter and restricts its flow round the earth. Zeus, on the other hand, is identified with the force responsible for controlling the process initiated by Kronos, and, thereby, bringing the ultimate cosmic balance between all these cosmogonic transformations. Hence, Zeus' dethroning of Kronos signifies (allegorically) that chaos has been replaced with order.

The above discussed interpretations of Rhea and Hermes show that Cornutus is inclined to treat conventional mythology and traditional religion as products of a society at a given stage of development. For him, myths preserve the world picture of the ancients: their beliefs, convictions, values etc. Although Cornutus believes

[54] *Ibidem*, 6.20–7.5. *Cf.* Cicero, *De natura deorum*, II 64.

[55] Hesiodus, *Theogonia*, 485–491. See also Lucretius, *De rerum natura*, II 638–639.

[56] Cornutus, *Compendium*, 7.10–12.

[57] *Ibidem*, 7.12–17. Struck stresses (*ad loc.*) the similarity between this interpretation and the Orphic cosmogony presented in the Derveni Papyrus, P.T. Struck, *Birth of the Symbol...*, p. 147, n. 15. For a discussion of Cornutus' allegoresis see also J. Pépin, *Mythe et allégorie...*, pp. 157–158.

[58] Let us recall that the original αἰνίττονται suggests that the ancient mythmakers revealed their profound wisdom through riddles and enigmas that were not properly understood by the poets who transmitted the myths.

[59] Cornutus, *Compendium*, 7.20–22.

[60] *Ibidem*, 7.22–8.2.

[61] *Ibidem*, 8.2–6.

that the poets sadly distorted this ancient wisdom, his ethnographic allegoresis of traditional mythology is designed to steer clear of the charges of blasphemy or atheism. As a matter of fact, the philosopher concludes his work with a clear admonition that young men be introduced to "piety" (τὸ εὐσεβεῖν) and not to "superstition" (τὸ δεισιδαιμονεῖν).[62] Thus, Cornutus puts it in no uncertain terms that his etymological and allegorical interpretations of mythology do not entail repudiating myths *in toto*. Nonetheless, the philosopher unequivocally recommends that the valuable and worthless be distinguished in every religion: Cornutus differentiates between authentic piety and crude superstition, as his hermeneutics aims to contrast genuine religiousness with ignorance, obscurantism, and shallow ritualism.

For Cornutus, as for all Greek thinkers, philosophy was not an academic discipline but rather a way of living and a way of dealing with everyday problems. In this regard, interpreting myths had for Cornutus a moral and ethical dimension: his exegeses were supposed to show the educational and didactic value of various ancient myths. In accord with Stoic philosophy, Cornutus assumed that when appropriately interpreted mythology could become an integral part of philosophical *paideia*.[63] When assuming that the particular myths preserve the wisdom of the ancients, Cornutus refused to treat mythology as a set of dark superstitions and ludicrous fables, for he was firmly convinced about the fundamental pedagogical role that myths play in every society.

4. Conclusions

As far as their specific content is concerned, Cornutus' analyses are today of purely *historical* value. However, the very idea that underlies his approach seems to deserve a more favorable assessment. Irrespective of how fantastic and naïve Cornutus' analyses might seem, his aim was to unravel the ancient ways of thinking about the world. When compared to Freud's psychoanalysis or Levi-Strauss' structuralism, Cornutus' analyses become less extravagant and bizarre, for whilst all these approaches aimed to show a certain continuity between the various forms of primordial thinking, the "evidence" was invariably provided by (more or less strained) interpretations of various ancient myths, rituals etc.

For Cornutus, etymological and allegorical interpretations of the gods' names and epithets made it possible to obtain insights into the archaic views that underlay the vocabulary of Homer and Hesiod. Accordingly, interpreting myths was supposed to provide the interpreter with a better understanding of the ancient as well as the present world. Bearing in mind the fact that Cornutus' etymological analyses were often completely fantastic and arbitrary, we should note that he belonged to the most influential philosophical school in the entire Hellenistic period. As for the development of ancient hermeneutics, we should, therefore, stress

[62] *Ibidem*, 76.12–13. The same idea has been expressed by Balbus, who also clearly differentiates between *superstitio* and *religio*, *cf.* Cicero, *De natura deorum*, II 71–72.

[63] I discuss this ethical and existential dimension of Stoic hermeneutics in: M. Domaradzki, 'Theological Etymologizing in the Early Stoa', *Kernos. Revue internationale et pluridisciplinaire de religion grecque antique* 25 (2012), pp. 143–147.

that Cornutus' work shows Stoicism to have contributed significantly to this area through etymological, allegorical and ethnographic interpretations of mythology.

Cornutus' analyses herald a new era in the development of ancient hermeneutics: Neoplatonism will likewise refuse to be content with the letter only. With that, a new ideal of a sage emerges: the one who realizes that knowledge requires a special exegetical effort, for beneath the literal veneer of various myths, images and practices, one can find profound wisdom expressed by the ancients in diverse symbols and enigmas. To reach this wisdom, one needs to have recourse to etymological and/or allegorical interpretation. A person capable of doing this properly deserves to be regarded as a sage.

Metaphysics

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

MAREK PIWOWARCZYK
The John Paul II Catholic University of Lublin

On the Difference Between Inseparability and Dependence

Abstract

This article is devoted to some issues concerning the difference between dependence and inseparability on the grounds of Roman Ingarden's ontology. The main problem is connected with the possibility of object parts which must coexist, but do not make a whole in the same sense as the qualities of objects do. Thus I will argue that the main difference between inseparability and dependence is based upon the difference between two types of whole: the absolute whole and the relative (summative) whole: inseparable ingredients constitute an absolute whole and dependent objects do not constitute a whole at all – or constitute only a relative (summative) whole. The two types of whole are defined in terms of the ties which integrate their ingredients: formal functions and relations. In the first section I present a general sketch of Ingardenian ontology and show the place of the inseparability/dependence distinction in it. The second section is a presentation of the main problem this article. In this section I also introduce the concepts of absolute and relative (summative) whole. The next part exhibits the differences between relations and formal functions and the last contains my solution to the problem and a formulation of possible arguments against this solution.

This article is devoted to some issues concerning the difference between dependence and inseparability on the grounds of Roman Ingarden's ontology. The main problem is connected with the possibility of object parts which must coexist but do not make a whole in the same sense as qualities of object do. Thus, I will argue that the main difference between inseparability and dependence involves the difference between two types of whole: the absolute and relative (summative) whole: inseparable ingredients constitute an absolute whole and dependent objects do not constitute a whole at all, or constitute only a relative (summative) whole. Two types of whole are defined in terms of ties which integrate their ingredients: formal functions and relations.

In the first section I present a general sketch of Ingardenian ontology and show the place of the inseparability/dependence distinction in it. The second section is a presentation of the main problem of the article. In this section I also

introduce concepts of the absolute and relative (summative) whole. The following part exhibits the differences between relations and formal functions and the final contains my solution to the problem and the formulation of a possible charge against this solution.

1. Basic ontological notions[1] and the place of the inseparability/dependence distinction in Ingardenian ontology

In Ingardenian ontology we have two concepts of an object. In the narrow sense, an object is subject of properties. In the broader sense, an object is everything what is "a trinity" of matter, form and mode of existence (I, §9, 68–72/22–27).[2] The last three notions are of the greatest importance in the Ingardenian ontological framework and are basic tools of analysis. Since all of them as primary concepts are indefinable, we can only evoke some intuitions and examples to explicate them.

Matter is a collection of qualities of object – its qualitative make-up (content). The word "qualitative" and "quality" are used in a very broad sense. We can distinguish the following groups of qualities (it is a provisional and incomplete list[3]):

1. Sense qualities: e.g. redness, wetness, smoothness, sweetness, soundness, whiteness, coarseness, coldness etc.

2. Shapes and geometrical qualities: roundness, squareness, triangularity, sphericity, rectangularity etc.

3. Dispositional qualities: fragility, resilience, thermal conduction, electrical conduction, electric resistance etc.

4. Quantities: magnitude, extensiveness, density, length, height, mass etc.

5. Quasi-natures: animality, humanity, being a living entity etc.

6. Constitutive natures: a) haecceities (possible at least in the case of human persons): petreity, socrateity; b) non-unique determinate constitutive natures: caninity, equinity, felinity etc. Quasi-natures are abstract and determinable aspects of constitutive natures.

As we can see the term "quality" is so broad that it refers also to quantities – this is one of reasons why the term "matter" is more suitable (although it is less suitable in other contexts).

Form is a result of the functions matter fulfills with respect to the object and

[1] A very good English introduction to Ingarden's ontology is given in D. von Wachter, *Roman Ingarden's Ontology: Existential Dependence, Substances, Ideas and Other Things Empiricists Do Not Like*, [in:] A. Chrudzimski (ed.), *Existence, Culture and Persons: The ontology of Roman Ingarden*, Frankfurt 2005, p. 55–81.

[2] Ingarden's opus magnum: *Spór o istnienie świata (The Controversy over the Existence of the World)* is cited in the main text. I cite the Polish third edition (Warszawa 1987). "I" stands for "vol. I", "II/1" stands for "vol. II, part 1". After the section number I list the page numbers. In the case of the first volume I also quote the page numbers from the English translation of parts of *Spór o istnienie świata* (R. Ingarden, *Time and Modes of Being*, trans. by H. Michejda, Springfield 1964). German version of *Spór* is: R. Ingarden, *Der Streit um die Existenz der Welt*, Tübingen 1964. I do not cite this edition but the section numbers are the same in both the Polish and German versions. Terminology of *Time and Modes of Being* differs from mine.

[3] Ingarden gives examples of matter in many places of *Controversy*. The list is my reconstruction.

vice versa (II/1, §34, 5–41). The basic form of an individual object is a subject-properties form (II/1, §39, 63–74). Let us consider two sentences:

1. This is a horse.
2. John is brave.

In the second sentence, the word "brave" stands for some quality John "has" (matter which can be distinguished in John): bravery. The word "is" stands for the function bravery plays in John. We can name this function "characterization". The first sentence can be regarded as referring to the relation of some object to its natural kind – but we must remember that for Ingarden an object belongs to a kind because of some immanent matter: just constitutive nature. Thus this sentence is about constitutive nature "equinity" and the word "is" refers to the function equinity plays with respect to the given horse. Ingarden calls it "immediate determination" but I will use the term "constitution". An object is immediately determined or constituted as a subject of properties. Hence it is not bare (non-qualitative) particular but as a subject it has its own matter different from the qualities which characterize it. In turn, the latter are called "properties". In other words, properties are only such qualities of object which play the function of characterization. Property is not a quality itself but quality performing some special role (the role is an abstract aspect of property). The difference between property and constitutive nature (considered as integrated with the function of constitution) is very hard to express. We can only say that constitutive nature makes an object this something (John, this horse etc.) and properties make an object such-and-such (red, big, brave, fragile etc.). In other words: nature gives "whatness" to object and properties give "suchness" to it (II/1, §40–41, 74–91).

We can easily notice it is an Aristotelian approach to the problem of object-qualities structure. According to Ingarden an object is not a bundle of qualities but has subject-properties structure. Yet a subject is not a non-qualitative substratum. Of course this view presupposes that nature cannot be reduced to matter of properties, even essential properties. The distinction between nature and properties is not modal: there are possible essential properties (in Ingardenian terms: absolutely proper properties) which do not belong to nature although are implied by it or "flow" from it. At this point Ingarden is an Aristotelian: peripatetic philosophers also postulated so called *propria*.

All characteristics of individual objects are also individual in the sense that they are characteristics only of a given object. Yet the individuality does not mean (nor imply) qualitative uniqueness. For example two tomatoes can be red in the same way: they have exactly similar rednesses of exactly similar shade, brightness etc. But we still have two rednesses, not one wholly present in two tomatoes.

In the case of properties, Ingarden often says that characterization is a form of property and being a subject is a form of an object (considered in abstraction from properties) and these two forms are inseparable aspects of one form: subject-of-properties. In my opinion these "partial" forms should be rather called "formal functions" and the name "form" better fits the whole complex of formal functions. In a sense, form is a way of organising qualities. Subject-properties is basic form but there some others.

Mode of existence is a complex of existential moments. And existential moment[4] is the way an object (in the broader sense) is existentially conditioned by other objects, or the lack of such conditioning (I, §11, 76–84/32–42). Ingarden distinguishes four pairs of existential moments:[5]

1. Originality/derivativeness (I, §13, 92–116/52–82) – an object is original iff its existence flows from its essence or belongs to its essence. Such an object cannot be produced by any other object. Of course originality is better known as scholastic *aseitas*. An object is derivative iff its existence does not belong to its essence nor flows from it. Such an object must be produced by other object(s).

2. Autonomy/heteronomy (I, §12, 84–91/43–52) – an object is autonomous iff it belongs to its essence that its matter is immanent to it. Ingarden says that autonomous objects have their foundation of being in themselves. An object is heteronomous iff essentially its matter is not immanent to it. Consider such a case: you are reading *Doctor Faustus* by Thomas Mann. One of the figures is Serenus Zeitblom. When you are thinking about him or trying to imagine him, you are ascribing to him some matter, for example humanity, calmness, the sense of responsibility etc. All Zeitblom's matter is projected by Thomas Mann and his readers. He does not have his qualitative content on his own but he needs our conscious acts which project him as an object of their intentional relations. Therefore he does not have his foundation of being in himself but in those acts. Real objects as trees, dogs and persons do not need such conscious acts to exist. Their matter is immanent to them and really "builds" them.

3. Separability/inseparability (I, §14, 116–121/82–89) – an object is inseparable iff it belongs to its essence that it must co-exist with some other object(s) within the larger whole. An object is separable iff it belongs to its essence that it does not have to coexist with any other entity within the larger whole. Consider some particular redness and colourness.[6] It flows from the essence of redness (it means: redness as redness) that it must exist with colourness forming the larger whole, namely red colour. There are several types of inseparability:

3.1. Taking into account the source of inseparability we can distinguish material and formal inseparability. For example, redness is materially inseparable from colourness (redness implies colourness). On the other hand, redness of some particular rose is only formally (as matter of property) inseparable from the rose (being a rose does not imply redness). There are two types of material inseparability:

3.1.1. Qualitative inseparability[7] (II/1, §40, 79–80, also notes 33–34) – this is inseparability which obtains between the determinable and its determinate, or

[4] Detailed and creative analysis of existential moments is presented in: M. Rosiak, *Spór o substancjalizm. Studia z ontologii Ingardena i metafizyki Whiteheada (The Controversy over Substantialism. Studies in Ingarden's Ontology and Whitehead's Metaphysics)*, Łódź 2003, p. 23–36. See also M. Rosiak, 'Existential Analysis in Roman Ingarden's Ontology', *Forum Philosophicum* 12 [1] (2007), p. 119–130.

[5] My presentation of existential moments draws from my article *Ontological Priority of Substances over Objects of Other Categories*, [in:] M. Szatkowski (ed.), *Dualistic Ontology of the Human Person*, München 2013, p. 203–214.

[6] In this article I accept that determinables are at least possible, but I know how many problems are connected with this view.

[7] The name is quite misleading insofar as it suggests that the second type of material inseparability has not its source in matter.

rather between determinable and determinans, so to speak. For example in human beings, animality is qualitatively inseparable from some determining aspect – let us say: rationality. Animality and rationality make up the larger whole: humanity which is a more determinate unit. A complete determinate is qualitatively separable (although it can be inseparable in other ways – for example formally). Notice that the determinable – determinans relation, as a relationship between qualities, is also some kind of formal function although it is not an objective form – it is not a form able to make the complete object (subject of properties). For by no means a determinable is a subject of determinans nor is determinans a property of a determinable. Let us call this kind of formal function "determination". Determination is possible both between material aspects of constitutive natures and between material aspects of properties. Only fully determinate (i.e. qualitatively separable) qualities can play the functions of either characterization or constitution.

3.1.2. Material inseparability in the broad sense[8] – not based on determination. For example, the ability of photosynthesis is inseparable from being green but being green does not determine the ability of photosynthesis(nor vice versa) although the latter needs the former. All essential properties are materially inseparable (in the broader sense) from constitutive nature although their matter is not an aspect of nature.

3.2. Taking into account the range of entities which an object can be inseparable from, we have:

— Rigid inseparability – an object must co–exist in a single whole with one specific object (with this–and–this specific object);

— Generic inseparability – an object must co–exist with one of the objects from the specific class 'M'.

3.3. Rigid inseparability can be mutual or unilateral.

4. Dependence and independence (I, §15, 121–123/89–92) – an object is dependent iff it is separable and essentially needs some other object to exist. An object is independent iff it is separable and does not need any object to exist. There are the same varieties of dependence as in the case of inseparability with the exception of 3.1.1. For example, Socrates is generically dependent on particles of oxygen. According to theistic conception of *creatio continua*[9] every being is rigidly dependent on God.

Every existential moment excludes some others (for example heteronomy excludes autonomy, originality and independence) and implies some others (e.g. inseparability implies derivativeness). A mode of existence is a complex of existential moments coherent (at least not excluding) with each other. Ingarden also analyses existential moments connected with existence in time (and non-existence in time) but they are not relevant to the topic of this article (I, §28–30, 189–232/102–156).[10]

[8] Ingarden does not use a distinct name for this second type.

[9] *Creatio continua* is conceived here as supporting existence and not as a continual production of new beings.

[10] Detailed considerations concerning temporal modes of existence one can find in M. Rosiak, *Spór o substancjalizm*, p. 48–58, and in F. Kobiela, *Filozofia czasu Romana Ingardena (Roman*

In a trinity of matter, form and mode of existence, the matter is ontologically the strongest in the sense that form and mode of existence are determined by matter. For example, roundness of a ball cannot be its constitutive nature but plays the formal function of characterization, or in other words: is a property. Playing this formal role, the roundness is rigidly inseparable from the subject which it characterizes. Thus it is also a derivative. Ingarden does not explicitly use the concept of identity-dependence but I think it can be applied in his ontology just in the case of the ontological status of properties.

2. Inseparability and dependence – problems with a specific borderline case

Inseparability is considered here as existential conditioning between some items within a larger whole. But sometimes we use to say that this larger whole is inseparable from its ingredients, for example a rose is inseparable from its necessary properties e.g. ability of photosynthesis. Thus, I think[11] we must distinguish inseparability between the ingredients of a whole (say I-I inseparability), inseparability between a whole and the ingredients it contains (say W-I inseparability) and analogical inseparability between an ingredient and a whole (I-W inseparability). In the second and the third case we have different meaning of inseparability than defined above. That this rose (considered as complete object, not as an abstract subject of properties) is inseparable from its ability of photosynthesis means that the rose essentially must be some larger whole containing the ability in question as one of its inner elements. Ingarden also recognizes the need of such clarification.[12] In this paper I am interested only in I-I inseparability but of course we must remember that the concepts of W-I and I-W inseparability are based on the fundamental I-I concept.

Dependence *ex definitione* excludes inseparability. It means that dependent objects must co-exist with other objects but they do not make up a whole with them. I think that the Ingardenian distinction between inseparability/separability and dependence/independence is very important. For example, it enables us to define the ontological status of properties more precisely. We intuitively feel that there is a great difference between the way in which God conditions the creatures and the way an object conditions its properties. A concept of dependence as it is used in analytic philosophy[13] is not enough even if we distinguish generic and rigid, existential and identity, dependence etc. To say that creatures are rigidly, existentially and identity-dependent[14] on God and that properties are in the same way dependent on their bearer, is to neglect very important difference between

Ingarden's Philosophy of Time), Kraków 2011, p. 151–199, 251–275.

[11] See my *Ontological Priority of Substances*, p. 208–209.

[12] See R. Ingarden, *O sądzie warunkowym (On Conditional Proposition)*, [in:] R. Ingarden, *Z teorii języka i filozoficznych podstaw logiki (On Theory of Language and Philosophical Fundaments of Logic),* Warszawa 1972, p. 297–299.

[13] The most extensive analytical study on dependence is: F. Correia, *Existential Dependence and Cognate Notions*, München 2005.

[14] In the case of *creatio ex nihilo* God is the only source of creatures, so He determines them completely, also with respect to their identities.

these two cases.[15] On the ground of Ingardenian ontology we have a tool to express the difference.

But I also think the Ingardenian distinction is only the beginning and can be easily misinterpreted.[16] Let us consider an object which is composed of some other objects. Artifacts are composed of mutually independent objects. But in the case of organisms some parts on higher levels of composition seem to have to coexist with each other. And even if we are mistaken with organisms known from common sense, it is possible to conceive such an organic whole which is composed of objects (in the narrow sense), the existence of which is conditioned by other parts of the whole: mutually and rigidly or unilaterally and generically. Let us call such a whole, a "strong organic whole". We do not need to search for such a whole in reality – its possibility is enough because ontology should provide a framework for ontologically possible entities.

The main problem is: are parts of a strong organic whole inseparable to each other or rather dependent?[17] At the first glance we should choose the first option: for parts of a strong organic whole must coexist with the other parts within the larger whole. If it is true, we still do not have a tool to distinguish the ontological status of properties from the ontological status of complete objects (subjects-of-properties) which are parts of strong organic wholes. But we intuitively feel that properties are in a different sense distinct from parts of a strong organic whole. The latter seem to have a stronger ontological status. Or, in other words: we have very a strong intuitive feeling that the "wholeness" of a strong organic whole is different than the "wholeness" of subject-of-properties and that the difference does not consist in the essential coexistence of parts or lack of it. We can conceive ontological atoms – in Ingardenian terms: primary individual objects. Such atoms are not composed of other objects in the narrow sense (subjects-of-properties) but still, we can distinguish different qualities in them although we resist to say that those qualities compose an object in the same way as organs compose a strong organic whole. On the other hand, in the case of secondary individual objects (composed objects – a strong organic whole is a specific case of such objects) the whole itself has some qualities which supervene on qualities and relations of

[15] Of course God is by no means dependent on creatures and this is the difference between the two aforementioned cases. But it makes no difference with respect to the distinction between the status of creatures and the status of properties. We can even conceive that God is dependent generically or even rigidly on His creatures – of course He would not be a God of Abrahamic faiths but such a situation is ontologically possible and in that case there is no difference between God-creatures conditioning and object-properties conditioning.

[16] In my opinion Peter Simons's approach to the distinction is based on misinterpretation. See P. Simons, *Ingarden and the Ontology of Dependence*, [in:] *Existence, Culture and Persons*, p. 39–53.

[17] I know (from private communication) that the problem was also recognised by Katarzyna Barska in her Ph.D. thesis *Momenty bytowe a formalna i materialna struktura przedmiotu w ontologii Romana Ingardena (Existential Moments and Formal and Material Structure of Object in Roman Ingarden's Ontology)*, Kraków 2012, manuscript. Unfortunately I do not know the results of the thesis – it is unpublished. Barska analyzed the concept of inseparability in: 'Niesamodzielność materialna i formalna jako naczelne typy niesamodzielności bytowej' ('Material and Formal Inseparability as the Main Types of Existential Inseparability'), *Principia* 47–48 (2007), p. 281–293.

parts, but these supervenient qualities do not constitute the next (higher) level of composition (II/1, §43, 102–137).[18] If we have only one notion of composition and one notion of wholeness we must admit that the qualities of primary individual objects are in fact primary individual.

These considerations suggest that we rather should say that properties are inseparable and parts of a strong organic whole are only dependent. Thus the crucial point of definition of inseparability should not be that inseparable items have to coexist in the larger whole, but that they have to coexist in the larger whole in very special meaning of "whole".

Ingarden introduces a distinction which can be relevant to this problem. He speaks about absolute whole and relative whole (II/1, §39, 67–69; §43, 104–105). An absolute whole is a whole in the sense of material (in Ingardenian sense), formal and existential completeness. Every real individual object is completely defined in its qualitative content, form and mode of existence. Matter is organized by some form, especially by the subject-properties form, and an object also has a defined mode of existence. It is wholeness in the sense of definiteness. Such completeness makes an object a distinct item of reality. It is an allusion to the Wolffian use of the word "determination" – cutting off (*de-terminatio*). An object as an absolute whole is closed in itself. Such a whole is absolute in the sense that it is not a whole with regard to some other objects in narrow sense, but rather with regard to its ontological ingredients like qualities joined by formal functions etc.

A relative whole is also called a summative whole – this is a whole with regard to some other objects in the narrow sense – this is a whole of these-and-these objects. A composed object can be grasped both as an absolute whole and as summative whole. It is an absolute whole with regard to its qualities (supervenient on qualities and relations of parts) and it is a summative whole with regard to its parts. Of course we can easily notice that the main non-controversial contribution of this distinction are only those names: absolute/relative whole. This distinction is based on the intuitions evoked above – but we need a stronger basis.

If we want to distinguish inseparability and dependence then the concept of an absolute whole must be entangled into the definition of the former. But only entangled because not every two items make up an absolute whole. There are degrees of completeness. Notice that absolute wholeness is the same as full concreteness. The larger whole in the definition of inseparability means a more concrete unit. Consider the colour red. We can abstract from it redness and colourness. Thus, there is a hierarchy of inseparability. Redness and colourness are at the bottom of the hierarchy. They are the highest *abstracta*. Red colour is on the second level but it is inseparable from the function of characterization, forming the larger

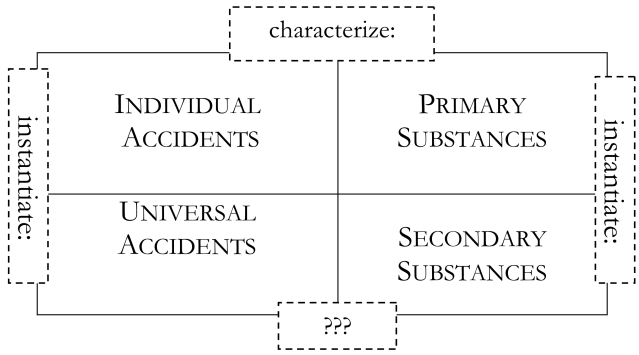
[18] Clarification and development of Ingardenian part-whole theory is contained in M. Rosiak, *Spór o substancjalizm*, p. 87–98. See also: M. Rosiak, *Ingarden Roman*, [in:] H. Burkhardt, J. Seibt, G. Imaguire (eds), *Handbook of Mereology*, München (forthcoming); M. Rosiak, 'Własności relacyjne, całości i przedmioty wyższego rzędu' ('Relational Properties, Wholes and Objects of Higher Level'), *Principia* 30 (2001), p. 117–133; K. Barska, 'Formalna i egzystencjalna analiza całości sumatywnej w ontologii Romana Ingardena' ('Formal and Existential Analysis of Summative Whole in Roman Ingarden's Ontology'), *Kwartalnik Filozoficzny* 35 [2] (2007), p. 109–121.

whole, i.e. the property "being red". And this property is inseparable from many other properties of a particular rose. The rose is the full *concretum* and is simply separable.

The parts of a strong organic whole are concrete entities. Although they must coexist with each other they do not supplement each other in the way qualities of concrete object do.

But do we have any non-circular criterion of distinction between an absolute and summative whole? We are tempted to define an absolute whole in terms of inseparability, but it automatically leads to circularity. All those words "*concretum*", "concrete", "absolute whole", "more concrete unit" cannot be defined in terms of inseparability, although we know it would be the easiest way of defining them. On the other hand, the "absoluteness" of an absolute whole so strongly contrasted with the "relativeness" of a relative whole turns out to be only another type of relativeness. An absolute whole is also a whole in relation to some parts – but parts of a very special type: the material, formal and existential ingredients of an object. Completely non-relative wholeness would be wholeness of an absolutely simple being – like, for example, God – conceived in the scholastic manner.[19] The distinction between two types of whole as based only on the distinction of two types of parts – qualities on one hand, and "normal" objects on the second, leads to circularity. Thus we are still at the starting point and do not know how to distinguish these types of whole without using inseparability/dependence terms.

It is not only the problem of Ingarden's ontology. It is well known that Aristotle, in *Categories* (*Cat.* 1a 20–1b 10), says about four kinds of entities: primary substances, secondary substances, individual accidents (non-substantial individuals – substance/accident distinction is not modal here), universal accidents. Relations[20] among these items form a so-called ontological square. The square was interpreted in many different ways. Here is my interpretation, similar in important respects to E.J. Lowe's four-category ontology:[21]



[19] Ingarden maintained God is not absolutely simple: He has a subject-properties structure.

[20] I will use this word because I think they are not relations. See the next section.

[21] See E.J. Lowe, *The Four-Category Ontology. A Metaphysical Foundation for Natural Science*, Oxford 2006, p. 40.

The characterization relation corresponds to Aristotle's "not-being predicated of a subject and being-in a subject". I interpret the relationship of being predicated of as a converse of instantiation – thus I insist universal accidents are predicated of individual accidents. I cannot agree with Lowe that universal accidents characterize secondary substances. We do not have space for discussion about my interpretation. What is relevant to the main topic of this article is that Aristotle's definition of being in a subject: "by 'in a subject', I mean what is in something not as a part and that it cannot exist separately from what is in it" (*Cat.* 1a 24–25). It seems Aristotle gives two conditions of being in: a) being in not as a part, b) the impossibility of being separated. Of course we cannot at the moment identify b) with Ingardenian inseparability.

The first condition is quite mysterious. Many scholars think that being in, not as a part, means being in not as a *physical* part – and being a physical part is often understood as being separable.[22] Of course on the ground of such interpretation, the second condition is already contained in the first. If we want to treat a) and b) as two distinct conditions we must admit that there are possible parts which cannot exist separately.

What items do we get when we cross two distinctions: being in as a part / being in not as a part and the possibility of separate existence/impossibility of separate existence? Obviously we get four types:

1. parts which can exist separately;
2. non-parts which can exist separately;
3. parts which cannot exist separately;
4. non-parts which cannot exist separately.

The first case is uncontroversial. The third seems to correspond to parts of our strong organic wholes and the fourth to properties or even to aspects of properties or other "non-part-like ingredients". We can try to complicate the division by interpreting "can exist separately" in terms of Ingardenian independence or separability. But what would the difference be between inseparable (now, in Ingardenian sense) parts and inseparable non-parts? Would there be possible dependent non-parts? Notice that in the Ingardenian framework, non-parts would mean "ingredients of an absolute whole". Thus, this qualification "not as a part" implies that the phrase "cannot exist separately" in 4. must mean Ingardenian inseparability. On the other hand, the possibility that *parts* of some kind cannot exist separately should be expressed in terms of Ingardenian dependence. On the ground of this ontology, 2. would be impossible in any way: non-parts cannot be separable nor independent.

The above considerations suggest once again that the modal distinction between parts and non-parts is not enough – there are possible necessarily coexistent parts of object parallel to its necessarily coexistent non-parts which however make up different types of a whole.

The trip to Aristotle teaches us that there were attempts to treat "metaphysical ingredients" as non-parts. Ingarden is also inclined to such an approach, but

[22] See R. Heineman, 'Non-substantial Individuals in the *Categories'*, *Phronesis* 26 (1981), p. 295–307.

sometimes he admits we can analyse an absolute whole in terms of "normal" part-whole relations (II/1, §35, 33–36), yet he immediately explains that this is a special variety of part-whole relationship and that in this case we impose on the object (as an absolute whole) the formal structure which does not express the real specificity of its objective form. For in analysis in terms of the part-whole relation, specific formal functions (characterization, constitution and determination) are lost. Formal functions *integrate* qualities and this is the reason why they seem to play an analogical role as a complex of relations which obtain among parts of some summative whole. Notice that part-whole relation is (trivially) possible only if some object is a part of another object and that parthood is not a relation on its own, but rather results from some relations in which the given object stands to other objects – relations sufficient to make a whole. In Ingardenian ontology such set of relations is called "form III" and objects standing in form III are called "matter III". Form analysed in section 1. is named "form I" and its correlate: "matter I". A part-whole relation is possible only because some objects are formed by form III. Because qualities of object are formed by form I they also seem to be very specific parts of an object and we think we can analyse their relationship with the object in terms of a part-whole relation. We cannot avoid thinking about them as special parts – even if we name them "metaphysical parts", "logical parts", "non-physical ingredients" etc. This is the cost of the possibility of analysis, which always involves some kind of distinction between some items – in contrast with other items. The only thing we can do is to show how specific parts they are and how specific wholes they make up, and that this part-whole structure is not fundamental but derivative.

A summative whole is a result of form III and an absolute whole is a result of form I. Thus if we want to find the difference between these types of whole, we must find the differences between form III and form I, or in other words, between relations and formal functions. The difference between inseparability and dependence can be expressed only in terms of the difference between formal functions and relations.

3. Relations and formal functions (formal relationships)

Ingardenian theory of relations[23] can be sketched as following (II/1, §55, 291–303):

In a relational state of affairs we must distinguish:

1. Relation bearers – objects among which a relation obtains. For example, in a relational state of affairs "Joseph is similar to his dog (Piwko[24])" relation bearers are Joseph and Piwko.

2. Foundations of relation (*fundamenta relationis*) – these aspects of relation bearers (their properties, constitutive natures or quasi-natures) due to which a relation obtains. Let us assume both Joseph and Piwko like beef. Thus the foun-

[23] Clear introduction to Ingardens theory of relations is given in: P. K. Szałek, 'Romana Ingardena ontologiczna teoria relacji' ('Roman Ingarden's Ontological Theory of Relations'), *Kwartalnik Filozoficzny* 34 [1] (2006), p. 31–60.

[24] "Piwko" was the name of Josef M. Bocheński's dachshund.

dations of similarity between them are Joseph's and Piwko's special food-habits.

3. Relational bond or a relation in the strict sense – this is what exists between objects participating in a relational state of affairs. In fact we do not have sufficient vocabulary to talk about this item because we almost always use names of so called relational features (see the next point) and they are more like monadic properties. In our example, a relational bond is just a similarity between Joseph and Piwko.

4. Relational features – features which characterize relation bearers because of their standing in the given relation. In our example they are accordingly: Joseph's being similar to Piwko and Piwko's being similar to Joseph. Now we can see how easily we can confuse these three things: relational feature, relational state of affairs, and relational bond. Let us consider a state of affairs: 3>2. We have a relational bond: > and relational features: 3's being greater than 2 and: 2's being smaller than 3. According to Ingarden, there is only one relations in the strict sense here and not, as many philosophers think, two relations: "being greater than" and its converse "being smaller than". There is only one relational bond and because this bond implies some order of bearers, they have different relational features. Notice also that in the case of Joseph and Piwko they also have two different relational features – although named with the same words: "being similar to...". Probably Ingarden would say that the broadly discussed problem of so called neutral relations[25] is caused by the fatal confusion of a relational bond with relational features or the confusion of relational states of affairs with relational features.

On the grounds of this theory, all relations are internal in Russell's sense but only insofar as they must be ontologically grounded in something non-relational. Of course there are possible such relations which have other relations as their foundations, but ultimately all relations are founded in non-relational qualities. Yet Ingarden's view is not reductionist – relational bonds are genuine elements of being although very weak and materially (in the sense of matter I), formally and existentially conditioned. Ingarden has very strong doubts whether some types of relational features are autonomous (non-intentional – real) but he never reduced relational bonds to non-relational qualities.

What must be emphasized here is that relational bonds are objects in the broad sense: they have matter, form and a way of existence. Relations (from now I will use the term "relation" instead "relational bond" or "relation in the strict sense") are distinguishable in terms of their matter (II/1, §55, 298–299). Of course it is very difficult to speak of matter of relations because it is very easy to confuse it with the whole relation (as "composition" of matter, form and way of existence). In the case of similarity relation, its matter is just similarity, its form is being a relation. We do not have a technical term to signify form of relation as we have the characterization in the case of properties. Redness plays the role of characterization with respect to some tomato and similarity also plays some formal function with respect to Joseph and Piwko – function different than characterization. And because being red belongs to the category of property due

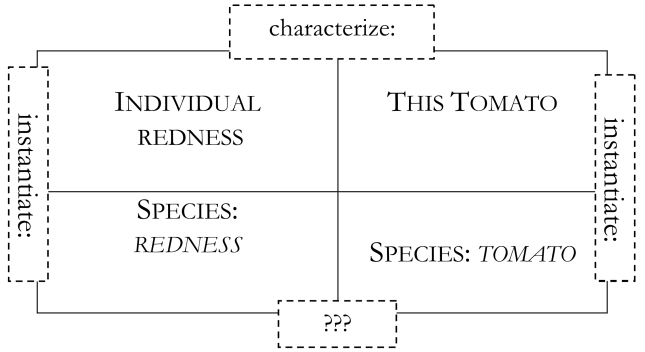
[25] See K. Fine, 'Neutral Relations', *The Philosophical Review* 109 [1] (2000), p. 1–33. Fine notices Ingarden's contribution to the ontology of relations – see p. 1, note.

to the function of characterization, then similarity belongs to category of relation due to its special form.

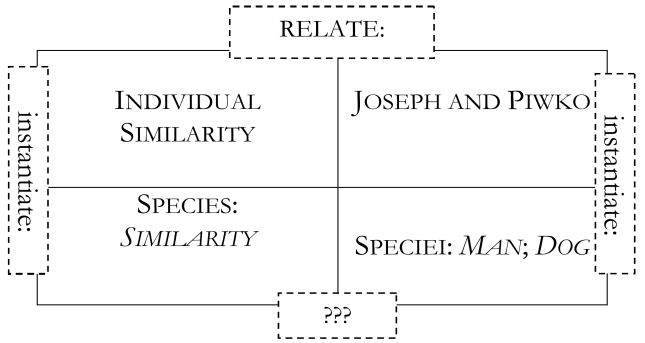
Moreover, that matter of relation is not the same as matter of relational feature (matter of similarity relation is not the same as matter of relational feature) although we sometimes use the same name to signify it. Matter of relational feature plays the function of characterization and does it only with respect to one object – whereas matter of relation plays its formal function with respect to at least two objects. Let us call the formal function of relation "RELATE".

The Ingardenian view of relations is highly complicated. But its main thesis that relations have matter and form can be grasped also in another way. Consider once again an ontological square. Universal accidents are of course all non-substantial universals. Let us compare two cases:

A.



B.



Redness plays the function of characterization with respect to its subject, and similarity plays the function RELATE with respect to its relation bearers.[26] The

[26] "This tomato" in A. is of course not the full object but an object considered as a subject, i.e. in abstraction from its properties. *Mutatis mutandis* we can say the same about Joseph and Piwko in B.

full relation is not only individual similarity from the upper left box in B. The whole relation is individual similarity as fulfilling the function RELATE.

It is time to ask a very important question: are formal functions relations? They seem to be but they are not (II/1, §56, 309–311). Formal functions do not have matter[27] but are ways in which matter is organized. It is evident in the examples A. and B. RELATE is not a further relation but is rather a much more fundamental tie by which relational matter (individual similarity) is attached to relation bearers (Joseph and Piwko). The same, *mutatis mutandis*, must be said about characterization. Formal functions are not instances of relational universals but are ultimate ties between instances of non-substantial (so also relational) universals and instances of substantial universals. Moreover, if formal functions were relations, we would be involved in infinite regress: their matter would have to be attached to the related items by further relation, and so on, ad infinitum.

Similar reasons lead to the thesis that instantiation also is not a relation. But Ingarden would say it is not a formal function in the strict sense. Ingardenian universals are transcendent ideas. Thus instantiation is not a tie which organizes matter of individual objects.

Strictly speaking, formal functions are not entities. I know it is a very controversial claim, but I do not mean that formal functions are simply nothing, or that they are completely out of being. They are not entities in a similar way that existence is not an entity. In a sense they are more fundamental than entities because due to them, an entity can be an entity.[28]

To sum up: we have relations and non-relational ties. Within the latter subdivision we should distinguish formal functions. The main difference between relations and non-relational ties is that the latter do not have matter. Thus they are not objects even in a broad sense.

4. A possible solution and a new problem

Given the distinction between relations and non-relational ties, an absolute whole can be defined as a whole formed by specific non-relational ties – formal functions – and a summative whole, as a whole integrated by relations. We must remember that there are also possible non-relational ties which cannot make up a whole of any type (for example instantiation) and relations unable to make up a summative whole (similarity is such a relation).

Hence we cannot define inseparability and dependence in purely existential terms. These concepts turn out to be existential-formal concepts. A new definition of inseparability/separability can be stated as following:

[27] In this sense they can be called "thin" relations. About thin and thick relations see K. Mulligan, 'Relations: Through Thick and Thin', *Erkenntnis* 48 [2–3] (1998), p. 325–353. Mulligan has some troubles with the criterion of distinguishing thick and thin relations. As we see Ingarden's concepts of matter can be helpful here.

[28] This approach to formal function is similar in many respects to E.J. Lowe's. See E.J. Lowe, *The Four Category Ontology*, p. 44–49; E.J., 'Some Formal Ontological Relations', *Dialectica* 58 [1] (2004), p. 297–316.

> INSep: An object is inseparable iff it belongs to its essence that it must co-exist with some other object(s) within the larger whole integrated by some formal functions.
>
> Sep: An object is separable iff it belongs to its essence that it does not have to coexist with any other entity within the larger whole integrated by some formal functions.

Literally, definitions of dependence and independence are not changed but of course the term "separable", as used here, has a more precise meaning:

> Dep: An object is dependent iff it is separable and essentially needs some other object to exist.
>
> INDep: An object is independent iff it is separable and does not need any object to exist.

Given these definitions dependent objects can be of two types:

1. Those which do not make up any whole.
2. Those which make up the whole but are integrated by relations.

At first glance it seems that the third type is also possible : those objects which make up the whole but which are integrated not by relations nor by formal functions but by non-relational ties of some other kind. But in that case, such a whole would be indistinguishable from an absolute whole and objects would not be dependent but inseparable – and those non-relational ties would turn out to be formal functions.

What are necessary conditions of 2.? There is only one condition: relations must be essential in that case. Parts of a strong organic whole are dependent (and not inseparable) only if they stand in such whole-making relations which are essential to them. What does this mean? Ingarden does not talk about essential relations but he proposed a very interesting approach to essential properties and we can apply this analysis to relations.

Ingarden maintains that the intrinsic/extrinsic distinction is too rough and replaced it with a four-fold division of properties (II/1, §57, 314–333):

1. Absolutely proper properties – their matter and existence is conditioned only by the object itself.
2. Attained properties – the beginning of their existence is caused by an external factor but their further existence is conditioned only by the object itself.
3. Externally conditioned properties – their beginning and further existence is conditioned by external factors.
4. Relational features.

For Ingarden, essential properties are some absolutely proper properties – but only those which immediately flow from the constitutive nature and, as Ingarden says, are equivalent to it (II/1, §58, 345–353). There are also possible absolutely proper properties which are implied by essential properties. We can say that they are essential in a broad sense. From now I will use the term "essential properties" as the equivalent of "absolutely proper properties". Of course the constitutive

nature and its determinable aspects are also essential characteristics of a thing, but they are not properties.

Essential relations would be relations which are either (1) contained in constitutive nature or (2) implied by constitutive nature or (3) implied by essential properties. I think that the first case should be excluded because of two reasons. Firstly I cannot accept that individual objects have relative identities (generic or individual). Secondly, ingredients of constitutive nature cannot be objects even in a broad sense. Yet I will not develop these problems.

The second and the third case are acceptable but we must refine some issues. Firstly, notice that "absolutely proper relation" is a contradiction in terms insofar as we conceive relations as entities conditioned at least by two objects. Thus essential relations are "absolutely proper" only in the sense that they cannot be attained and lost. Moreover, the thesis that essential relations are implied only by constitutive nature or essential properties does not only mean that nature and those properties are in that case foundations of relations. For there are possible non-essential relations which fundament is constitutive nature or essential properties. Every two objects of the same natural kind are similar to each other also because of their constitutive natures and absolutely proper properties, but of course such similarity is not necessary to them: for example, all dogs with one exception can be terminated and the survivor will not be similar to other dogs – although this similarity was based on its constitutive nature. Essential relations are implied by the nature or essential properties in the sense that nature or essential properties are mutually materially inseparable (in the broad sense) from the relations in question.

But there is a very serious problem of this new approach. If RELATE is a formal function, then relations are inseparable from their bearers. This means that matter of relations must coexist in the absolute whole with their bearers so (given transitiveness of inseparability) that the bearers are also inseparable from each other.[29] Then, parts of strong organic wholes are inseparable and we cannot express our intuitions (mentioned in section 2) in ontological terms.

I do not want to analyse this problem, but I will mention some possible solutions:

1. We can refute the transitiveness of inseparability.
2. We can accept that parts of a strong organic whole are in fact inseparable.
3. We can accept the view that RELATE is a non-relational tie, but refute the opinion it is a formal function.

I am inclined to 3. although I am aware of problems connected with this solution. As I have mentioned above, non-relational ties which are not formal functions could not be whole-making ties. But notice that such non-relational ties are ties between objects of different categories. For example, instantiation is a tie between individual objects and universals (Ingarden's ideas) but there are also non-relational ties between objects and processes, objects and events, and so on. In these cases, entities of different categories tied in such non-relational ways do

[29] Of course it is not true in the case of non-essential relations because object does not have to coexist with them.

not make an absolute whole. But entities of some categories can be means by which entities of other categories are integrated into a summative whole. Relations are such means by which individual objects can be integrated. Processes can also be means of integration: objects can be unified by process in which they participate. In such cases, integrated objects are dependent (if relations and processes are essential to them) but not inseparable. Of course the cost of such a solution is that integrating relations or processes can only be dependent on their bearers or participants.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

WOJCIECH MAŁECKI
University of Wrocław
DAVID WALL
University of Northampton

Narrative Necessity

Abstract

In the paper we argue that narrative structure is a necessary condition for meaningfulness. That is, it is necessary for a life to have meaning that there is a sufficient degree of coherence between the different actions, events, and attitudes in an agent's life, and that these can be understood in terms of a narrative. Furthermore, we suggest why having such narrative structure is important for meaningfulness: where someone's life is meaningful this seems to be in virtue of her being appropriately related to some meaning-providing feature. Finally, we try to show that what we have argued sheds light on why various suggestions that have sometimes been made by philosophers about what having a meaningful life consists in often strike us as implausible or overly abstract.

1. Introduction

Suppose that Albert is suffering from existential angst. He thinks his life lacks meaning and he finds this distressing. This distress is increased when he considers his friends and acquaintances: they do not share his feelings of meaninglessness and he thinks his life is empty in comparison with theirs. But what is missing from Albert's life (or that he thinks is missing) that is present in those of his friends (or is thought to be by them and by Albert)? What is it that provides someone's life with meaning?

Some people take this to be the crucial problem that philosophers are, or should be, concerned with. Yet, when philosophers, either professional or lounge-room, do address this problem their suggestions often strike us as nonsensical, false, or useless. Proposals such as the pursuit of eudaimonia (e.g., Aristotle) or of the greatest pleasure (hedonism), the maximizing of the happiness of the greatest number of fellow-humans (utilitarianism) or the fulfilling of a God-given purpose (various kinds of supernatural philosophies), seem overly abstract and difficult to grasp; others such as contributing to the glory of one's nation, or effecting a particular social change, or having a particular career, seem too specific indeed

and thus vulnerable to counter-examples. We will not attempt to make any similar suggestion here. Rather, we argue that it is necessary for a life to having meaning that it has a certain formal property, that is, that it has narrative structure. Furthermore, we hope to explain why this is necessary, and hope that this will be instructive as to why these other, particular suggestions might seem prima facie unconvincing.

But note some other things that we will not try to do here. We will not argue that narrative structure is a necessary property of human experience. Galen Strawson has recently argued that it is not and claims that he experiences his life episodically and without narrative.[1] Perhaps this is true of Strawson, and it probably makes our claim more interesting if it is possible for one's experience to be episodic and non-narrative. Second, we will not try to say with any exactness what kind of property meaningfulness is in this context. Intuitively this is different from linguistic, or sentential meaning, aesthetic meaning, or intentionality. It also seems different from merely making sense, or being intelligible.[2] We assume that most people have some pre-theoretic grasp of the notion, and it is being meaningful in that pre-theoretic sense that we argue narrative structure is necessary for. Finally, we make no claim about what is involved in a life having value.[3] Perhaps there is some relation between meaningfulness and value but they are not the same and the first does not seem necessary for the second: we would balk at the idea say, that it was permissible to kill Albert because his life has no meaning. However, we will not attempt to investigate how these notions are related.

2. Examples of Meaningful Lives

When do we say that someone has a meaningful life? Consider some examples: perhaps the paradigm case where we think that someone is living a meaningful life is the spiritual or religious person. This seems to be related to the agent's dedication to, or involvement with, something that is described as "greater than themselves." Similarly, someone who has dedicated her life to doing good work is often said to have a meaningful life. This might include people who, say, do voluntary work overseas or within their community, or who work to make some contribution that improves the quality of life for others. But it does not seem necessary that one is related to some greater purpose such as a religious or ethical project for one to have a meaningful life. It seems possible for someone to have meaning through something more personal or circumscribed. For instance, we

[1] G. Strawson, 'Against Narrativity', *Ratio* 4 (2004), pp. 428-452; *cf.* G. Strawson, *Episodic Ethics*, [in:] G. Strawson, *Real Materialism and Other Essays*, Oxford–New York 2008; P. Lamarque, 'On Not Expecting Too Much From Narrative', *Mind & Language* 19 (2004), pp. 393–407; A. Rudd, 'In Defence of Narrative', *European Journal of Philosophy* 17.1 (2009), pp. 60–75.

[2] So what we are arguing for is different from what McIntyre argues for in *After Virtue* (esp. ch. 15). There he argues that someone's life must have narrative structure for it to be intelligible, for determining what particular actions she is performing, and for it to be properly unified as a single life of an individual. While having meaning might be related to these other things it is intuitively distinct from, and additional to, them.

[3] Of course, being meaningful is one respect in which a life can be valuable. But we are concerned to put aside discussion of the relation between meaningfulness and moral value.

sometimes think that someone's life has meaning because she has children, or perhaps a career that she enjoys or that is important to her.

Of course, in all such cases a radical sceptic, who denies that it is possible for someone to have this kind of meaning, might question whether these individuals do actually have meaningful lives. But in ordinary situations, where we accept that there is such a thing as a meaningful life, such radical sceptical worries do not apply: these seem to be the kinds of individuals for whom concerns about meaninglessness would not arise. To be sure, there are philosophers who think it is only they who know what the real meaning of life is, and that the rest of humanity needs to be enlightened by them, but we ourselves do not have any such ambitions.

These paradigm examples seem to have a similar structure. In all cases the agent seems to stand in some relation, R, to some feature, F, and it is in virtue of standing in R to F that her life is meaningful. To illustrate, it is in virtue of bearing the *dedicated to* relation to her religion that the religious person's life is meaningful; it is in virtue of having the *performing* relation to her good deeds that the ethical person's life is meaningful; it is in virtue of standing in the *parenting* relation to her children that the parent's life is meaningful; and it is in virtue of having an *occupying* or *pursuing* relation to her particular career that the career-person's life is meaningful. We can acknowledge that there may be some cross-cultural variation, and even variation between individuals within a particular culture regarding the features that could make someone's life meaningful. For example, in traditional Tibetan culture dedicating oneself to religion and becoming a monk is seen as an especially meaningful way of life, whereas in 1980's USA dedicating oneself to one's career (especially a career of a certain kind) was seen as especially meaningful. And there may be some pursuits that cannot provide meaning in any culture: it is problematic to think that someone might have a meaningful life in virtue of collecting brown jelly beans, for instance, although it seems that even this possibility cannot be ruled out altogether (it is neither logically inconsistent, nor refuted by any empirical data). Nonetheless it seems to be the structure of standing in an appropriate relation, R, to some relevant feature, F, that is what provides someone's life with meaning.

Can someone be mistaken about whether or not her own life is meaningful? Recall, in the original example, although Albert thought that his own life was meaningless he believed that the lives of (at least some of) his friends were meaningful; and they themselves believed this about their own lives. But consider Albert* who in addition to believing that his own life was meaningless, similarly to Albert, also believed that no human life could have meaning. So he judged that Albert's friends were mistaken about the meaningfulness of their own lives. Can Albert* be right about this: Is it possible for someone to make a mistaken judgment that her life has meaning? Or is having a meaningful life similar to having a particular phenomenal experience, such as an experience of red: if it seems to you that you are having a red experience then you are (almost certainly) right, and if it seems to you that your life is meaningful then it (almost certainly) is? Conversely, can someone be mistaken that her own life is meaning*less*: might Albert in fact

have a meaningful life, perhaps in virtue of being appropriately related to some meaning-giving feature, even though he believes that (at least in his case) being related in this way to that feature is not sufficient for having a meaningful life? This is an interesting, and difficult question. It raises a number of issues about the distinction between first-person and third-person assessments of someone's life, about the extent to which meaningfulness depends on psychological satisfaction, and about the kind of phenomenon meaningfulness is and how we have epistemic access to it. However, we will remain largely neutral about these issues for the sake of this discussion: we recognise that there can be, and often is, disagreement about the meaningfulness of someone's life but will not consider how to adjudicate such disagreement.

3. Narrative Structure

A further claim that is sometime made in this context is that it is necessary for having a meaningful life that one's life has narrative structure. For example, according to Charles Taylor, "a basic condition of making sense of ourselves is that we grasp our lives in a narrative".[4] And John Campbell says "identity [through time] is central to what we care about in our lives: one thing I care about is what I have made of my life"[5].

Related to this, Dennett says that it is being able to give a narrative to someone's life that allows us to interpret her behaviour as the behaviour of a single coherent subject, and encourages us to recognise or posit a unified agent.[6] Alasdair MacIntyre seems to be of roughly the same opinion when he says that "It is because we live out of narratives in our lives and because we understand our lives in terms of the narratives that we live out that the form of narrative is appropriate for understanding the action of others".[7]

What is meant by "narrative structure" in this context? This is not altogether clear but having narrative structure seems to involve at least being experientially diachronically extended, that is, extended across time in one's self-experience. To have this kind of diachronicity, according to Strawson, "one naturally figures oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future".[8]

But diachronicity alone is not sufficient for narrativity. In addition there must be some intelligible relation between the events that occur at subsequent times, and intelligible relations between the agent's actions at different times. For someone's actions to be part of a particular narrative these actions, and the events they bring about or are responses to, must cohere together in some way. For example,

[4] Ch. Taylor, *Sources of the Self*, Cambridge 1989, p. 47.

[5] J. Campbell, *Past, Space, and Self*, Cambridge 1994, p. 190.

[6] D.C. Denntett, *Consciousness Explained*. New York 1991, p. 418; *cf*. T. de Villiers, P. Cilliers, 'Narrating the Self: Freud, Dennett and Complexity Theory', *South African Journal of Philosophy* 23 [1] (2004), pp. 34-53.

[7] A. MacIntyre, *After Virtue*. London 1984, p. 212; *cf.* J.B. Schneewind, 'Virtue, Narrative, and Community: MacIntyre and Morality', *The Journal of Philosophy* 79 [11] (1982), pp. 653–663.

[8] G. Strawson, 'Against Narrativity', p. 430

they might be coherent in terms of aiming towards bringing about some goal, or as proceeding from some significant event, or as focused around some particular concern.

Note that the fact that diachronicity, as described here, is a property defined in terms of first-personal experience does not entail that the first-person perspective is privileged when it comes to determining whether or not someone's life has meaning. Diachronicity is necessary for narrative structure, which in turn we will argue is necessary for meaningfulness. But it is not sufficient for narrative structure or for meaningfulness. Whether or not there exists sufficient coherence between someone's diachronically extended experiences for her life to have narrative structure is an objective matter: it is something that she might be mistaken about and is not necessarily authoritative about. Indeed, it is possible that a third-person is able to make a more accurate judgment about this in some cases. So it remains an open question how to compare first-person or third person assessments of coherence and therefore meaningfulness.

This appeal to coherence helps explain why, when someone does something that does not cohere with her other actions, and that she cannot understand in terms of the narrative of her life, this action is typically either (a) marginalised as insignificant, unintentional, or out-of-character, or (b) it causes the narrative to fall apart, perhaps leading her to create a new narrative, or leading to crises of identity and self-understanding at the extreme (see Rorty, chap. 8). Consider for example, someone who spins her story as that of a decent citizen, but on a particular occasion acts in a way that a decent citizen would not ordinarily do. How she reacts to this will depend on how far it seems to her to fail to cohere with her narrative of being a decent citizen. For instance, if she has dropped some litter then she might respond by marginalising the action as unimportant, unrepresentative, or a minor aberration. In contrast, if she has murdered someone then she might find this harder to reconcile or excuse away. It might require her to reconstruct her self-understanding, perhaps as a rebel against society. Or it might initiate a personal breakdown as we sometimes see in people who commit a heinous crime having had a law-abiding, regular life up to that point. Or consider someone who had dedicated his life to religion, joined a monastery, and understood his life as organised around his ascetic dedication to God. Now suppose that he finds himself falling in love and having an affair with another of the monks, something explicitly forbidden by his religion. One way we might expect him to react is by rejecting his previous, passionless life as misguided. He might replace this by coming to understand himself as a caring partner in a loving relationship. But if he could not find this or another idea to organise his life around, that was acceptable to him, he might lose any sense of his life having meaning altogether. A real-life example of this latter reaction might be seen in the case of Daniel James. James was a young, promising rugby player who had prospects of a career as a professional sportsman. Unfortunately in a training accident, a collapsed scrum, he suffered a spinal injury that paralysed him. As his mother described, he "found his life so unbearable" and in 2008 sought assisted suicide in a clinic in Switzerland.[9] One

[9] See 'Mother Defends Rugby Suicide Son', *BBC News England*, March 10, 2010, Web. 30 Aug. 2011.

way of understanding James's reasons for this is in terms of his loss of the narrative structure in terms of which his life had meaning. He had previously understood his life in terms of his being a rugby player, and within the story of becoming a professional, perhaps striving for international honours, and so on. With this no longer being a plausible narrative for him after his accident we can understand him as being unable to conceive of some other narrative to fit himself into, finding his life to be meaningless as a consequence, and being unable to bear this.[10]

Of course, coherence and intelligibility are a matter of degree, so it is plausible that lives can have more or less narrative structure. Perhaps there is some threshold level of coherence that must obtain between the various actions and events within someone's life for it to be appropriate to describe it as having narrative structure. And perhaps what kind of reaction someone has to particular actions that do not cohere with their narrative will depend on the degree to which those actions do not cohere. But the notion should be clear enough.

There is support for this idea if we think about how we understand literary and cinematic characters. It seems that authors have to describe some degree of coherence in the behaviour and attitudes of the characters they create in order for those characters to be plausible.[11] Of course, one may always say that too much coherence can make a character seem artificial and contrived, but this objection would not pose a problem for the general idea that having narrative structure is necessary for a life to be meaningful. It merely suggests that not every aspect of a person's life need contribute towards its meaning. As mentioned above, there can be actions that are marginalised as unimportant. It also suggests that achieving meaning in one's life may be an active, ongoing process: that the agent, or an observer, must actively interpret her various actions and attitudes in terms of a coherent narrative structure that she adopts or is applied to her in order to achieve narrative coherence and meaning. At any particular time someone may have aspects of her life that have yet to be successfully interpreted in that way, or she may have aspects of her life that she is unable to interpret in that way. We might then expect to find that real people do have some aspects of their lives that fail to cohere with their main narrative. But we would still expect to find that they do have a main narrative with which these aspects fail to cohere.

Neither is there an objection to this view in certain recent approaches to narrative in literature and film. Recall Jean-Luc Godard's response to the question whether his movies have a beginning, middle, and ending: "Yes, but not necessarily in that order."[12] This attitude might be taken as showing that non-linearity and incoherence is not necessary for something to be a narrative, or to be meaningful (on the assumption that works created in that way can be meaningful). But it is useful here to evoke the distinction between fabula and *sjuzet*, or between

[10] It can be a risky business to speculate on the motivation underlying real cases, especially those of a sensitive nature. So we do not insist that this is actually what was going on in Daniel James's case. Rather, it is one plausible way of understanding it.

[11] See, e.g., L. Edelstein, *Writer's Guide to Character Traits*, Cincinnati 2006, pp. 10–12

[12] Cited in D. Sterritt, *The Films of Jean-Luc Godard: Seeing the Invisible*, Cambridge 1999, p. 20.

the narrative discourse and the story it (re)presents, even if it not without its disadvantages.[13] In most cases, the narrative incoherence in literary or cinematic works appears on the level of discourse, and the task of the reader is to reconstruct the (coherent) story that is represented by narrative discourse in a more or less twisted way.[14] If this is right, it lends support for the importance of coherence for narrative and meaning. It suggests that it is by finding, or creating, coherence in the prima facie incoherent events that are presented that the audience is able to find meaning in such fractured films and literature.

This allusion to artistic creation highlights another distinction that it is important to draw here. We must not confuse having narrative structure with being interesting. Characters in literature and film are typically, and hopefully, interesting: they do things that we want to watch and to read about. But someone's life might be dull yet have narrative structure nonetheless, if there is sufficient, and the right kind of, coherence between the uninteresting actions, attitudes, and events in her life.

We also find support for the claim about the necessity of narrative by considering ordinary language. In everyday situations, it is plausible that discourse about meaninglessness is related to the absence of narrative structure, or a failure to be able to recognise narrative structure. As MacIntyre describes, "When someone complains – as do some of those who attempt or commit suicide – that his or her life is meaningless, he or she is often and perhaps characteristically complaining that the narrative of their life has become unintelligible to them, that it lacks any point, any movement towards a climax or a *telos*. Hence the point of doing any one thing rather than another at crucial junctures in their lives seems to such persons to have been lost".[15] Relatedly, when trying to comfort someone who makes such a complaint, we typically try to offer them ways in which the lost narrative can be recovered, or propose new narratives that might be applicable and appealing to them. Suppose your friend's partner has left him, which has led him to quit his job, regularly drown his sorrows in alcohol, and become reclusive. He complains that his life is meaningless. To try to persuade him otherwise you typically would not say something like "but life is beautiful" or "life is good" or, "life has such-and-such feature." That would be a pointless strategy. Rather, you would say things like "Don't worry, she'll come back to you: she just needs sometime to think things over," or "there are so many beautiful girls, you'll find someone else. In fact, I know that X has been interested in you for a long time. . . ," or "This only proves that she is not your other half. You still have to search for the one who is truly your destiny, your soulmate," or even "There are other important things in life besides having a girlfriend, why don't you join the French Foreign Legion, or dedicate yourself to your music, for example?" You try to tell a story, or describe

[13] See J. Pier, *On the Semiotic Parameters of Narrative: A Critique of Story and Discourse*, [in:] T. Kindt H.-H. Muller (eds.), *What is Narratology? Questions and Answers Regarding the Status of a Theory*, Berlin–New York, pp. 73–98; *cf*. H.P. Abbott, *The Cambridge Introduction to Narrative*, Cambridge 2002, pp. 14–17.

[14] Which of course is not to say that there are no incoherent stories.

[15] A. MacIntyre, *After Virtue*, p. 217.

a narrative, with your friend as the protagonist in the hope that he will find it appealing and will be able to adopt it as a way of understanding his life. In that way he will once again be able to see himself as "having a future," rather than merely facing a series of empty days ahead of him.

Furthermore, we can see the practical relevance of this.[16] If someone understands her life in terms of a particular narrative this provides her with a decision procedure. When faced with difficult decisions we can reflect on what someone within the kind of narrative that we take ourselves to have would do, and what is the necessary choice for achieving what our narrative is directed towards. This helps to narrow down the space of alternatives under consideration, if not to determine the option we choose. And it prevents certain problems even becoming choice situations: say, in everyday situations, the way in which someone in such-and-such narrative acts is not open to question, so as someone within such a narrative this is how one will act. Again, we can agree with MacIntyre when he stresses the importance of stories for our practical and social education:

> We enter human society, that is, with one or more imputed characters – roles into which we have been drafted – and we have to learn what they are in order to be able to respond to us and how our responses to them are apt to be construed. It is through hearing stories about wicked stepmothers, lost children, good but misguided kings..., youngest sons who receive no inheritance but must make their own way in the world and eldest sons who waste their inheritance on riotous living and go to exile to live with the swine, that children learn both what a child and what a parent is, what the cast of characters may be in the drama into which they have been born and what the ways of the world are. Deprive children of stories and you leave them unscripted, anxious stutterers in their actions as in their words.[17]

The importance of narrative structure for meaningfulness, and the kind of role it might have towards that, is plausible and wide ranging.

4. A Conflict?

How does this claim about narrative structure relate to our earlier claim about the paradigm examples of individuals with meaningful lives? In fact, there seems to be a conflict. Recall, in those cases it seems that the agent's life has meaning in virtue of her standing in some appropriate relation, R, to some feature, F. But if this interpretation of those cases is correct, then it supports a claim about what is necessary *and sufficient* for having a meaningful life, that is, standing in R to F. Prima facie, this condition does not include having narrative structure as a clause, in which case narrative structure would not be necessary for having a meaningful

[16] *Cf.* R. Shusterman, *Regarding Oneself and Seeing Double: Fragments of Autobiography*, [in:] G. Yancey (ed.), *The Philosophical I: Personal Reflections on Life in Philosophy*, Lanham 2002, pp. 1–22.

[17] A. MacIntyre, *After Virtue*, p. 216.

life. Furthermore, it is not obvious what role might be played by narrative structure in the meaningfulness of those lives: what is it that having narrative structure adds to a life that might make it meaningful *if it is not* already? Although we were able to posit a practical role for narrative structure by considering ordinary discourse, perhaps this was not in relation to meaningfulness in the correct sense, the kind of meaningfulness in the examples and that we are concerned with.

This worry is brought out by reflecting on the original case of Albert experiencing existential angst. His life appears meaningless, at least from his perspective. Yet we can imagine that it has narrative structure. While it was not claimed that having narrative structure was sufficient for being meaningful it is puzzling what it could add such that it might plausibly be necessary for being meaningful.

5. Instant Agents, Patchwork Lives, and a Role for Narrative Structure

Consider the example of a meaningful life that you take to be most convincing from those above and the feature it involves. For the sake of exposition we will consider the example of the parent, but this specificity is unimportant. What seemed to make the parent's life meaningful was her relation with her children: her children were the feature to which she was related in an appropriate way. Now suppose an agent were to come into existence, persist for an instant and then go out of existence again. And suppose that this instant agent was an instant parent: for her brief existence she had children (perhaps they themselves persisted for an instant like she does). So she has the feature that we think confers meaning in the ordinary case. We might think of her as a duplicate of a momentary time-slice of the ordinary parent, who does have a meaningful life. But in contrast to that original example of the ordinary parent, it seems that the instant parent's life lacks meaning.[18] How can this difference be explained?

The problem concerns the relation that the instant parent stands in to her children: this is different from the relation that the ordinary, temporally extended parent stands in to hers. Despite the presence of the same, appropriate feature in each case (their respective children) only the ordinary parent stands in an appropriate relation to this feature. And the reason the instant agent fails to have an appropriate relation to this feature, despite being a duplicate of a momentary time-slice of the ordinary agent who does have such a relation, is that the appropriate relations are narrative relations. This is clear in the particular case of the parent: the ordinary parent's life is meaningful in virtue of her standing in a *parenting* relation to her children. And parenting relations are temporally extended, and narrative. For example, it involves, in addition to giving birth to them, nurturing one's child, ensuring that they have necessities such as food and clothing, teaching them to tie shoe-laces, teaching them to read, watching them

[18] At least, it will seem this way to a third person observing the case. Again this raises a difficult question about how first- and third- person assessments of meaning are related, and whether one or other is authoritative. But even if the instant parent were to believe, for the instant she existed, that her life was meaningful, and assuming this was not based on false beliefs about the duration of her existence, then intuitively we would judge that she is mistaken.

get older, setting up their first bank account, and so on.[19] The point generalises: it is not merely an artefact of this particular example of a way in which a life can be meaningful. Consider some of our other examples. For instance, dedicating one's life to religion involves regularly attending a place of worship, abstaining from certain impious behaviour, surrounding oneself with like-minded individuals, studying scriptures, and trying to spread "the word" and encourage others to also behave more piously. Undertaking some ethical project and performing good deeds involves doing things such as discovering and researching some particular need or deprivation, developing plans for addressing it, initiating the project and carrying it through to completion, following it up to see its effects and to find out if further work is needed and repeating similar projects elsewhere. Pursuing a particular career involves things such as imagining and planning the career path in advance, studying to achieve the necessary qualifications, starting at an entry-level position or apprenticeship, taking on various roles, projects, and responsibilities, seeking promotions and changes of employer, and so on. Note that these are not merely ad hoc choices, or examples gerrymandered to fit with our claim: they were all pre-theoretically plausible examples of meaningful lives. But they support the idea that having narrative structure is necessary for meaningfulness, in a particular way. In each case the agent's life has meaning in virtue of her standing in some appropriate relation, R, to some appropriate feature, F. And in each case the relation R is a narrative relation. So it is plausible that in the case of the instant agent, the reason why her life strikes us as lacking meaning is because she cannot stand in such a narrative relation to any meaning-giving feature.

But is the example sufficient to show that it is narrative structure that is necessary for meaning? The instant agent also lacks diachronicity, or temporal extension. Perhaps that alone is necessary for having meaning and it is merely a contingent, but irrelevant, fact that it tends to be coherent in real cases. In fact diachronicity without narrative structure will not do. We might understand someone being diachronic simply in terms of her being temporally extended such that her life involves a sequence of a number of events. This makes no commitments about the coherence of such events. But consider an agent who's life was diachronic but who's life lacked any narrative structure. It is composed of unrelated events with the agent as protagonist, each itself part of the kind of life that we recognise as meaningful, but randomly arranged with no intelligible relation between them.[20] So imagine the different ways in which Karl could live a meaningful life: in one case he is a devoted father; in another case a successful barrister dedicated to achieving recognition among his peers and a position as a judge; yet another possibility would see Karl as a freedom fighter and revolutionary trying to liberate the people of a Central American country against a despotic regime; in another

[19] Or some relevant subset of these and similar features. It is plausible that parenting is a family resemblance concept so as to allow for various ways in which particular parental relations can differ yet still be instances of the same kind of relation. Nonetheless, all such parental relations are narrative ones: an instant, non-narrative relation would not be sufficiently like others to be properly described as a parental relation.

[20] *Cf.* A. MacIntyre, *After Virtue*, ch. 15.

case he spends his life as a monk. Now try to imagine Karl's life as composed of sequence of events from these various possible lives in such a way that no two events from any particular alternative would follow one another. For example, on a particular day Karl might get-up at sunrise, daub himself in jungle camouflage, then head to the temple to meditate for an hour, next make some sandwiches for his children's lunch, then meet with one of his guerrilla comrades to plan a night-time ambush, next change from camouflage to court gown and wig, before returning to the temple for another hour of meditation, and so on. Intuitively such a patch-work existence is not meaningful despite the fact that it is diachronic and that each particular event is an event from a potentially meaningful life. The reason is the lack of coherence between the various events in patch-work Karl's life: his life lacks narrative structure.

So it is plausible that having narrative structure is necessary for having the kind of relation, R, to an appropriate feature, F, that is necessary and sufficient for having a meaningful life. The kinds of relations that are appropriate for providing meaning are narrative relations.

6. Conclusion

We have argued that narrative structure is a necessary condition for meaningfulness. That is, it is necessary for a life to have meaning that there is a sufficient degree of coherence between the different actions, events, and attitudes in an agent's life, and that these can be understood in terms of a narrative, for instance, as aiming towards some goal, or focused around some concern. Furthermore, we suggested why having such narrative structure was important for meaningfulness: where someone's life is meaningful this seems to be in virtue of her being appropriately related to some meaning-providing feature. But such appropriate relations are narrative relations.

This leaves open the possibility for a great degree of individual and cultural variation between different ways in which someone's life can be meaningful, in respect to the different features that can play the appropriate role, and the subsequent (narrative) ways in which the individual can be related to these features. This seems to be a positive feature of the argument: intuitively there is such variation so an account of meaningfulness should accommodate it.

Finally, we think that what we have argued sheds light on why various suggestions that have sometimes been made by philosophers about what having a meaningful life consists in often strike us as implausible or overly abstract. These suggestions often describe a particular feature, for example the pursuit of the highest good, or of the greatest pleasure, or the maximizing of the happiness of the greatest number of fellow-humans . But they neglect to describe how someone is supposed to relate to this feature, and how people can incorporate this feature into a coherent narrative. So these suggestions seem impractical and divorced from how we actually understand our lives whether or not the particular feature described, in fact, has some relevance to meaning. And in contrast it is revealing that those suggestions that people *have* found more persuasive, such as those of Nietzsche and Plato, were presented in the context of a narrative. For example, Nietzsche

used the story of Zarathustra to present the idea that one's life could be meaningful by striving for self-overcoming.[21] And Plato used the life and death of Socrates to present the idea that one's life could be meaningful by seeking knowledge of the form of the Good.[22] In each case the use of the narrative illustrates how the protagonist organizes his actions and attitudes in a coherent way in relation to a particular feature or goal and thereby this life seems to take on meaning. Yet independently of this narrative structuring the feature or goal would seem obscure and irrelevant like the other failed suggestions presented in an abstract way.

[21] On the role of narrativity in Nietzsche's conception of the meaningful life see, e.g., R. Rorty, *Contingency, Irony, and Solidarity* Cambridge 1989 (chap. 2) and A. Nehamas, *Nietzsche: Life as Literature*Cambridge 1985.

[22] Of course, it is worth noting here that Plato, as well as many other ancient philosophers, conceived of philosophy as a way of life. See, e.g., P. Hadot, *Philosophy as a Way of Life*, transl. M. Chase, Oxford 1995; R. Shusterman, *Practicing Philosophy: Pragmatism and the Philosophical Life*. New York–London 1997, and R. Shusterman, *Philosophy as Literature and More Than Literature*, [in:] W. Jost, G. Hagberg (eds.), *A Companion to Philosophy of Literature*, Malden –Oxford–Chichester 2010, pp. 7–21.

Epistemology and Philosophy of Science

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

STANISŁAW JUDYCKI
University of Gdańsk

Arguments in favour of global antirealism

Abstract

The aim of the paper is to give an analysis of the main arguments for so called global antirealism which are discussed in contemporary philosophy. As arguments for global antirealism are presented: (1) the fact that conformity of opinion, actual or potential, in all domains of discourse does not exist and thus objectivity must be interpreted as intersubjectivity, (2) the existence of different conceptual schemes which are used to interpret different kinds of data, (3) the replacement of correspondence theory of truth with so called epistemic theory of truth, (3) the verificationist theory of meaning, (5) the possibility of reduction (mainly naturalistic) of entities of some kind to entities treated as more basic. The end of the paper is devoted to the analysis of relations between contemporary forms of antirealism and forms of (old) idealism and between forms of global antirealism and forms of domain specific antirealisms.

1. Historical background

The traditional differentiation between realism on the one hand and the opposing positions, i.e. idealism and phenomenalism, on the other hand, was subject to significant transformations in the 1970s and 1980s due to events in English-language philosophy. The main impact was exerted by Michael Dummett (1925–2011), who not only introduced the very term 'antirealism', but also proposed that the ongoing discussion regarding realism and solutions opposite to it might be more fruitful if it were reformulated as a problem of two opposing theories of meaning.[1] Although in the course of further discussions, Dummett's terminological proposal was accepted, the scope of 'antirealism' was extended beyond the theory of meaning, and so in the 1980s 'antirealism' was referred to in many different ways.[2] In the early 1990s, Michael Devitt, one of the discussion participants, wrote:

[1] *Cf*. M. Dummett, *Realism* (1963), [in:] M. Dummett, *Truth and Other Enigmas*, London 1978, p. 145–165 (Polish translation: 'Realizm', *Principia* 6 (1992), p. 5–31); W.P. Alston (ed.), *Realism & Antirealism*, Ithaca–London 2002.

[2] *Cf*. E. Craig,*Realism and Antirealism*, [in:] E. Craig (ed.), *Encyclopedia of Philosophy*, vol. 8, London–New York 1998, p. 115–119.

> Dozens of different metaphysical, semantic, and epistemological theses jostle for the name 'realism' [...] The British School continues to write as if *the* realism issue were all a matter of what a speaker can or cannot 'manifest'. [...] Hilary Putnam ingeniously derives antirealism from just about everything.[3]

Today, the realism-antirealism opposition is already a natural part of the technical philosophical language, and the very term 'antirealism' has partly driven out both 'idealism' and 'phenomenalism'. However, the words 'idealism' and 'phenomenalism' are still used, especially with reference to metaphysical idealism (Plato, Hegel) or in the context of historical analyses relating for instance to Kant's epistemological idealism or Berkeley's phenomenalism. I believe that the following, tentative characteristics of antirealism can be proposed: according to an antirealism follower, nothing is what it is, irrespective of how we describe it with various notions, irrespective of how we think about it, and irrespective of how we talk about it. Alternatively, a positive approach may be used: *nothing can be real to us without reference to some theoretical conceptual scheme*, and every attempt at stating that something is simply real should be regarded as senseless. Such a formula of antirealism, which surely reflects one of its most general forms, is referred to as a conceptual relativity thesis. Recently, however, Michael J. Loux, in the third edition of his introduction to metaphysics, perhaps abrasively calling antirealists 'conceptual schemers', wrote:

> That moral is that there is something self-defeating in the conceptual schemer's account of conceptual representation. If the conceptual schemer is correct in claiming that the activity of conceptual representation bars us from an apprehension of anything we seek to represent, then why should we take seriously the schemer's claims about conceptual representation?[4]

Antirealism, as conceptual relativism in the above sense, on the one hand differs from the idealisms that preceded it, in particular that of Kant, although, on the other hand, it may be interpreted as a generalization of such idealism. Kant is known to have claimed that everything that we can perceive owes its structure to forms of intuition and the categories that the human mind is equipped with. Despite this, Kant did not negate realism in its entirety, as he regarded the existence of reality 'in itself', but, at the same time, he denied that such reality can be accessed. Contemporary followers of antirealism – antirealism in the sense of conceptual relativism – have generalized Kant's view as much as possible, claiming that everything that we regard as cognition is dependent on conceptual structuralizations that we do. Contemporary antirealists, however, no longer say anything about Kantian 'things in themselves'.

[3] M. Devitt, *Realism and Truth*, Oxford 1991, p. VII.

[4] M.J. Loux, *Metaphysics. A Contemporary Introduction*, Third Edition, New York–London 2006, p. 9.

Secondly, to Kant, structuralization that is done by the human mind to sense-data was something the only one, i.e. there existed the *only conceptual scheme* (a scheme composed of forms of intuition and categories)*, independent of human decisions*, by means of which we structuralize perceptions (*Empfindung*), while modern antirealism assumes that structuralization can occur by means of a multitude of conceptual schemes available to perceiving subjects.

Thirdly, antirealism, understood as conceptual relativism, repeats in a way Kant's distinction between that which is *empirically real* and that which is *transcendentally ideal*. When we remain within spontaneous applications of some conceptual scheme, the image of the world we thus obtain is to us something real, but only when we realise that a given scheme is one of many possible; it is also then that we see its constituting role for our cognition, and so we see that a spontanously produced image of the world is only something 'ideal'. This Kantian distinction between that which is empirically real and that which is transcendentally ideal, H. Putnam replaced with a distinction between internal realism on the one hand, and metaphysical realism on the other, claiming that we can only be internal realists, and this happens when we remain within a specific conceptual scheme, while we can never know what the world is like, irrespective of any conceptual scheme, and thus we can never be metaphysical realists.

One can remind here a concept of Rudolf Carnap, who distinguished between *external* and *internal questions* regarding the existence or reality of entities.[5] Internal questions are asked from the point of view of a certain accepted language, whereas external questions are only asked by philosophers who want to know, for instance, whether physical objects truly exist independently of the language of physics. According to Carnap, such questions, however, are either nonsensical or they are about – in a hidden way – whether a given language well suits our practical objectives. It seems, however, that in this manner one cannot eliminate the sense of the dispute between realism and idealism: if one intends to formulate statements by means of a certain language within a purely ideal system, i.e. a system that does not state anything about the world, then the problem of realism and idealism does not indeed occur, but if one wants to formulate statements about the world by means of a language, one has to specify criteria for when something truly exists and when it does not.

Fourthly, there exists *a bilateral independence between antirealism understood as conceptual relativism and idealism, resp.* metaphysical phenomenalism in the spirit of Berkeley. One may be a conceptual relativist, and so claim that everything we regard as real depends on one of many equally possible manners of conceptualization, and at the same time one may not recognize that everything that exists has an exclusively mental (spiritual) nature, and so, if one is an antirealist one does not have to claim that in reality, i.e. autonomically, there exist only God and the spirits created by Him. An independence also occurs in the other direction: one may be a follower of Berkeley's system, i.e. claim that there are only minds (spirits) and that which is dependent on them, but not recognize that everything

[5] *Cf*. R. Carnap, *Empiricism, Semantics and Ontology*, [in:] R. Carnap, *Meaning and Necessity*, Chicago 1956, p. 225–221.

we perceive (or everything that we think we perceive) is dependent on our, human ways of conceptualization, and so while being – in accordance with the above terminology – a metaphysical phenomenalist, one may claim for instance that although a tree in the yard exists depending on God's mind, but in its existence and endowment it is independent of our human ways of conceptualization of sensory excitements.

It seems therefore that if we take account of the above clarifications regarding possible relations between former idealisms and modern antirealism, we should draw the conclusion that in the case of the latter we have to do with a certain form of relativism that has been known since ancient times and that the entire discussion should move to the area defined for a long time now by the concept of relativism, which goes back to the time of Protagoras and is contained in the formula that man is a measure of everything (*homo mensura*). This is partly so, but only partly, because, as it follows from the above quotation from the book by M. Devitt, today there exist many concepts referred to as antirealistic which do not, at least not directly, have the nature of a conceptual relativism.What is more, if in relation to antirealism as a conceptual relativism doubts arise as to whether it replaces the issues of idealism in the style of Berkeley and Kant, the question becomes even more pertinent with regard to other positions referred to as antirealism. Thus, does for instance the position regarding the theory of meaning which is today referred to as semantic antirealism have anything to do with varieties of former idealism? Because if we assume that in the case of various uses of the concept of 'antirealism' we do not deal with something that can be referred to as *terminological nihilism*, i.e. an approach claiming that everything can be called by means of any term and that in the process one does not have to take account of tradition, then one should presume that modern antirealism has had to preserve some common 'intuitions' with the former idealism (*resp.* phenomenalism).

2. Varieties of antirealism

I would like to claim that the most general division that should be applied to modern types of antirealism is the distinction between global or universal antirealism on the one hand and domain (particular) antirealisms on the other. From the point of view of the former terminology, this distinction is to a certain extent a repetition of the distinction between agnosticism or universal scepticism, stating that all problems that can be formulated well are unresolvable in a finite number of cognitive steps, and agnosticisms, resp. particular scepticisms, in the case of which some form of essential unresolvability of problems in a specific object domain is claimed. The difference between domain antirealisms and particular agnosticisms (scepticisms) is such that the discussion that has been taking place in the last few decades has led to the creation of many new problems unknown to the previous generations of philosophers arguing for one of the possible detailed agnosticisms. The former agnostic positions already had – what we would call today – an antirealistic aspect, because the essential unresolvability of problems regarding some object domain had to arise from the nature of our mind, and so

from what it is capable of and from what it is not capable of; from what concepts it uses and what their epistemic scope is.

As far as *global antirealism* is concerned, its three principal varieties seem to exist in contemporary discussions, namely, the above-mentioned fact of conceptual relativity is used as a universal (supra-domain) argument in favour of antirealism. Secondly, of global antirealistic character is the position that the truth should only be interpreted using epistemic terms, and, thirdly, the thesis that the meaning of propositions should be defined in categories of that which can be rationally asserted seems to lead to universal antirealism. Thus, the global antirealist will claim that nothing can be real to us outside the use of a specific conceptual scheme, whereas the realist will negate that thesis. Similarly, the antirealist will state that only the epistemic notion of the truth is an operative notion, that is for instance such according to which the truth is equal to an agreement of scholars in a given domain. The realist will, however, assert that the truth does not exhaust itself in an agreement reached by scholars competent in a given domain, but that above all it consists in the agreement of our thoughts (judgments) with facts that are indepedent of us. The antirealist in the theory of meaning will claim that meaning (cognitive or factual meaning) should be defined in the categories of proposition (judgment) assertability conditions, while his opponent – the realist – will negate that, stating that the meaning of expressions should be defined by means of the so-called truth conditions.

As far as *domain* (local) *antirealisms* are concerned, it is obvious that there exist numerous very significant areas of human knowledge in relation to which one can adopt either an antirealistic position or a realistic position, but in fact realism and antirealism are propagated with regard to important, centuries-old areas of philosophical discussion. Today, the issue of realism and antirealism is being discussed with reference to the area of sensory perception, with reference to unobservables whose existence is assumed by natural sciences, in the area of abstract entities, where nominalist positions pass for antirealism, while realist positions coincide with either Platonic realism or Aristotelian realism. Realism and antirealism are being discussed in the field of philosophy of mind and, naturally, in the field of axiology, particularly in relation to the sphere of moral values.

The issue of realism and antirealism considered in individual domains enables one to be a realist in one or several domains, and an antirealist in other areas. Thus, for instance, one can be a (Platonic) realist as regards the existence of abstract objects (like the ones talked about by mathematics), at the same time being an antirealist in relation to objects that are not directly observable, and so objects whose existence seems to be assumed by natural sciences.

Arguments in favour of antirealism or realism in a given object domain may be unique to that domain, and so for instance a discussion about realism and antirealism related to objects that are not directly observable will make use of the argument of underdetermination of theory by data, but with reference to a specific domain of issues there may also be used *supradomain arguments*, for instance the above-mentioned conceptual or epistemic concept of the truth. On the other hand, both conceptual relativity, the theory of meaning formulated in the assertability

condition categories, and the epistemic concept of truth will have antirealistic consequences for all domains. If somebody claims antirealistically that the truth is nothing more than a product of an agreement of scholars competent in a given domain, then obviously he cannot be a realist in the domain of axiology, because that which is positively or negatively valuable (antivaluable) will not then consist in the agreement of evaluative judgments with objective states of things, but will become a result of a decision of some circle of people regarded as competent. Although the axiological realist will admit that the agreement of judgments with objectively existing states of things that he supposes, for instance an agreement with objectively existing values, must somehow be mediated by an agreement of the competent, but at the same time he will emphasize that it cannot come down to it, because only then can we assume that even an axiological decision of competent persons may be wrong.

It is clear then that if we refer the realism-antirealism opposition to various important object domains, then being in favour of realism or antirealism will require exploration of basically all topics fundamental to systematic philosophy. There is, however, another, even more depressing issue. If we take into account individual domains of philosophical discourse, the terms 'realism' and 'antirealism' will start to look only as redundant labels, because anyway in each of such domains the discussion has its own character and appropriate nomenclature for individual positions. It seems, therefore, that 'mass' talking about antirealisms in individual domains is really related to propaganda rather than any purely substantive or terminological needs. And what would happen if we looked at this domain antirealism cumulatively and tried to picture a philosopher who is an antirealist in all areas? Would that not come down to forms of radical relativism – that were known as early as the ancient times – reflected in the following formula: nothing exists, and even if it exists, we are incapable of perceiving it.

However, this is not the end of the complications generated by the contemporary uses of the term 'antirealism', as today the notion of 'antirealist' is applied to *reductionist positions*, which propose that it is possible to reduce one type of expressions and objects described with such expressions to expressions and objects of another type. That other type of expressions and objects is then regarded as more basic, both in the theoretical and ontical aspects. Reductionist antirealism also suggests that reduced objects are in this or another way fictional objects. Modern reductionist (or eliminative) antirealism is mainly naturalistic (physicalist or materialist) in nature. Today's reductionist antirealists do not approach their antirealism in Berkeley's spirit of reductionism, but in the opposite direction.

A model example of reductionist antirealism is the so-called *eliminative materialism*, which claims that mental states, as they present in internal experience, are fictional. Such fictionality of mental states will be proved in the future by empirical studies, which will clearly show that the only real states are neuronal states of the brain. Similar reductionist and antirealistic argumentation appears, as it were, on the opposite end of this topic, namely with regard to the status of moral values. In this case, it is claimed that one can prove that moral values are a useful fiction – a fiction created during the first socialization and the further

development of humankind; at the beginning, i.e. prior to socialization, there only existed egoistic needs, but a concession had to be made in favour of other entities – a concession aimed exclusively at satisfying one's own egotism. Thus, such values as, for instance, honesty or faithfulness were adopted so that everybody might be assured of a more reliable satisfaction of their own needs. Thus 'generated' moral values were then assimilated (internalized) as part of upbringing, and then they started to appear not as useful egotistic fictions, but as an objective 'being in itself'.

As far as *naturalist antirealism* with regard to mental states is concerned, it is certain that it is not based on the argument from the universal relativity of every conceptual scheme, as according to a vast majority of naturalists (physicalists), the language used by physics – present or future physics – accurately describes the ultimate nature of reality. One may think that no defence of this position, indicating that in the history of physics there have appeared various languages, such as the language of Aristotelian, Newtonian or Einsteinian physics, will help, because either such various languages have something in common and they may be called the language of physics or there exists no union between them whatsoever, and then any language, for instance the one used today to describe the production of dairy products, can be called the language of physics.

It seems that here one can see the limits of the ingenious derivation of antirealism from everything – as Devitt referred to Putnam's antirealist efforts – namely, if the argument of conceptual relativism were to be applied to the language of physics, it would unavoidably lead to the conclusion that followers of physicalist reductionism dogmatically and realistically take the position of the absoluteness of the language of physics, and so antirealism as conceptual relativism and antirealism as materialist reductionism (eliminationism) are mutually irreconcilable.

Antirealism, i.e. antirealism in the sense of conceptual relativism, is based on the assumption that notions constituting a given conceptual scheme exist and, what is more, truly exist, and so exist independently of any specific description, and thus conceptual relativists cannot accept the conclusion that notions as mental states do not exist, that they are fictions, while in reality there only exist states of human brains. Furthermore, the antirealists who substantiate their position with conceptual relativism cannot claim that everything is relative, i.e. that everything depends on a specific set of notions, because this thesis cannot be applied to such data as the very notions certainly are. Notions can also be regarded as a peculiar type of data which can be described from a superior level of reflection.

Apart from the division into global antirealism and domain antirealisms, and regardless of the principal types of global antirealism (conceptual relativism, epistemic interpretation of the truth, the theory of meaning of expressions and propositions referring exclusively to the so-called assertability conditions, and not to the conditions of truthfulness), there is also a division into *ontological* antirealism, *epistemological* antirealism and *semantic* (resp. logical) antirealism.

It might seem, which is in fact in line with the traditional, i.e. epistemological understanding of relativism, that relativist antirealism is about the extent of human knowledge, and this is principally the sense of the thesis of global antirealism

– namely, that we are unable to know the (ultimate) nature of reality, because we always have to use one of the many possible sets of notions in order to interpret data. Consequently, we can never know if a given conceptual scheme aptly captures the (ultimate) nature of reality. We can only substitute one set of notions for another. It should be pointed out, however, that a reference to the existence of a multitude of possible conceptual schemes is not necessary to justify agnosticism as to the ultimate nature of reality, because even if a multitude of equivalent sets of notions enabling the description of the same object did not exist, but only a single one such set existed, for instance one devised by Kant, i.e. a set composed of two forms of intuition of space and time as well as categories, then the problem of the ultimate nature of reality could still be formulated. Kant suggested that the problem only arises, because we realise that there is a difference between the excitation of our mind and the real cause of such excitation. We, humans, together with our mental states, are situated at the end of a certain chain of causes and we do not know the first and the real cause of our excitations. Objects of daily experience that we regard as causes of our states of mind are divided into a number of further causal components by natural sciences, but in this way we can never arrive at the first and proper cause of that which happens in our mind.

> It is also not necessary for us to limit the kind of intuition in space and time to the sensibility of human beings; it may well be that all finite thinking beings must necessarily agree with human beings in this regard (though we cannot decide this), yet even given such universal validity this kind of intuition would not cease to be sensibility, for the very reason that it is derived (*intuitus derivativus*), not original (*intuitus originarius*), thus not intellectual intuition, which for the ground already adduced seems to pertain only to the original being, never to one that is dependent as regard both its existence and its intuition.[6]

In the light of Kant's above remarks, it seems that contemporary antirealism, finding support in the argument of the anthropological conceptual relativity and thence deriving the inability of metaphysical realism, is rather like a product intended for a more common recipient and so it does not 'touch upon' the essence of the issue of metaphysical realism, which boils down to the question: what is, metaphysically, the ultimate source of not only the data we receive, but also of various sets of notions that are at our disposal?

In this context, the introduction into this discussion of antirealism in the sense of naturalistic reductionism that was referred to above complicates the picture of antirealism and seems to pass the problems in question from the epistemological issue to the area of metaphysical questions as to what truly exists. The answer provided by naturalistic antirealism is that *only 'matter in motion' truly exists*, including human brain neurons in motion, but this answer seems to lead to a total confusion of the discouse levels. This type of answer is a dogmatic metaphysical

[6] I. Kant, *Critique of pure reason*, B 72 (transl. by P. Guyer).

thesis, as then antirealists become those who recognize the existence of matter already at the starting point of the entire discussion, while the principal question with regard to realism and idealism seems to be one about what truly exists, and one should attempt to answer it without making the mistake of mixing up the discourse levels.

Not only metaphysical and epistemological aspects are discussed in contemporary publications devoted to realism and antirealism; they also talk about the *semantic side of the dispute*. In accordance with the predominant part of the philosophical tradition, until the mid-20th century, that which is semantic belonged to the sphere of cognition (knowledge), and so whatever semantics would be beyond that, it was understood as belonging to epistemology. Because, however, of a significant interest in language in recent philosophy, semantic issues have become independent, and some philosophers have even started to suggest something more serious – namely, that the semantic plane might be that neutral ground on which so far unresolved metaphysical and epistemological issues could be settled.

In the case of realism and antirealism one would have to be in favour of one of the two theories of meaning of expressions and propositions, i.e. the theory referring to the so-called truth conditions or to the theory explaining meaning by means of the so-called assertability conditions. This very distinction requires separate considerations, but it seems that the theory of meaning regarded as a '*gate to antirealism*', i.e. the theory of meaning interpreting meaning in the categories of the assertability conditions, is nothing else than a little 'retouched' verificationism of the Vienna Circle. This were to look in such a way that propositions whose sense (meaning) is impossible to be given in the categories of their assertability conditions are unresolvable, and so such propositions prove that the scope of our cognition has to limit itself to that which we are able to resolve by means of experience within the domain of real (natural) sciences or by means of definitions and constructions within the domain of formal sciences (logic and mathematics). Consequently, all the other problems fall beyond the human epistemic scope, and such a position is antirealism (previously called: verificationist agnosticism).

It must be emphasized, however, that it is not clear why unresolvability of some issues should be something antirealistic. On the contrary, one may argue, as evocatively done by Nicolai Hartmann, that the very existence of unresolvable problems is an argument for realism, as such situations show the limits of human knowledge and human mental activity; they are situations in which we come across something that is independent of our epistemic activity.[7] The modern antirealist thesis, however, has a different dimension – namely, it is a metatheoretical thesis, because if we assume that only those problems and those propositions have a sense for which one can specify the conditions of their assertability, based on experience or construction, then, as it were, the problem of metaphysical realism becomes automatically an unresolvable pseudoproblem. In the light of this cryptoverificationism, all the big philosophical issues will also become unresolvable – but this is an entirely different matter altogether. This is really not antirealism,

[7] *Gibt es ein Unerkennbares, so muss dieses notwendig unabhängig vom Subjekt dastehen. Es muss ein Ansichsein haben* (N. Hartmann, *Zur Grundlegung der Ontologie*, Berlin 1935, p. 176).

but meta-antirealism, i.e. it is a metatheoretical assertion that we are unable in any manner to go beyond the immanence of our experience and beyond our constructions in order to try to prove what truly exists, and so what is the ultimate source of excitations of our mind.

In the closing fragments of the monograph devoted to Dummett's semantic antirealism, Tadeusz Szubka writes:

> The essence of semantic antirealism, which is a consequence of the justification theory of meaning, are usually two theses. According to the first one, the meaning of propositions should be characterized in the categories of their substantiation or evidence attesting to their truthfulness, i.e. categories of warranted assertability. Thus, to know the meaning of a proposition is to know the conditions in which its assertability is warranted. Meaning is therefore recognizable by nature, i.e. – as it is sometimes put briefly – has an epistemic character. [...] According to the other thesis of semantic antirealism, reality is not fully determined, and so there are no grounds for the full application of the principle of bivalence and the law of excluded middle.[8]

At the same time, however, Tadeusz Szubka approvingly states that what Dummet has managed to achieve is:

> to connect traditional epistemological and metaphysical categories of disputes on idealism (independence of cognition, objectivity, evidence, etc.) with semantic and logical categories (meaning, truth conditions, bivalence, etc.) in a way that has resulted in the constitution of new, original philosophical positions: semantic realism and antirealism.[9]

Szubka defends Dummett against the accusation that he quotes that semantic antirealism, being a consequence of the justification theory of meaning, is nothing else than a set of discredited neopositivist ideas which have been slightly differently formulated and deprived of antimetaphysical rhetoric.[10] An entirely different opinion is expressed by the already quoted Loux, who writes:

> Of course, the philosophical community's assessment of the contributions of positivism might turn out to be a mere matter of intellectual fashion. It might turn out to be the expression of a prejudice that future generations of philosophers will manage to overcome. But what is significant is the fact that this negative verdict on positivism was part of the culture of Anglo-American philosophy in the 1970s, 1980s, and 1990s. If prejudice, it was deeply entrenched prejudice. How, then, was it that a bundle of claims so close to the central themes of logical

[8] T. Szubka, *Antyrealizm semantyczny. Studium analityczne*, Lublin 2001, p. 245–246.
[9] *Ibidem*, p. 244.
[10] *Ibidem*, p. 256–247.

> positivism could have succeeded in finding the prominent place on the philosophical stage that Dummett's work occupied?[11]

Loux finishes with:

> The fact is, I think, that we do not as yet have a satisfactory resolution of the historical puzzle of the 'Dummett phenomenon'. It is a puzzle that future historians of twentieth-century Anglo-American philosophy will need to address.[12]

Thus, we have three versions of global antirealism, domain antirealisms and the division into metaphysical (ontological) antirealism, epistemological and semantic (logical) antirealisms. Taking account of the divisions, it is impossible, however, to combinatorily generate an appropriate number of positions. The reason for this is the fact that contemporary philosophy has been developing spontaneously – just like the philosophy of the former times – and it seems that in the word 'antirealism' philosophers have 'sniffed out' easy spoils, which allow them, on the one hand, to get rid of socially awkward connotations associated with the word 'relativism', and on the other hand, it provides them with ample room for manoeuvre as regards presentation of various theories, which previously were referred to as agnosticism or scepticism.

3. Arguments: 'no convergence'

From the standard perspective, antirealism (or antirealisms) should be treated as a theoretical (philosophical) position (positions), and if positions are to be recognized as warranting discussion, they should have some justification, and so the thesis (theses) of antirealism should be separated from its (their) justification. In accordance with what has already been said, in the case of global antirealism such justification is to be conceptual relativity, an epistemic concept of the truth and a peculiar theory of meaning of expressions and propositions, whereas in the case of domain antirealisms either global arguments or arguments specific to a given domain will apply. When talking about contemporary antirealism, one obviously cannot present all the important domains of philosophical discourse in order to say at the end that those who challenge the autonomic existence of objects in any of such domains can be called antirealists. Such an undertaking would be impossible to do due to its scope, which would practically encompass the entire philosophy and would unnecessarily promote the term 'antirealism'. Thus, arguments in favour of global antirealism remain to be dealt with.

If reasons for global antirealism are concerned, solid candidates for arguments in this respect are: (1) the lack of convergence of views in a specific area of discourse (and perhaps in all areas); (2) the fact of existence of a multitude of alternative conceptual schemes used to interpret various types of data; (3) the rejection of the correspondence concept of the truth in favour of the so-called epistemic concept;

[11] M.J. Loux, *Realism and Antirealism: Dummett's Challenge*, [in:] M.J. Loux, D.W. Zimmerman (eds.), *The Oxford Handbook of Metaphysics*, Oxford 2003, p. 661–662.

[12] *Ibidem*, p. 662.

(4) a verificationist theory of meaning of expressions and propositions, and the related rejection of the principle of bivalence and the principle of excluded middle. This repertoire of arguments in favour of global antirealism also comprises: (5) a reference to the capability of carrying out a regressive analysis of some levels of data and notions to levels of data and notions regarded as more basic. Antirealism in this sense has been discussed above.

It seems that apart from the above arguments, there are no other arguments that could be recognized as reasons justifying global antirealism. As an argument for antirealism that might be a supradomain one, one might attempt to put forward for instance the *issue of vagueness*. One might consider whether the existence of borderline cases only indicates that our notions are not sufficiently precise or rather that borders between existing objects are vague. We would certainly not regard three grains of sand lying together as a dune, but how many grains of sand make a dune? Is it a problem of the fuzziness of our notions or rather a problem of 'blurred' boundaries between objects, which make up that which we call 'reality'? Timothy Williamson writes about it in this way:

> we may wonder whether an account of vagueness can distinguish in any principled way between some borderline cases that make reality vague and other that do not. At any rate, it has become obvious that the question 'Is reality vague?' must eventually be answered by comparing theories of vagueness overall.[13]

The problem with vagueness, as an argument for antirealism, seems to consist, among other things, in the fact that it is not clear whether there exist any borderline cases that would warrant the conclusion about the vagueness of the very reality. It is also not known in advance whether the word 'reality' within the discussion about vagueness exclusively refers to the world of spatial-and-time objects or whether it concerns all the possible regions of reality, and so also abstract objects, the psychological sphere presenting itself in internal experience, etc. Because of such doubts, I do not regard the problem of vagueness as a possible area of discussion on global antirealism.

As arguments in favour of global antirealism, universal constructivism or conventionalism might be used. It seems, however, that both constructivism and conventionalism fall under antirealism as conceptual relativism. Because it may only relate to certain specific object domains, I shall pass over the argument that the existence of certain objects should be rejected due to their '*queerness*'. A reference to the argument of 'queerness' would be for instance a basis for antirealism in the domain of moral values: a realist in this domain would assume the existence of 'queer' objects, i.e. objectively existing moral values which – according to him – would constitute a justification for standards of behaviour and exert a a quasi-causative impact on our consciousness. Against this, the antirealist – arguing in the spirit of naturalistic reductionism – would claim that objects of this kind do

[13] T. Williamson, *Vagueness in Reality*, [in:] *The Oxford Handbook of Metaphysics*, p. 705.

not exist, because 'non-queer' causative relationships are only possible between physical things.[14]

It should also be borne in mind that with reference to every argument in favour of global antirealism, each time the realist position will be different, and so – at least to a certain extent – the realist will claim something else as an opponent of a conceptual relativist, as an opponent of a follower of the epistemic theory of the truth, as an opponent of a follower of the verificationist theory of meaning, etc. These preliminary and only systematizing deliberations are to show how tangled up the contemporary uses of the terms 'anti-realism' and 'realism' are and that this state of affairs may effectively discourage one from making any attempt to conduct an ordering synthesis.

It is obvious that a discussion of the above-mentioned ways of thinking, even if only limited to the 'core' of each argument in favour of global antirealism, goes beyond this text, and so I shall limit myself to the first issue, i.e. the argument referring to the so-called lack of convergence.

A nearly obvious argument for antirealism seems to be the fact of divergence of opinion among specialists with regard to the existence and the nature of objects belonging to some specific object domain. The antirealist will thus claim that when one considers the issue of existence and the nature of abstract objects (e.g. mathematical objects), moral values or the issue of existence of beings assumed by natural sciences, it is clear that opinions on this topic differ, and so it should be assumed that either such objects do not exist at all or that they have a nature which is entirely different to that assumed by followers of mutually divergent views. It is also not true, as a supporter of this argument will point out, that one can assume that in the future, as a result of further discussions and studies, the different opinions about objects in a given domain will approach unanimity. Therefore, there is neither a present nor a potential agreement as to the existence and the nature of objects in a specific domain. This is rather a divergence similar to that which occurs in the area of sensory perception of colours, where a divergence of views about what colour a given thing is results from disparate conditions of perception.

Another antirealist approach in this spirit might be a statement that although there is some convergence of opinion in individual domains, it only results from the human nature we all share, not from the fact that our judgments autonomically refer to existing object domains. Thus, what we achieve is not an objective applicability of judgments, but their *intersubjective applicability*. We are unable to settle, for instance, whether there exist universalia independent of us and our agreement as regards various classifications only results from the fact that the classification schemes we apply come from the fundamental life needs that all people share. This argument hints at the existence of an unbridgeable gap between the idea of objectivity and objectivity as intersubjectivity.

As an example of particularization of the argument from the lack of convergence, we can consider the so-called pessimistic meta-induction, which is to be

[14] Thus argued by J. Mackie, *Ethics – Inventing Right and Wrong*, London 1977.

directed against the so-called scientific realism. The thesis of weak scientific realism can be formulated as follows:

> Most of the essential unobservables of well-established current scientific theories exist mind-independently.[15]

The *thesis of strong scientific realism*, on the other hand, would read as follows:

> Most of the essential unobservables of well-established current scientific theories exist mind-independently, and mostly have the properties attributed to them by science.[16]

Referring to Putnam, Devitt calls the following thesis a pessimistic meta-induction:

> The unobservables posited by past theories do not exist; so, probably the unobservables posited by current theories do not exist.[17]

If we take into account all the past scientific hypotheses, such as the phlogiston theory of burning, and we notice that its rejection was equivalent to the rejection of the existence of the phlogiston itself, as an unobservable factor that was to enable burning, then – according to the opponent of the weak scientific realism – it is clear that we cannot assert that most unobservables posited by current scientific theories exist, because it may turn out in the future that their existence will be rejected. Even the more so, we cannot claim that most of such unobservables possess the properties attributed to them by current scientific theories, because it is clear that modern natural theories argue about the properties possessed by such objects. In such a situation, the follower of scientific realism might, however, weaken his thesis and assert that some unobservables must exist, and although neither current nor past science could agree as to what such objects were and what properties should be attributed to them, on this basis we cannot claim antirealistically that such objects do not exist at all.

At this point, it should be emphasized that ontological (metaphysical) antirealism on the one hand and agnosticism on the other hand are, after all, two different positions: it does not follow from the fact that we are unable to determine what unobservables exist and what properties they have that no such objects exist. Will

[15] M. Devitt, *Scientific Realism*, [in:] F. Jackson, M. Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, Oxford–New York 2005, p. 769. At least two other arguments are put forward against scientific realism: an argument from underdetermination of theory by data and an argument against the principle of inference to the best explanation. The first one is about the fact that there may exist many different, but empirically equally adequate theories, and so theories stating different things about unobservables, and this is supposed to mean that we do not know the nature of such objects. In the other case, the scientific realist claims that the best explanation for prognostic and technological successes of natural sciences is to assume that unobservables (e.g. elementary particles) truly exist, while the antirealist will challenge the value of conclusions based on the principle of inference to the best explanation.

[16] *Ibidem*, p. 769.

[17] *Ibidem*, p. 784; this refers to H. Putnam *Meaning and the Moral Sciences*, London 1978.

such defence of scientific realism not be, however, the defence of the fig leaf realism? The term '*fig leaf realism*' should be understood as the thesis that some unobservables must exist, but what type they are and what properties they have we do not know now and we will not know in future, because views on this will never converge.

In this way, one may arrive at juxtaposing entity realism[18] or realism of causes and realism of theories, i.e. assert that scientific natural theories do not have to be true as regards properties attributed to unobservables, but have to be true as regards the existence of any beings at all, which should be treated as causes of specific, and not any other, results of experiments. Is it, however, so easy to separate beings posited by some theory from the very theory itself, and can the antirealist not – in the spirit of operationism and instrumentalism – claim that when for instance we use a so-called electron gun in order to determine the existence of some particles, then the word 'electron' does not refer to anything other than the fact that in future we will be able to carry out similar experiments and to take similar measurements and that we can credibly predict that such experiments and measurements will probably provide us with data similar to those we have obtained so far?

As regards the principle of inference to best explanation, the antirealist can claim that it is a principle not about the truth, but about probability, and that it can be interpreted as referring not to beings existing independently of us, but as a principle that refers to instrumentalistically understood credibility in the sense that the best explanation of the credibility of our predictions so far is their empirical adequacy, which in turn means that in future there will be experiments and technologies that will confirm the experiments done so far. In the eyes of the antirealist (agnostic), the realist that invokes the principle of inference to the best explanation as a principle about the truth of natural theories, in advance resolves the discussed issue of realism and antirealism in favour of realism.[19]

In this way, the main scheme of the current wide-spread distinction: realism-antirealism has been discussed. In conclusion, I presented in more detail one of the arguments in favour of antirealism, i.e. the lack of convergence of views in a specific domain of discourse. The issue of antirealism as interpreted by Dummett has only been 'touched upon', and above all no mention has been made of the main directions of the discussion on the argument for global antirealism consisting in the reference to the fact of existence of various conceptual schemes. I also give no illustrations of the discussion on the epistemic understanding of the truth. A discussion of such issues, even a brief one, goes beyond the limits imposed on this text.

[18] *Cf*. I. Hacking, *Representing and Intervening*, Cambridge 1983.
[19] *Cf*. B.C. van Fraassen, *The Scientific Image*, Oxford 1980.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

MAREK ROSIAK
University of Łódź

The Truth – a Correspondence, Coherence or Unconcealment?

Abstract

In the following paper I am trying to look at, from a systematic point of view, three influential and radically different, conceptions of truth with the aim to compare their pros and cons. I am referring to the classical (correspondence) theory of truth, to its chief rival – the coherence theory – and to the notorious, would-be pre-Socratic theory of truth as *aletheia* (Unconcealment), put forward by Heidegger. I do not regard the pragmatic theory of truth as an important counterproposal to the classical theory. Consensual theory can be ignored as a special (and obviously incorrect) case of the coherence theory. It shows some kinship to the pragmatic theory of truth as well if the last one is interpreted in an idealistic way.

There is no need to account for the utmost importance of the notion of truth. In spite of this, existing interpretations of this fundamental notion differ radically. In the following paper I am trying to look at, from a systematic point of view, three influential and radically different, conceptions of truth with the aim to compare their pros and cons. I am referring to the classical (correspondence) theory of truth, to its chief rival – the coherence theory – and to the notorious, would-be pre-Socratic theory of truth as *aletheia* (Unconcealment), put forward by Heidegger. I do not regard the pragmatic theory of truth as an important counterproposal to the classical theory.[1] Consensual theory can be ignored as a special (and obviously incorrect) case of the coherence theory. It shows some kinship to the pragmatic theory of truth as well if the last one is interpreted in an idealistic way.

A critic of the classical theory – for example, a proponent of the coherence theory – is not someone who just understands the word "truth" differently – he has some motivation for this. The controversy between different options about the nature of the truth is not purely verbal. Opponents of the coherence theory are

[1] An extensive (based on recent literature) critique of the pragmatic theory: D. Leszczyski, *Antyreprezentacjonizm, pragmatyzm, korespondencja* [*Anti-representationism, pragmatism and correspondence*], [in:] D. Leszczyński (ed.), *Prawda*, "Studia Systematica I", Wrocław 2011, pp. 235–278.

not satisfied either by its consequences or by its assumptions. The last objection is raised by idealists, who support the coherence theory. For them, there is no radical difference between thought or its linguistic expression on the one hand and reality on the other. Moreover, idealists point out that if thought were of a radically different nature than reality, their correspondence could not happen. Adherents of the pragmatic theory of truth require an effective ("empirical") method of establishing the truth of a statement. Thus, in my opinion, the pragmatic theory is not an alternative to the correspondence theory but is rather its (not quite successful) supplement.

The coherence theory, denying the realist assumption, accepts the general idea of a concordance between a true statement and something else, albeit not so radically different as in the correspondence paradigm. This element of a concordance can be found even in an odd conception of truth as "aletheia" in an etymological sense, as proposed by Heidegger.

1.

The noun "truth" should not suggest that truth belongs to the category of independently existing objects. What category, then, does it belong to? In principle, truth could be a property or a relation. According to the correspondence theory it is a relative property of corresponding, depending upon the correspondence relation between its bearer and its correlate. If we are ready to accept a common opinion that there are (also) temporary truths, it could perhaps be also a process or an event. "Becoming" of truth is obviously a Hegelian idea.[2] In Hegel's dialectic there is no sharp difference between language and reality. It is very symptomatic that Hegel entitled his chief metaphysical treatise: "*The Science of Logic*". On one hand, the items which are basic components of language – concepts – manifest some inner dynamism – a tendency towards dialectical development. On the other hand, reality itself is composed of concepts – categories in Hegel are not only the conceptual skeleton of reality, but also its content. So the relation of a concept, or of a sentence – which is for Hegel just a synthesis of concepts – to reality is not a relation of a correspondence, but that of a dependent part to a superior whole which is "more real" than its parts. A sentence or judgment is true insofar it is a part of some bigger system, but this system is itself a part of a still bigger system and so on. Moreover, the aforementioned dynamism of concepts consists in their tendency to "grow" – every part tends to become a bigger whole. And because a whole is "more true" than a part (as Hegel puts it – "*das Wahre ist das Ganze*") truth is a process of becoming more and more true, approaching the absolute truth which is identical with the totality of reality.

It seems that in Heidegger's *Being and Time*, the phenomenon of truth is described as if it were an event.[3] Heidegger underlines that truth happens to be discovered or – to put it differently – is a kind of an encounter. Truth is a revelation

[2] See *Preface* to *The Phenomenology of Spirit*, section II and the *Introduction* to *The Science of Logic*, part I: *The General Concept of Logic*, *passim*.

[3] A concise, essential description of Heidegger's conception of truth can be found in: J. Dębowski, *Prawda w fenomenologii* [*Truth in phenomenology*], [in:] *Prawda*, pp. 91–95

– something that appears momentarily. A typical example of such a revelation of truth can be found in the Bible, e.g. *Acta Apostolorum* 9, 3: the conversion of Saul on the way to Damascus. In that very moment Saul grasped the truth.[4] But, no matter how intense is such an epiphany of truth, no matter how important such an experience is for a conscious subject, it is not the only situation where the term "truth" is applicable. An epiphany of truth does not always have to precede a formulation of a true statement: the Copernican theory has not been revealed. Instead, it had to undergo a strenuous process of corroboration which proved that it was right from the very moment of its formulation. We can even say something true quite blindly, without any justification. Living in Poland I can say that it was a very hot summer in Norway this year and it can happen to be true notwithstanding my complete lack of knowledge in this field.

Both Hegel and Heidegger stress an existential aspect of truth – a dynamic character of its *modus existentiae*. For them, truth implies a change. This character cannot of course be ascribed to so-called eternal truths – they are explicitly denied by Heidegger.[5] According to the correspondence theory, not only eternal but temporal truths as well, have no dynamic character – they simply express something factual: truth is regarded here as a property of expressing something which there is. Nevertheless, if truth is a property of (certain) sentences, which are temporal beings, their truth is temporal as well. It could seem absurd to say that the truth of eternal truths is temporal, but this is an effect of a simple equivocation. The expression "eternal truth" refers usually to that which can be pronounced in a sentence, e.g. an axiom of geometry. In this objective sense, a truth can be eternal.[6] But the truth of a sentence expressing this axiom is temporal: without a sentence its truth cannot exist.

Categorial status of truth according to the theories taken into consideration can be characterised as follows:

Correspondence theory:	Truth as a property of a temporal object (property of expressing)
Coherence theory (Hegel):	Truth as a process (of development or growth)
Epiphanic theory (Heidegger):	Truth as an event (of an encounter or a disclosure)

2.

As we have seen, in all theories taken into account, truth is not an independent object – it must be referred to such an object as its qualification. Hence the next issue of our investigation: what kind of subject does truth belong to? In

[4] It is interesting to what extent the theological background of Heidegger influenced his conception of truth. His "epiphanic" conception is akin to Descartes' conception of evidence as a criterion of the veracity of an idea.

[5] M. Heidegger, *Being and Time*, transl. J. Macquarie and E. Robinson, London 1962, part I, Ch. VI, §44 c.

[6] I shall deal with this "objective" sense of truth in due time.

correspondence theory, truth qualifies a sentence or an act of judgment. In the coherence theory, the process of the dialectical development concerns concepts treated as components of the universe. The process of augmenting the truth is at the same time the development of reality itself. This is implied by Hegel's famous saying: "What is rational is real; and what is real is rational".[7] Thus, the proper subject of the development of the truth is reality itself. If "The truth is the whole", then the notion of truth is of no use: everything is partially true (as capable of further development – as a part of a bigger whole) and at the same time partially false (as not yet developed fully, as taken in isolation).

In the Heideggerian epiphanic theory "*das Dasein 'in der Wahrheit' ist*".[8] This means that truth is principally a human mode of being. Heidegger does not reject the interpretation of truth given by the correspondence theory, but he regards it as derivative, resulting from an encounter of *Dasein* with being. In a sense he is right, because it is we who produce true sentences. But being a source of linguistic phenomena does not necessarily imply being the ultimate subject of all qualifications concerning these phenomena. A description of an encounter of *Dasein* with being, which Heidegger gives us, reminds rather of the acquiring of knowledge. Of course, knowledge is *ex definitione* true, but knowledge, being a sufficient condition of truth, is not its necessary condition. The situation which Heidegger points to has too particular a character to account for all possible relevant instances.

3.

In the correspondence theory, truth is a relative property. What is then, the correlate of a true sentence? It is the reason of the sentence's truth, and as such it can be called a truth-maker (of a sentence). For example, what does the sentence "This pencil is yellow" say? It points out a particular object (my wooden pencil) and qualifies it as yellow. "Yellow" does not designate another particular object, somehow connected with the pencil. Such an understanding would be completely wrong, although Aristotle's mode of expression may suggest that: "who thinks the separated to be separated and the combined to be combined has the truth".[9] The "combination" here does not refer to a combination of parts making one whole. Such a connection is a symmetric relation, but the relation of subject and predicate is not: you cannot predicate a pencil of the color yellow. "Yellow", unlike "this pencil" is a general term, which can refer to different shades of yellow. If, after being exposed to a direct sunlight for a long time, the yellow color of my pencil has faded a little, the sentence "this pencil is yellow" will not become false. So a sentence like this does not refer directly to a particular in question, but *via* some of its general characteristics, for example, by being yellow.

[7] G.W.F. Hegel, *Preface*, [in:] *The Philosophy of Right*.

[8] *Cf*. M. Heidegger, *Being and Time*, II, 6, §44 b, p. 263. These words of course remind the biblical "I am the truth" (John 14, 6) or "Everyone on the side of truth listens to me" (John 18, 37).

[9] Aristotle, *Metaphysics*, 1051b 2–4, transl. W.D. Ross.

If the meaning of general predicates such as yellow, had purely conventional character (the standpoint of nominalism), all sentences having the subject-predicate form could express only analytical truths. Aristotle claims: "It is not because we think truly that you are pale, that you are pale, but because you are pale we who say this have the truth."[10] This implies the rejection of nominalism – and conceptualism as well – and thus, as far as any language contains some non-conventional general terms, it assumes a realistic option with regard to the status of universals.

In the coherence theory, it seems, there is only one general truth-maker – The Absolute. It contains all partial truths and even opinions which seem evidently false. If their falsity shows to be obvious, their negation has to be accepted and in this way a falsehood contributes to reach the truth. For Hegel the contradiction is the driving force of reality: there would be no progress towards truth without trying and testing wrong options. The ultimate truth-maker of the coherence theory has, however, a potential character only – it is a goal to which all reality strives in a process of self-improvement. Such a conception of the truth-maker allows for distinguishing different degrees of truth and falsity in terms of its relative proximity to the end of the process or something like a proportion of the part to the whole. A part is always less true than its superordinate whole. The less developed (less complete) a whole, the more false it is. Thus, to estimate the ratio of truth to falsity of a given item (statement) we must take into account the ultimate result, which has not yet been realised. It is doubtful if such a method of establishing the truth and falsity of a given statement is of any use.[11]

For Heidegger, the truth-maker of a given statement is an event of discovering such a truth by *Dasein*. But how can you discover (experience) negative states of affairs such as "There are no unicorns", for example? At best you can check that no existing evidence can support the belief in the existence of unicorns. But the nonexistence of unicorns is hard to experience. However, surprisingly, Heidegger claims that *Dasein* can somehow experience non-being. Nothingness as if exerts some action on *Dasein* – nihilation (*Nichtung*). The reaction of *Dasein* consists in some primitive anxiety (*Angst*). "Anxiety reveals the nothing."[12] The nothing is – in Heidegger's opinion – a kind of a surrounding or a context for the totality of being: "in anxiety the nothing is encountered at one with beings as a whole."[13] A result of the encounter with the nothing is *Dasein*'s understanding that beings are finite and destructible, as if they hover in nothingness: "The nothing itself does not attract; it is essentially repelling. But this repulsion is itself as such a parting gesture toward beings that are submerging as a whole. This wholly repelling ges-

[10] *Ibidem*, 1051b 7–9.

[11] However, during discussions with adherents of Hegel's dialectics, one often hears opinions like this: "Your standpoint is wrong, but it is understandable from a higher point of view." Saying so, the advocate of Hegel's views suggests that he can understand his opponent's opinion in a wider context and that from this point of view, the opinion in question cannot be accepted without reservations, but there are some limited reasons to maintain it.

[12] M. Heidegger, *What is metaphysics*, Part II: *The Elaboration of the Question*, §23, [in:] D.F. Krell (ed.), *Martin Heidegger: Basic Writings*, London 1993, pp. 93–110.

[13] *Ibidem*, Part III: *The Response to the Question*, §28.

ture toward beings that are in retreat as a whole, which is the action of the nothing that oppresses *Dasein* in anxiety, is the essence of the nothing: annihilation. It is neither an annihilation of beings nor does it spring from a negation. Nihilation will not submit to a calculation in terms of nihilation and negation. The nothing itself nihilates."[14]

In spite of the suggestiveness of this way of talking, there are serious doubts if we have here a description of any experience at all and not just an elaborate conceptual construction. Let us notice that *Dasein's* mode of being consists, *inter alia*, in *In-der-Welt-sein*. This means that *Dasein* exists among particular things and cannot reach – so to say – the border of the world (compare the memorable episode from the trailer of Monty Python's *The Meaning of Life*). In spite of this, Heidegger claims that there are special moods in which the totality of being is somehow given to us: "Even and precisely then when we are not actually busy with things or ourselves this "as a whole" overcomes us – for example in genuine boredom. Boredom is still distant when it is only this book or that play, that business or this idleness, which drags on. It irrupts when one is bored. Profound boredom, drifting here and there in the abysses of our existence like a muffling fog, removes all things and men and oneself along with it into a remarkable indifference. This boredom reveals beings as a whole."[15]

It is true that sometimes we feel bored or tired with "everything", but here this term only means "everything we have encountered so far", or even less: "everything we remember at the moment". By no means have we referred here to the experiences yet to come. Some people in this mood make extraordinary decisions, searching desperately for a radical change. Even if you feel so ultimately bored or tired that you do not even think of starting something new, this only means that you do not expect anything interesting in the future. This is a completely different perspective on things and the matters which surround us than a would-be "revelation of all beings as a whole". Without reaching this overall perspective on the totality of being, one cannot experience "the other side" of it – the nothing. This is the opinion of Heidegger himself: "Nihilation is not some fortuitous incident. Rather, as the repelling gesture toward the retreating whole of beings, it discloses these beings in their full but heretofore concealed strangeness as what is radically other with respect to the nothing."[16] Thus, the nothing is just the "opposite side" of the totality of being. I cannot grasp the totality without its border (if there is any) and the border refers to the "other side".

But, even if it were somehow possible to realise the totality of being, a perspective on the nothing would not necessarily open. If the totality is the Absolute (Hegel's opinion) there is no borderline between being and nothing. Having no experience of the totality of being we do not know if it is finite or infinite.

It is surprising that Heidegger, who so confidently refers to a pre-Socratic Parmenides' conception of *aletheia*, so radically differs from ancient thinkers and Parmenides in particular, with regard to the views on non-being. If one thinks

[14] *Ibidem*, §31.
[15] *Ibidem*, Part II, §18
[16] *Ibidem*, Part III, §31

that pre-Socratic thinkers managed to grasp philosophical matters better than their successors, one should not ignore their authority without explanation.

One can see that Heidegger's conception of experience as the ultimate truth-maker, cannot withstand criticism both with respect to the general idea, which has too narrow a range of applications and with respect to certain particular applications as well.

4.

"To Dasein's state of being belongs *falling*. Proximally and for the most part, Dasein is lost in its 'world'. [...] That which has been uncovered and disclosed stands in a mode in which it has been disguised and closed off by idle talk, curiosity and ambiguity. Being towards entities has not been extinguished, but it has been uprooted. Entities have not been completely hidden; they are precisely the sort of thing that has been uncovered, but at the same time they have been disguised. They show themselves, but in the mode of semblance. Likewise, what has formerly been uncovered sinks back again, hidden and disguised. *Because Dasein is essentially falling, its state of Being is such that it is in 'untruth'.*"[17] Here we have a conception of falsity (untruth) as principally *Dasein*'s giving up being in truth, i.e. instead of experiencing things personally, rather to form unjustified opinions or repeat someone's else statements about the matter. We have already observed that Heidegger's conception of truth is too narrow and now we see that as a consequence of this fault, his conception of falsity becomes too wide. For if I just guessed that it would rain the next day, according to Heidegger my statement was false, even if it actually rained.

When I am looking for information, I just need an adequate report how things really are and I do not care if my informer has experienced relevant facts personally or maybe got to know them from some reliable source. In this second case nobody (with the possible exception of Heidegger's believers) would complain that he was misinformed. Heidegger is completely wrong in claiming that his own conception of truth allows us to understand the correspondence theory better: "Our analysis takes its departure from the *traditional conception of truth*, and attempts to lay bare the ontological foundations of that concept. In terms of these foundations, the *primordial* phenomenon of truth becomes visible. We can then exhibit the way in which the traditional conception of truth has been *derived* from this *phenomenon*."[18] These theories are incompatible, and moreover, there are strikingly counterintuitive consequences on the side of Heidegger's theory.

While unnecessary subjectivisation of falsehood is the principal flaw of Heidegger's theory, the correspondence theory has its own problems with falsehood and they have objective character in turn. In Plato's *Sophist*,[19] we have an interesting analysis on the problem of non-being. According to Plato it must exist somehow because otherwise there would be no false sentences.[20] Every declarative atomic

[17] M. Heidegger, *Being and Time*, I, 6, §44 b, p. 264.
[18] *Ibidem*, p. 257.
[19] Plato, *Sophist*, 236e–241e, transl. B. Jowett, http://classics.mit.edu/Plato/sophist.1b.txt.
[20] *Ibidem*, 237a.

sentence asserts something definitive. If it is true, it says something about being. If it is false however, then there is no being corresponding to what it says. We cannot claim that false sentences do not say anything – in that case they would have no sense, while being false implies having sense. On the other hand, if we agree that false sentence actually refers to non-being, there should be non-being as its correlate and, moreover, it should possess characteristics ascribed to it in the sentence in question.

But Plato did not intend to postulate a separate region of non-being, populated by subsisting objects like dragons and square circles. Unlike Heidegger, who frightens his readers with horror stories of the nothing which nihilates being, Plato (who is, by the way, a much better storyteller than Heidegger) decides to root non-being in being. He makes it by accepting negative states of affairs, like for example Socrates' not being handsome:

> Stranger: Then, as would appear, the opposition of a part of the other, and of a part of being, to one another, is, if I may venture to say so, as truly essence as being itself, and implies not the opposite of being, but only what is other than being.
> Theaetetus: Beyond question.
> Str.: What then shall we call it?
> Th.: Clearly, not-being.[21]

Then the false atomic sentence saying that Socrates is handsome refers in fact to an existing object – Socrates – and only wrongly ascribes to him the quality of being handsome which actually he isn't. This sentence is false in virtue of the objective negative state of affairs that Socrates is not handsome instead of being false because of the nonexistent state of affairs that Socrates is handsome.

This theory cannot help in the case of non-existent, but possible objects like the first man on Mars. Its further elaboration, overcoming this flaw, can be found in Wittgenstein's conception of negative facts formulated in *Tractatus logico-philosophicus*. Even before studying this advanced version of Plato's conception of non-being, one can expect that this approach is more promising than stories of the nothing which nihilates.[22]

In the coherence theory, the falsehood of a sentence consists in its neglecting mutual relations of the situation reported by the sentence with the rest of reality. The coherence of the theory means here its completeness. The only unconditional, 'full truth', is a definite description of the total reality. The Absolute Idea is such a description and the total reality at once. We can see that the coherence theory has strong metaphysical monistic assumptions. The theory postulates not only what the essence of truth is, but also what the ultimate truth is. This does not happen with the correspondence theory which, nevertheless, has its own metaphysical assumptions too.

[21] *Ibidem*, 258b.

[22] A systematic and detailed account of ontological aspects of the correspondence theory of truth is: M. Piwowarczyk, *Prawda i ontologia* [*Truth and ontology*], [in:] *Prawda*, pp. 99–126.

5.

The problem of the criterion of truth belongs to the essentials of the every theory of truth. Adherents of non-classical theories often think that this problem cannot be satisfactory solved by the correspondence theory – and that this is its fundamental flaw. On the other hand, some adherents of the correspondence theory try just to ignore this problem, claiming that the definition of truth does not have to refer to it. In other words, they seem satisfied with a non-operational definition of truth.

However, it seems that the correspondence theory does not leave the problem of a criterion open. The truth of a sentence depends on its correspondence with the sphere to which the sentence refers. To check if the sentence is true one has to know what the sentence says and if the situation to which the sentence refers has happened. When I say that I feel cold and I actually feel cold, I have said the truth. But the critics of the correspondence theory want an effective application of the criterion to the extra-mental reality. We must remember that the correspondence theory itself does not assume the standpoint of metaphysical realism. Of course, Aristotle was a metaphysical and epistemological realist and the correspondence theory was basically his idea, but such a "personal union" does not necessarily imply the logical interdependence of the theories in question. When Aristotle says: "It is not because we think truly that you are pale, that you are pale, but because you are pale we who say this have the truth"[23] he obviously interprets his definition of truth in a manner of metaphysical realism, but this is only a possible particular interpretation of a general formula. Snow can be white not only for a realist, but for a subjective idealist as well. The formula: "The sentence "Snow is white" is true iff snow is white" does not assume any particular *modus existentiae* of snow. In particular, it does not assume an existential independence of the object of judgment from the act of judgment.

The second important thing is that non-applicability of the criterion in certain special cases does not have to be regarded as its flaw. Certain philosophical standpoints can disable a criterion and if somebody puts forward an alternative criterion, it should be checked carefully if the alternative really works in the same context where the former tool failed. The case of the coherence theory is a very good example of such an oversight. An adherent of this theory is right that the correspondence theory cannot help in case of metaphysical realism combined with epistemological idealism (a possible example of this standpoint can be Lebnitian monadology deprived of the pre-established harmony). But also the coherence theory does not help in the slightest way in that case. No coherent system of statements about the external world can guarantee its own relevance. The context in which this theory (and its criterion) works is metaphysical idealism, like this of Hegel, where there is no absolute difference between thought and reality. The combination of metaphysical realism and epistemological idealism disables all criterions of truth altogether. No criterion of truth can help to get access to things in themselves and this is not a fault on the side of the theories of truth.

[23] Aristotle, *Metaphysics*, 1051 b 7–9.

6.

As we have seen, theories of truth can presuppose some metaphysical or epistemological conditions. The idealistic option is obvious in case of the Hegelian coherence theory. But, if according to this theory, being true (to a certain extent) means just to be a certain part of reality, the term "truth" shows to be superfluous.

F.H. Bradley tried to make absolute idealism more intuitive and intelligible than in original Hegel's version. In his essay *On Truth and Coherence*[24] he shapes the doctrine of absolute idealism in a manner similar to that of subjective idealism: a subject is given a stream of sense data which can be interpreted as representing external reality in different manners. No sense data determine their interpretation in a unique way, so the criterion of choice for the best interpretation of the data given is its comprehensiveness and coherence: "With regard to the two aspects of coherence and comprehensiveness [...] I have merely urged that it is necessary to use them in one, and that here, and here alone, we have the criterion of perceived and remembered truth. I have argued that, in principle, any judgment of perception or memory is liable to error."[25] There can be no direct infallible knowledge of external reality for him, so truth can only be understood as a property of a totality of statements interpreting in a coherent way the sense data registered until a given moment. Then, as new sense data continually come, there can be no definite truth about the external world. A description of reality can always have only a tentative character.

This view, however, is not radically opposed to the standpoint of the correspondence theory. Truth, according to Bradley's conception, is nothing more than just a hypothetical truth according to the correspondence theory. Moreover, we must remember, that for Bradley, comprehensiveness is as important as the coherence. This means that we cannot ignore sense data – they limit the range of possible interpretations. If I had an impression of a patch of red I cannot interpret it as a waving of Prophet Muhammad's flag. In effect, although no statement within the framework of Bradley's interpretation of the coherence doctrine can be definitely true, some statements can be definitely false. This is some kind of a (negative) correspondence between the sentence and some sphere different than the system of statements alone.

The consequence that the Bradleyan version of coherence theory truth can be only hypothetical stems from the assumption of subjective idealism. The correspondence theory as such makes neither realistic nor idealistic epistemological assumptions – this is not its task to engage into the realism-idealism controversy. However, it makes another assumption: that thought and language, because of their intentional character, are not direct parts of reality. The relation of intentionality creates a distance or difference between thought or linguistic expression and the object they refer to. The identity of thought and being, postulated by Parmenides is excluded. That his formula of identity cannot be reconciled with the existence of a language was observed already by Plato:

[24] F.H. Bradley, *On Truth and Coherence*, [in:] F.H. Bradley, *Essays on truth and reality*, Oxford 1914, pp. 202–218.

[25] *Ibidem*, p. 218.

> Stranger: To distinguish the name from the thing implies duality.
> Theaetetus: Yes.
> Str.: And yet he who identifies the name with the thing will be compelled to say that it is the name of nothing, or if he says that it is the name of something, even then the name will only be the name of a name, and of nothing else.
> Th.: True.
> Str.: And the one will turn out to be only one of one, and being absolute unity, will represent a mere name.
> Th.: Certainly.[26]

Hegel tried to restore the unity of thought and reality, although within a dynamic paradigm. But his gain is of not much value: The coherence theory is in fact a kind of a deflationary theory *avant la lettre* – truth becomes, in principle, a redundant term.

The above comparison of the correspondence theory of truth, with its two much disputed counterproposals, shows that it can withstand criticism quite well. Instead of uncover its disadvantages, alternative theories showed their own unintuitive consequences and arbitrary presuppositions. Reports of the death of the correspondence theory have been greatly exaggerated.

[26] Plato, *Sophist*, 244d. It's funny, that Heidegger tries to root his epiphanic theory of truth in Parmenides' theory while the radical monism of the latter in fact does not allow of any theory of truth (and, indeed, no theory at all).

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

ROBERT PIŁAT
IFiS PAN, Warszawa

Truth as an internal standard of judgment

Abstract

Donald Davidson was sceptical about the possibility of having a definition of truth and useful criteria of truthfulness at the same time. Davidson's conclusions seem right in relation to truth expressed with a single-argument predicate indicating a certain property of a sentence. In the article, I defend the view inspired by Edmund Husserl's deliberations that 'truth' is best expressed with a two-argument predicate, as it belongs to the internal structure of a judgment. I understand a 'judgment', in Husserl's spirit, as a relation between thought and the object it is captured by. I argue that this way of understanding truthfulness is the best one to reveal its prescriptive aspect. I present truthfulness as an internal standard of judgments. I assert that truth in this meaning can be reconciled with the disproportion of criteria for establishing truthfulness in various areas of knowledge. Truth as a standard is applicable to ordered pairs of cognitive states – the ordering expresses the fact that the occurrence of the first state produces a peculiar cognitive obligation to accept the other one. It does not seem that the notion of truth as a standard might be constructed in this way for sentences. The problem is that our practices of turning from untrue sentences to true sentences are incommensurable – they are not subject to any one principle which would allow the construction of a set that might be recognized as an extension of the notion of truth. However, it seems that such a principle may be sought for judgments. A certain line of critique by Alfred Tarski of the semantic definition of a true sentence, presented for instance by Ernst Tugendhat, indicates the assumption regarding the truthfulness of judgments embedded in the definition. I am looking for such a description of judgments that would allow me to verify the intuition that the essence of a judgment consists in its being subject to the standard of truthfulness and that the material content of the standard is shared by all types of judgment. Based on Edmund Husserl's views presented in his *Logische Untersuchungen*, *Formale und transcendentale Logik* and *Erfahrung und Urteil*, I formulate a hypothesis that the pair <intention, fulfilment> and the phenomenological notion of truth in the context of a full presentation of an object is applicable to all types of judgment and carries the standard of truthfulness regulating cognitive activities. In *Erfahrung und Urteil*, Husserl compares such activity

to the satisfaction of desire, namely, the desire to possess an object in its self-presentation more and more fully. I think that the comparison may be given a less metaphorical sense by a retentive-and-protentive analysis of the structure of acts of judgment.

1. Four challenges as regards truth

The 20th century enriched and complicated our reflection on truth. The notion of truth was subject to fourfold critique:

1. *Relativization to language.* Alfred Tarski proved that it is impossible to formulate a definition of a true sentence for a natural language due to the impossibility of providing an unambiguous translation of sentences in such a language into appropriate sentences in the metalanguage containing names of expressions of the source language and other expressions needed to describe the source language. A definition of truth is only possible for an ordered language, in which a formal principle of creating the metalanguage is given.

2. *Irreducible multiplicity of truth criteria* Criteria for justifying scientific sentences differ depending on the discipline. We have to respect various standards of justification, in accordance with the nature of the very objects subjected to study as well as with the current state of the theoretical development of a given science. Thus, the notion of truth has no uniform meaning. At attempt at providing a definition of truth in a situation in which various ways of justification fail to share a common denominator leads to the formulation of a notion that is empty and redundant in science.

3. *Truth as part of a power system.* Truth and other cognitive values are ways of standardizing human mental activities, but we do not have cognitive states directed at the very property of truth – the very truth is not subject to real perception. This is why when talking about truth, we do not talk about the properties of all true sentences. We rather refer to entire systems of knowledge. However, we recognize or reject these on a principle that is entirely different than that in the case of true sentences. Michel Foucault indicated for instance the role of social power – according to him, truth is one of the elements of legitimizing and exercising power.

4. *Pragmaticist reduction of truth.* In the tradition of pragmatism, truth is the basis of activity. Here, the equivalent of truthfulness is the resolution of controversies or selection from among alternative solutions. Cognition is used to find practical solutions, to reach a consensus, to plan the future, to assess the degree of responsibility, to make decisions, etc. Truth is not something separate or superior to such goals, but simply a fulfilment of such goals. In the pragmaticist perspective, truth retains its prescriptive power, but only as a set of other cognitive norms.

The above-mentioned sceptical arguments with regard to truth concern either definitions of truth, the criteria for truth, or the prescriptive power of truth. It is also emphasized that these three aspects of truthfulness support each other insufficiently. For instance, Tarski's formally correct definition of truth has no peculiar consequences either for the problem of the criteria for truth or the problem of the

prescriptivism of truth. In turn, a focus on perfecting the criteria for determining truthfulness in science requires that one rejects any definition of truth, although undoubtedly it brings something into the understanding of the normativity of truth. The thinkers who start from the prescriptive power of truth apply it to practice rather than cognition. To them, truth is a set of criteria for justification or consensus adopted by a given community.

In the deliberations that follow, I try to substantiate the view that the notion of truth in all of the three aspects (definition, criteria and norm) has a uniform sense, which becomes visible when we refer the notion of truth to judgments and the act of judgment, and not only to sentences. However, the reference of the predicate 'truth' to judgments requires that the efforts undertaken by Edmund Husserl in *Formale und Transcendentale Logik* and in *Erfahrung und Urteil* be continued – the efforts that have been pushed away unfairly to the margin of the contemporary analytic philosophy.

2. Deflationary consequences of the critique of truth

The following opinion is characteristic of contemporary scepticism with regard to truth:[1] The multiplication of problems resulting from the desire to define truth (to capture its essence or its aprioristic relations to other values) is unnecessary. The dictate related to truth simply requires that true sentences be formulated in the best possible way – in order to describe reality as best as possible in a given situation.[2] This position can be split into the following mental steps: (1) the truth norm applies to situations in which it is possible to determine truth, but – taking account of possessed information – it is also possible to formulate an untrue sentence; (2) all information about the content of the notion of truth and truth norm lies in conditions that differentiate true sentences from false ones; (3) the conditions can ultimately be brought down to the best possessed justifications of the articulated sentences; (4) the justifications should be understood non-atomistically (with regard to a specific sentence under consideration), but as elements of broader descriptions of reality.

This position is very suggestive and, as a certain generalization of scientific practice, legitimate. I do not think, however, that it satisfies the philosophical aspirations that lie at the base of deliberations on truth. First of all, we should notice that conditions allowing true sentences to be distinguished from false ones may be based on something entirely different than truth – namely, on an accidental correspondence of a set of sentences uttered by a given person with the set of

[1] The opinion summarized below reflects the spirit of the entire family of deflationary solutions that differ from one another in material detail. However, such details are of no importance to our present deliberations.

[2] I would like to express my gratitude to Prof. Andrzej Grzegorczyk for a clear formulation of this position in an informal conversation about the essence of Tarski's results. According to Prof. Grzegorczyk, talking about a *definition* of truth is not very accurate in this case – it is rather about a formally correct generalization of the intuitive understanding of truth to sentences created on the basis of a given language. A somewhat similar direction is followed by an analysis by Adam Nowaczyk, which I will refer to further on in this paper, when I attempt to better understand that intuitive understanding of truth by referring to Husserl's analyses.

true sentences. Such correspondence may be founded on certain properties of the world that really limit the possibility of error, but do not enable insight into the justification of the sentences uttered. In other words, we are always under threat from the aporia formulated by Davidson: The reasons for which we recognize something as true are something different from that which makes such sentences true.

One of the ways of dealing with this aporia is to weaken the requirement that the reasons for recognizing sentences as true and the reasons for which such sentences are true should match each other exactly. If we combine such a weakened requirement with a certain optimism as regards the legitimacy of scientific methods, we will arrive at a conclusion that every sentence is true in only one way that is unique to it: by correspondence to a certain state of things described in an appropriate methodology. Such a formulation of the practical clarity of the notion of truth does away with the need to look for its definition. Anyway, we do not need any more precise definition than that contained in Aristotle's classic formulation. Attempts at making it more specific for certain languages in the way described by Tarski would not have any special impact on cognitive practice and would not be helpful in formulating cognitive ideals or norms.

However, the solution referred to above is unsatisfactory. Even if a sentence can only be true in one way, it can still be false in many ways. The norm of truthfulness calling for the determination of truth refers to pairs of sentences, in which one is not true and the other is true. Although this situation can be presented as a situation of choice, because a sentence can be false in many ways, the choice becomes multifaceted and not very clear. It rather consists in a complicated process of rejecting inferior hypotheses, which calls for standard scientific procedures: designing experiments, improving measuring instruments, better conceptualization and visualization, etc. Ultimately, we replace pronouncement of truth and falsehood with pronouncement of better or worse justifications.

Thus, we come to a conclusion that is slightly too strong in relation to the intent of some sceptics. Davidson claims for instance that it is possible to use the notion of truth without defining it, i.e. without establishing relations of semantic nature. However, by reasoning as above, we not only do away with the definition of truth, but also the very notion of truth, replacing it with the notion of a warranted assertability. At best, we leave truth as a certain metaphor that describes all our cognitive efforts. This is not what Davidson intended.

An interesting line of critique of the reduction of truth to the validity and warranted assertability was adopted by Schnädelbach. He claims that all holistic, pragmatic or verificationist interpretations of truth limit truthfulness to validity (*Gültigkeit*). That last notion is of importance to science and probably sufficient for it. This means that a given judgment can be proved for all the relevant cases. However, it does not work in a general philosophical sense. Truth as a general norm of cognition has neither attributed subject nor a set of relevant cases. This is why it has always been so important to differentiate the expression 'true' from 'true

as' or 'true with relation to'.[3] And so an intellectual challenge is posed by Kant's concept, who tried to combine these two senses of truthfulness. The philosopher's reasoning included the following three steps: (1) he defined truthfulness as importance, as truth consists in a certain reference to reason; (2) he gave conditions of universal validity (for all rational beings); (3) thanks to the previous step, he did away with the relativizing condition and was able to talk directly about truth as a certain specific idea – equivalent to universal validity.

The assessment of the Kantian solution by Schnädelbach is critical, because the replacement of the predicate 'valid', resp. 'binding', for the predicate 'true' has some undesirable consequences, as it leads to a change in the logical form of the predicate 'true'. It becomes a three-argument predicate P(x,y,z), which can be interpreted as: Sentence (judgment) x is binding on person y in conditions z. According to the philosopher, this logical revision loses an important intuition that lies at the base of the notion of truth. Schnädelbach defends the intuition that 'truth' is a single-argument predicate, drawing on the universal language expression 'Sentence "p" is true', which does not introduce any additional arguments. However, Schnädelbach pays an exorbitant price for his solution. Emphasizing the single-argument nature of the predicate 'true', he loses sight of the prescriptive nature of truth – it is simply a property. From no property, even the most valued one, there follows a norm ordering that such a value should be realized, at most an inclination to do that. The expression of prescriptivism calls for the use, which Schnädelbach rejects, of a two- or three-argument predicate, depending on interpretation. I believe that Schnädelbach's critique should be recognized – indeed, the transfer to truthfulness of prescriptivism appropriate for morality, i.e. prescriptivism understood as significance and validity, is inappropriate, but the proposed solution (the single-argumentness of the predicate 'truth') does away the entire prescriptivism in one go – both ethical and epistemological.

It should also be remembered, which has already been pointed out, that single-argument interpretation of the predicate 'true' automatically falls victim to Tarski's critique. Donald Davidson has suggested an interesting way to circumvent Tarski's objections without abandoning the single-argument interpretation of the predicate 'true'. He recognizes truth as a primeval and indefinable notion. Thus, no complication of the logical structure with a view to connection with other variables is then necessary. This solution, however, necessitates the proposal of another principle of constructing a universal set of true sentences – a principle that would not have to depend on a (an impossible) definition of truth. To Davidson, the word 'truth' is extended by a collection of fuzzy sets of true sentences in various idiolects. In turn, such sets are only defined by a certain indication; namely, as a majority of sentences accepted by a given person (provided such person is rational). However, the Davidsonian principle of construction of a set of true sentences is not a sufficient interpretation of the notion of truth. Undoubtedly, the principle is understandable; it provides a certain feeling that we know what we are talking about when using the word 'truth'. However, it is not an effective

[3] H. Schnädelbach, *Rationalität und normativität*, [in:] H. Schnädelbach, *Zur Rehabilitierung des "animal rationale"*, Berlin 1992.

concept, either in the epistemological sense (it does not implicate the criteria of truthfulness) or in the semantic sense (no criteria for the correct use of the word 'truth' arise from it).

Therefore, we seem to face the following possibilities:

1. The property 'true' expresses the single-argument predicate: P(x). In this view, truth is a property of a sentence. This interpretation is susceptible to Tarski's argumentation. It cannot be defined for colloquial language and so it is of no special importance to epistemology, although it obtains a precise meaning as a semantic notion.

2. The property 'true' expresses the two-argument predicate: P(x,y). In this view, truth is a norm ordering cognitive states. This is the interpretation I intend to defend in these deliberations.

3, The property 'true' expresses the three-argument predicate: P(x,y,z). In this view, truth is a cognitive value to a certain subject in a certain context. This perspective reflects the prescriptive power of truth, but makes the remaining components of this notion incomprehensible or redundant. (It is for this reason that Schnädelbach rejects this interpretation.)

Below I defend the model of truthfulness as a two-argument predicate whose arguments are certain moments of judgment. I define a judgment after Husserl as a certain relation between thought and its object. Truth characterizes the relation between some constituents of this relation, and so it is an internal property of a judgment. Before undertaking the argumentation regarding the phenomenological interpretation of the truthfulness of judgments, I would like to justify the transition from considering the truth of sentences to considering the truth of judgments, which I do in the following paragraph.

3. Transition from truthfulness as a property of a sentence to truthfulness of a judgment

Today's popular deflationary concepts of truth, partly inspired by Tarski's achievement, hold that the predicate 'true' does not possess any content or that its content is created as a result of an ordinary replacement of the name of the sentence '*p*' for sentence *p*. Seemingly, this removal of quotation marks is an operation identical to the so-called Convention T: '*p*' is true if and only if *p*.

However, the issue seems more complicated. When we establish that there occurs the relationship: '*p*' is true if and only if *p*, and we establish that *p*, then we obtain the right to regard sentence '*p*' as true. Therefore we may determine that *p* does not refer to '*p*', but to *p*, which is what the sentence '*p*' is about. Acknowledging that the predicate truth can be applied to a certain sentence depends on the satisfaction and recognition of a judgment whose expression is the sentence 'p'. In other words, '*p*' without quotation marks is not the same sentence, but a judgment. This is why it is possible to replace 'Snow is white' for 'Śnieg jest biały' on the left-hand side of the sentence ' "Śnieg jest biały" if and only if snow is white'. There exists a certain formula: '[ ] is true if and only if snow is white', into which expressions in various languages can be inserted. The criterion for the

possibility of this replacement is the content of a certain judgment – namely, that snow is white.

The objection made above is connected with remarks by Adam Nowaczyk who pondered the question: Did Tarski really formulate a semantic definition of truth? Commenting on the critique formulated by Putnam and Echemedy, Nowaczyk observes that Tarski's concept assumes certain semantic intuitions, but it does not express them directly. Tarski aimed at providing a purely morphological definition of truth, without assuming any other semantic notions. Two notions of language need to be distinguished here: 1) only morphology and meanings, with no pragmatic elements; 2) morphology and semantics plus pragmatics (Ajdukiewicz-style). When talking about the language of class theory, he meant the first one (language with no pragmatics connected with the normal practice of mathematics). The other view of language is related to having language models. In his work on truth, Tarski did not use the notion, but implicitly it is present in the semantic definition of truth, because Tarski makes use of a certain intuitive model, when he translates expressions from the object language of the class theory into the metalanguage understandable by the reader[4] (he refers to objects of the syntactic class [z] as sentences 'z', creating a metalanguage class ['z'], etc.). In fact, it is not necessary at all to translate expression z into expression 'z', understood as expression 'sentence' (expressing the notion SENTENCE, or meaning sentence) in the metalanguage. The metalanguage could be entirely arbitrary in the sense that it would determine unintuitive ranges of names in the metalanguage; the names would only have to meet the condition of coherence with the syntactic rules of the source language. Actually, however, we use translations, which will be useful in a way – language models and pragmatics lie at the basis of such decisions.

'In the model theory semantics, all notional constructions are based on the notion of a (formalized) language model, i.e. (in the simplest case) the organized pair M = <U, D>, where D is any relation between language expressions and objects that is in line with the principle of categorial conformity (i.e. correspondence between the syntactic category of the expression and the ontological category of the object defined relatively to the scope of U). Each such relation is designated as *denotation* and maintained that it is a semantic relationship.'[5]

The conclusions from Nowaczyk's analysis are as follows: (1) With Tarski, there is no reduction to morphology – object notions are still used; (2) 'The defining of the notion of a true sentence in a given language requires a reference to the denotation of all syntactically simple terms of such a language.'[6] Tarski refers to the intuitive interpretation of the denotation.

When formulating the definition for the language of calculus of classes, Tarski simultaneously negates the ability to capture the universal definitional property of

[4] A. Nowaczyk, *Semantyczna czy asemantyczna*, [in:] J. Hartman (ed.), *Filozofia i logika. W stronę Jana Woleńskiego*, Kraków 2000, p. 301.

[5] According to Adam Nowaczyk, there is an analogy here with the situation in which we compare the formal notion of marriage (a certain set of ordered pairs) with the ordinary sense of marriage. The relation between them consists in a generalization resp. idealization. *Cf. ibidem*, p. 302.

[6] *Ibidem*, p. 303.

truth, which property would be responsible for the ability to implement Convention T in various languages. However, Nowaczyk's analysis brings in a certain solution to the problem – even if there is no such property, there is a certain theory binding various semantics – namely the general principles of the model theory semantics.

In the same spirit, Tarski's position is analysed by Ernst Tugendhat. This philosopher is, however, more radical in relativizing Tarski's result. Above all, he questions the privileged function of the definition of truth in seeking an explanation of the essence of truth. He thinks that there is no single definition of a 'true statement.'[7] Furthermore, the definition of a true statement, resp. sentence, refers – according to Tugendhat – to 'the truthfulness of a judgment.'[8] Tugendhat advances an argument similar to that which I formulated against deflactionism at the beginning of this section. He claims that the statement ' "x" is a true sentence if and only if "p" is true' is based on an equivocation – first, it talks about a sentence, then about a judgment. According to Tugendhat, the difficulty is clearly seen when we turn around Tarski's convention and say: 'p = "p" is true'. In this way, the truthfulness of judgment p is made dependent on the truthfulness of 'p'. However, it has to be pointed out that: '...if we reflect on the actual status at the base [of this equation], the meaning of the word 'true' that is present in it acquires a sense which cannot be seen in the equation itself.'[9] The relation referred to above is as self-evident, as it is elusive. A judgment can be expressed by means of many sentences and it is impossible to limit such a set, and so a full understanding of the content of a judgment by only designating a certain set of sentences, without indicating the sense determining the set, is impossible.

However, the formula: 'p = "p" is true' reflects still another property of key importance to our present deliberations about the prescriptivism of truth. Namely, there exists a close connection between the truth of a judgment and the truth of sentences – in each judgment there is a potential reference to language, or perhaps even stronger: only thanks to a connection with language can a judgment reveal a property of truth. It does not mean, however, that the truth of a judgment boils down to the truthfulness of sentences, and this is because of the connection between a judgment and an unspecified multiplicity of sentences in a given language.

Ultimately, according to Tugendhat, the purely formal definition of truth given by Tarski is, from the philosophical point of view, trivial, but the scheme it is based on ('x is a true sentence if and only if p'), is not trivial and 'leads to a legitimate definition of the truthfulness of a sentence – a definition that subordinates a judgment, resp. the captured state of things, expressed by a sentence to that sentence. The correspondence established here should, however, be distinguished from the adequacy that regards the relation of the captured state of things, a judgment, to the very thing.'[10]

[7] E. Tugendhat, *Tarskiego definicja prawdy I jej miejsce w historii problemu prawdy w pozytywizmie logicznym*, [in:] E. Tugendhat, *Bycie, prawda. Rozprawy filozoficzne*, tłum. J. Sidorek, Warszawa 1999, p. 169.

[8] *Ibidem*, p. 172.

[9] *Ibidem*, p. 174.

[10] *Ibidem*, p. 177.

Thus, truth concerns a judgment and a judgment is a relation between a thought and a certain thing, which is formed thanks to a cognitive interpretation of a certain fragment of reality. A reference to the very reality is a different matter – the real occurrence or nonoccurrence of states of things that a judgment refers to is only of significance to a certain (although a very important) class of sentences, which in particular includes the sentences of science. Within science it is enough to talk about the truthfulness of sentences, which practically becomes indistinguishable from the type and degree of justification of such sentences. However, within the entire scope subject to human judgment, it is important that judgments remain in the relation of truth to that which they concern, regardless of whether they are about existing or nonexisting objects, whether they are empirically cognizable or not, whether they are independent of or entirely dependent on the perceiving subjects or even forming part of the very cognitive processes. For this broad sphere encompassing all thoughts and their pretensions to truthfulness we should explain the property of truthfulness and the standard of truthfulness that regulates the very process of forming such judgments. In my earlier remarks, I outlined the negative basis for further reasoning, stating that attempts at defining truth as a property of sentences assume certain intuitions as to the nature of judgments. Now, I would like to focus on the positive determination of the truth of judgments, in particular the normative nature of truth in relation to judgments.

4. Phenomenological interpretation of the truthfulness of judgments

Truth is a cognitive norm. When uttering true sentences on any grounds, we simultaneously undertake to form our judgments (ways of referring thoughts to their objects) in the way that is implicitly contained in true sentences. This does not mean that we know or that we can say that implicitly present norm. As we will see further on, the formulation of such a norm is not a trivial task, regardless of how often we comply with the norm every day. I will try to show – principally, by referring to Husserl's analyses in *Logische Untersuchungen*, *Formale und transcendentale Logik* and *Erfahrung und Urteil* – that truth can function as a norm for judgments in a more literal and stricter sense.

The task is to prove, by a certain analogy with Tarski's efforts in respect of the truthfulness of sentences, a certain general property of judgments, which can be a carrier of the properties of truth and an implementator of the norm of truthfulness. This property is, to a certain extent, visible in the very unity of a judgment understood not as an act of judgment, but as the meaning of a sentence. Judgments contain notions which Frege proposed to interpret as functions, i.e. objects containing an unsaturated element. This element makes a judgment preserve its fundamental identity in many acts of its shaping and formulation. Recently, John Searle returned to the idea in order to explain the unity of judgments.[11] The unity of a judgment arises from the unsaturation of a conceptual function. This function is satisfied by a certain objects posited by certain set of acts of judgment. The

[11] J. Searle, *The Unity of Sentence*, [in:] J. Searle, *Philosophy in a New Century*, Cambridge 2008, p. 181–196.

set of satisfying objects can be interpreted statically (without any assumptions regarding the temporal aspects of acts of judgment or time in general) or dynamically. The very notion of unsaturation does not implicate this dynamics, but only its static fabric. We do not know yet why or how the state of unsaturation of a conceptual function (and hence judgment) gets restored after each satisfaction, i.e. after each statement of a given predicate about some thing. It seems that our efforts embodied in judgments have a certain teleological aspect – pursuance of truth has always been a way of intuitively expressing this teleology. However, if we want to understand truth not only as an ideal we strive for (the general teleology of cognition), but as an internal norm driving cognitive acts (the internal teleology of each judgment), then we cannot refer back to the notion of truth. The question is now: At what form does a judgment aim that such dynamics can be regulated normatively, i.e. that a given judgment can be perfected because of truth? It is not enough to indicate a certain property of judgments: truthfulness, rationality, clarity, etc., and to say that the property should be realized. Obligations do not arise automatically from valued properties – just like cognitive obligations do not arise automatically from valued cognitive properties.

Edmund Husserl proposed an interpretation of judgments in a language accepted by him generally for all acts of consciousness. Of fundamental importance to this interpretation is the notion of the pair: <intention, fulfilment>. Using this, in *Logische Untersuchungen*, Husserl distinguished four meanings of the word 'truth'. These were extensively discussed in Polish literature by Andrzej Półtawski[12] (who took account, among other things, of the analyses of Ernst Tugendhat in *Der Wahrheitsbegriff bei Husserl und Heidegger*), and so I will only focus on the elements of alethiology presented in *Logische Untersuchungen* which lead to the explanation of the teleological dynamics of judgments, and so they contribute to the explanation of the prescriptivism of truth.

According to Husserl's most general definition of truth, truth is a total coincidence of intentions and fulfilment (*die volle Übereinstimmung von Gemeintem und Gegebenem*). This coincidence has to be given in experience in the form of a certain intuition (*Evidenz, Intuition*). The most important thing is that such intuition – the intuition of truth – is at the same time corrected in the light of the very experience – i.e. in the light of other truth intuitions. Phenomenological analysis is tasked with bringing about the clarity of certain types of coincidence of act intentions and appropriate fulfilments.

At the level of sensory experience, the process takes place in accordance with the perception norms. The dynamics of operation of such norms is shown by Husserl on the example of kinesthesia. As Andrzej Półtawski writes: 'The general dynamics of intention and fulfilment corresponds there to the dynamics of indeterminacy and determinacy, and the prerequisite for a closer determination of the content of experience is "the process of inclusion in the remaining knowledge, which becomes habitual" [...] Thus, this is about the process of constituting sense, which encompasses "that which is new", when a closer determination brings

[12] A. Półtawski, *Aletejologia Edmunda Husserla*, [in:] A. Półtawski, *Realizm fenomenologii. Husserl-Ingarden-Stein-Wojtyła*, Toruń 2001.

a new element of sense.'[13] In this way, the sense builds in perception. However, a kinesthetic experience includes, on the one hand, unfulfilled elements, and on the other – a presentation of an object goes beyond what is included in the intention. 'This which here, in the course of kinesthetic perception, goes beyond the very conjecture (*Vorziechnung*), beyond that which is expected, does not characterize simply as falsehood, but as a closer determination.'[14] That last one as such has a nature of fulfilment. The dynamics of kinesthetic experiences refers one both to the very object and to the movement of the subject (kinesthesis as an experienced form, resp. a time-and-space scenario organizing the surroundings of an object). On the one hand, an object is given only to the extent to which it accompanies a certain bodily movement, and on the other hand – it is not simply a correlate of such movement. The movement of a body is the most important factor – it is this movement that triggers a constant spreading of intentions and fulfilments, and, at the same time, constitutes the cognitive movement of a sensory judgment from one state to another.

In order to explain truth as the prescriptivism of judgment we now need a transition from rudimentary judgments for which we have perception norms (based on body movement and explainable at the naturalistic level) to a whole class of judgments whose truth and other cognitive values can be stated. What is the relation of the regularities operating at the sensory level with the general theory of judgment?

Husserl's thinking runs in two directions here: firstly, he examines the genesis of the form of judgment. This topic is his focus of attention in *Formale und transcendentale Logik.* Secondly, he studies the content-related genesis of judgment (principally in *Erfahrung und Urteil*). Both of these trains of inquiry are to establish the bases for logic and the theory of perception by outlining the possibility of reconstructing the source experience – the original obviousness that lies at the basis of all types of judgments. The most important instrument of analysis in both cases is intentional analysis. It assumes that in each perception there takes place intentional modification, which – by transforming interpretations characterized by original obviousness – generates areas of senses expressed in possible act variants.[15] Husserl's key assumption is that such modifications can be traced in an appropriate insight. The subject of consciousness, and so the phenomenologist as well, who 'imitatively understands such a form of consciousness', can strive towards the

[13] *Ibidem*, p. 152–153.

[14] *Ibidem*.

[15] 'But intentional modifications have, quite universally, the *intrinsic* property of *pointing back* to something unmodified. The modified manner of givenness, when, so to speak, we interrogate it, tells us itself that it is a modification *of* original manner of givenness, to which it points. For the subject of the consciousness (and consequently for everyone who places himself in that subject's place and understands that mode of consciousness in following him), that makes it possible, starting from the particular non-original manner of givenness to strive toward the original one and perhaps make it explicitly present to himself in fantasy at the same time making the object-sens "*clear*" to himself. The fulfilling clarification takes place with the transition to a synthesis in which the object of non-original mode of consciousness becomes given eithe as the same as the object of consciousness in the mode of "experience" (the mode of "it itself") or else as the same object "clarified" – that is to say as it "would" be itself-given in a "possible experience". E. Husserl, *Formal and Transcendental Logic*, transl. D. Cairns, The Hague 1969, p. 314.

source form of experience. This source form has, in line with the general idea of intentionality, a certain objective sense. The explanation consists in revealing the source objective sense. This happens in a synthetic transition from something that is given in a non-sourceful manner (as represented, mediated) to a form in which it presents itself 'as it itself' – namely, in a way in which 'it would be given itself in a possible experience'. The possible experience is the most important notion here. Intentional analysis does not seek a new justifying experience, but hopes to reveal the source form in the experience we have.

The same strategy, as I have already outlined, applies to thinking, and in particular to the logical form of thinking. However, here the explanation requires the assumption of the unity of thought and language. It is not a trivial unity and it is not given in a natural way. It is rather the object of one of the most interesting questions asked by Husserl in *Formale und transzendentale Logik*, namely: 'What universal essential character must a proccess of consciousness have in order to be capable of taking on a significant function?'[16] What is needed here is a typology of thoughts, relationships of consciousness to various types of objects, and, above all, an analysis of constructing pure forms in mental operations. It is on the basis of pure forms that there arises the capability to outline relationships of judgment of all types (belonging to direct experience and to organized scientific experience). These pure forms emerge as 'pure possibilities of a cognitive life.'[17] If they are captured in their essential features, they will become a source of clarity – other than that proposed by Carthesius, as he drew on a certain property of the mind, and not on the ideal domain of pure forms. It is only from the perspective of clarity understood in this way that it is possible to undertake deliberations on truthfulness. "Truth and falsity are predicates that can *belong only to judgement that is distinct* or can be made distinct, *one that can be performed actually and properly*. Logic has never made clear to itself that this concept of the judgement is at the basis of the old thesis that truth and falsity are [...] the predicates of judgements."[18] Thus, a theory of judgment in general, whose elaboration would be a theory of scientific judgments, is called for.

Obviously, formal sciences have worked out their own ways of substantiating their sentences – by proving them; a philosophical project may not contribute any bigger degree of substantiation or certainty. However, it may supplement formal methods with a parallel train of reasoning, which will not show that something is true, but rather that which is true. One can reduce every judgment down to the very object it talks about, and use the grasp of the object to assess a positive or negative adequacy of judgment in question.[19] In the first case, a judgment

[16] *Ibidem*, p. 25–26.
[17] *Ibidem*, p. 28.
[18] *Ibidem*, p. 66.
[19] This fragment will best show the strategy of Husserl's explanation. When considering the law of noncontradiction in the second part of the work, he writes as follows: 'The law of contradiction expresses the general impossibility of contradictory judgements being true (or false) together. If we ask for the evidence in which it is grounded, we see that this impossibility involves the following: If a judgement can be brought to an adequation in a positive material evidence, then, a priori, its contradictory opposition not only is excluded as a judgement but also cannot

is self-evidently true – thanks to a satisfying-and-confirming agreement with the categorial objectivity which it regards as an opinion and which now offers itself as self-given; in the other case, it is self-evidently false – namely, because simultaneously with a partial fulfilment of an opinion (a surmised categorial objectivity as such) categorial objectivity revealed itself as a self-given, and it negates the entire opinion and, by necessity, it 'does away' with it. In other words, for each thought it is possible to find a certain basic form, which presents itself with self-obviousness and combines with such thought via a constitutional sequence of mental forms.

The focusing and intentional function of an object makes theory of truth refer to ontology. In the perspective of truth there lie final objects whose status needs to be described in an appropriate ontology. This obligation of ontology to describe objects that lie within the horizon of experience as target objects – ultimate fulfilments of intentions – means that all ontology also becomes a theory of truth.[20] According to Tugendhat, the basic notions of object domains are regarded by Husserl as counterparts of the basic forms of experience; they cannot be separated from types of self-presentation of objects. This is why Husserl's ontology, which is not based on differences between the objects themselves, but on the types of experience, is in fact a theory of truth.[21]

Deliberations on the logical form can be expected to negatively determine the limits of truth in logic – to show forms that in no conditions find an appropriately fulfilling intuition (*Evidenz*). However, a step further should be taken and a correspondence between the justification lying in the object and the justification lying in the form of judgment should be shown. According to Husserl, it is necessary to transfer interest to the very objective importance, resp. fulfilling intuition [*Evidenz*]. It is important to distinguish the more primeval forms of judgments that in themselves carry results of previous judgments from the very objective obviousness [*Evidenz*].

The studies in *Erfahrung und Urteil* are, to a certain extent, parallel to those in *Formale und transcendentale Logik*. In the latter work, Husserl was looking for a general theory of the logical form of judgment, while in *Erfahrung und Urteil*, he focused on the objective obviousness that lies at the basis of judgments. Husserl begins his deliberations by distancing himself both from Hume's and Brentano's concept of judgment. To the first, judgment is a certain given of consciousness,

be brought to such an adequation, and vice versa. That is not yet to say that, without exception, every judgement can be brought to adequation. But just this is involved i the law of excluded middle, o the subjective side, its evidential correlate" *Cf*. *ibidem*, p. 193.

[20] In the very history of mathematics, Husserl saw a certain movement leading towards an increasingly richer expansion of the objective sphere to which mathematical theorems refer. The best example of this is the extension of the notion of geometry with non-Euclidean spaces (Riemann, Łobaczewski).

[21] A. Półtawski, *Aletejologia Edmunda Husserla*, p. 159. This motif also appears in *Formale und transcendentale Logik*. Husserl looks for basic forms of experience to consolidate logical operations in them. The task brings basic object intuitions closer to formal sciences. Ultimately, mathematics is to Husserl a formal ontology. According to Husserl, from the notion of an object in general derive such notions as: multiplicity, number, relation, series, combination, entity and part, etc. In turn, Husserl calls logic apophantics (a science about forms of judgment) and an apriori-and-formal theory of object.

to the other – a pure activity of I. Both concepts are regarded by Husserl as false. To him, a judgment is always a certain reference of thought to an object, and it always contains some doxastic element, i.e. an embedded belief about an object. From the perspective of an object, Husserl's thesis says that even a passive contitution of a datum standing out of the background of the unity in immanent passivity has a passive *doxa*.'[22] This relation possesses an appropriate dynamics – doxastic states are never stable – '... this domain of doxa is the domain of the fluid. A passively given unity of identity is not yet one which is grasped as such and retained as an objective identity. On the contrary, this apprehension, e.g., the perceptive contemplation of the pre-given sensuous substrate, is already an activity, a cognitive performance at the lowest level.'[23]

In the field of philosophy, an analytically natural tendency would be to interpret such dynamics of judgment by means of linguistic categories – as a semantic-and-pragmatic order of sentences and statements. Husserl does not deny that a reference to language is key to the building of a certain field of possibilities that constitutes a basis for the dynamics of judgments. Yet the issue of linking judgments with language is deliberately treated as marginal by Husserl.[24] The founder of phenomenology is more interested in cognition (formulation of judgments) as a certain activity. In paragraph 48 of *Erfahrung und Urteil*, Husserl even suggests that cognitive activity belongs to the same broad category as practical activity. The following feature is pointed out by Husserl as a difference: "The predicative achievment of cognition has been characterized as an action, and this is justified in that the general structures of all action are also capable of being exhibited in this cognition, though in other respects cognition is still to be distinguished from action in the ordinary sense of that term. We prefer to think of action as an external doing a bringing-out of certain objects (things) as self-giving from other self-giving objects. In cognitive activity, new objectives are indeed so pre-constituted, but this production has an entirely different sense from that of the production of things from things; and what is here important above all – this production of categorial objectives in cognitive action is not the final goal of this action. All cognitive activity is ultimately referred to the substrates of the judgement.'[25] The striving for cognition is analogous to the striving for something out of desire. In both cases it is about a certain possession. As Husserl says: 'in the progress of the action the striving fulfils itself more and more from the initial mere intention to realization.'[26]

However, if truth were only to be a total fulfilment by the objective sense, and thus a full possession of an object, then truth would still be only a certain ideal – a fanciful possibility of skipping – as Husserl wrote in the fragment already quoted – a certain number of cognitive steps and the imagining of a total fulfilment. In

[22] *Ibidem*, p. 61.
[23] *Ibidem*, p. 59.
[24] 'The predicative operations will be examined purely as they phenomenally present themeselves in lived experience, apart from all these connections, nely as subjective entities'. E. Husserl*, Experience and Judgement Investigation in a Genealogy of Logic,,* transl. J.S. Churchill, Evanston 193, p. 199.
[25] *Ibidem*, p. 200.
[26] *Ibidem*, p. 201.

order to extract a realistic basis of truth, we need to show this process of cognitive appropriation of an object in a dynamic relation with its previous, imperfectly fulfilled forms. A judgment has indeed to be presented as an activity characterized by an internal connection of its individual phases.

Thus, time is constitutive in relation to the act of judgment. According to Husserl's thought that is of key importance to these deliberations, in an act of judgment there occurs a peculiar symmetry between retention and protention. As known from Husserl's analyses of time, the very protention-retention structure is not a full development of the time experienced – however, it reflects the dynamics of the present moment, going from a certain field of possibilities in protention to the indeterminacy of content given in the passing. The tracing 'back' of a modification of a judgment is also an exploration of a protentive possibility. An increasingly fuller presentation of an object is simultanously a protention-based disclosure of constantly new content relating to a given object and the constitution of a constantly fuller presentation of the object in a series of retentive views. Conversely, the tracing of intentional modifications is at the same time a disclosure of a certain way of a possible, and so a prospective, experience.[27] The goal set by Husserl – to disclose the source objectivity – is also the pursuance of truth as a full presentation of an object. Truth is a standard of judgment that demands its development towards a full presentation of appropriate objects. At the same time, reflection leads to a gradually disclosed source of each judgment – the object (also its parts and moments) constituting the ultimate establishment of a given judgment.

The retentive-and-protentive symmetry of an act of judgment is closely related to the dialectic of intention and fulfilment. A field of possibilities given in protention corresponds to an intention, while the retentive depth of a given act corresponds to fulfilment. In the case of a judgment, both the first and the latter are connected with language. The protentive field of possibilities is given as a field of meanings, and the retentive field of fulfilments is given as a certain structure of notions – co-stated in a certain order about a given object.[28] Developments of a statement can theoretically extend into infinity, as each use of a notion refers to successive notions. In practice, the content of an experience is limited by a certain convention, which requires that the optimum subject of presentation (a set of expressions that come together in a full presentation of a given object) be placed at a certain level of notional presentation.

We have reached the key moment of our reasoning: Not only does the field of possibilities of determination actualize (fulfil) in a retentive series of notionally organized determinations, but also the opposite relationship occurs – a given state of judgment in the form of a retentive series of determinations is another expression of that which is protentively present as a field of possibilities of determination. I believe that Husserl's description enables us to understand that we go beyond

[27] Because both retentive and protentive developments still belong to a present experience, here we do not deal with the past or the future in their strict sense, although in a way the retentive-protentive symmetry indicates an open horizon of temporality.

[28] *Cf.* L. Eley, *Nachwort*, [in:] E. Husserl, *Erfahrung und Urteil. Untersuchung zur Genealogie der Logik*, p. 515–516.

a simple symmetry, namely towards a true dynamics of an act of judgment directed as a cognitive act towards a full presentation of an object. My argument is as follows: the very mechanism of determination that makes it to occur at a level of world categorization appropriate for a given cognitive situation is also a source of opening of a field of possibilities – that which is left in the indeterminacy of retention is also the openness of protention. Conversely, a field of protentive possibilities indicates the lack of a sufficient determination of an object in the retentive depth of an experience. This lack is the source of a cognitive need, which Husserl puts in one group with other desires. Here, we have to do with a peculiar experience of a partial possession of an object and a partial satisfaction of a need. Striving for a full presentation of an object is not based on an enigmatic sense of a lack, but on a specific development of a linguistic (categorial) indeterminacy into a field of protentive possibilities.

7. In conclusion

Considering the modern criticism of the notion of truth, can it be regarded as an immanent and, at the same time, effective standard that drives our cognitive acts? Is truth not simply a justification? In *Oneself as another*, Ricoeur observes that prescriptivism is embedded in the very nature of acting[29] – each act can be improved, the acting party can be advised, etc. However, such an analogy between the prescriptivism of cognitive acts and the general prescriptivism of acting is not sufficient – truth and all cognitive standards would then boil down to perfectionist standards (how to best perform a given act), but one would lose sight of the attitude to the world – it would, at most, be taken account of implicitly, without ever mentioning what it is. The vertical axis of the connection with reality and the horizontal axis of perfection has to cross in the explanation of the notion of truth. We need a realistic consolidation of the notion of truth. The very naturalistic relativization to a cognitive system is not enough, as it is not sufficient that a system operates in a c c o r d a n c e with the truth – it has to operate in a c c o r d a n c e w i t h t h e t r u t h, b e c a u s e o f t h e t r u t h. I think that in spite of the fact that over 100 years have passed since the publication of *Logische Untesuchungen*, Husserl's phenomenology is still a source of inspiration to those who try to find a solution to this problem. I also think that the phenomenologically interpreted dynamics of intention and fulfilment is helpful in understanding the prescriptive and, at the same time, realistically based function of truth. It is, however, important to apply the notion of 'truth' above all to judgments and acts of judgment, and only secondarily to statements and sentences. Each phase in the constitution of a judgment implicates the imperative that an increasingly fuller presentation of an object should be strived for. This striving is reflected in a constant self-turning (in a field that can only exist thanks

[29] P. Ricoeur, *Oneself as another*, transl. K. Blamey, Chicago 1992, p. 169. Later, by referring to the notion of immanent goods of Alister McIntyre, Ricoueur stresses even more the importance of the internal teleology of action, which, combined with the idea of an immanent good, generates the notion of a standard of action; *cf. ibidem,* p. 176–177.

to language) of a retentive indeterminacy into a protentive field of possibilities. The very truth operates opposite to such striving, turning a field of possibilities into a gradually consolidating knowledge.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

MATEUSZ KOTOWSKI
University of Wrocław

Toulmin's Evolutionary Theory of Scientific Development: The Study of a Failure

Abstract

This paper discusses main issues concerning Toulmin's theory of evolutionary development of scientific knowledge. The reconstruction of Toulmin's assumptions, which underlie the theory, and the presentation of the theory itself, is followed by the analysis of the difficulties of the conception in general. The author tries to show that these difficulties consist primarily in the discrepancy between what Toulmin aimed at and what he actually achieved. The problem is that although incorporating the means of Darwin's theory may itself be – in the context of describing the growth of scientific knowledge – justifiable, it is not consistent with Toulmin's demand for the rationality of science. The purpose of creating the evolutionary theory was, as Toulmin claimed, to give an account of the development of scientific knowledge which would be non-relativistic and rationalistic. However, this purpose is, as the author argues, not achievable on the grounds of Toulmin's accounts of rationality.

Ever since Charles Darwin published *On the Origin of Species,* more and more philosophers have found it tempting to employ his ideas and conceptual tools in their attempts to explain the problem of intellectual development of humankind. The tremendous explanatory power of Darwin's theory, when employed within its proper scope of application, has allured philosophers like Thomas H. Huxley, Ernst Mach or, more recently, Karl R. Popper to transplant some of its concepts and tools from the domain of the development of biological life to the domain of the development of scientific knowledge. However, such attempts had barely achieved the explanatory success of the original. This fact casts some doubts on the very possibility of founding a viable theory of the development of knowledge on Darwin's theory (or, for that matter, any other theory of biological evolution). However, answering the question of whether a theory of scientific development founded on the theory of evolution – if possible at all – would require the scope of analyses far extending the limitations of one article. Additionally, conclusions of such extended analyses would most likely offer us not an exhaustive answer

to this question, but rather many specific answers to the question of how such theory is impossible. Therefore, in this article I would like to confine myself only to the one case study exemplifying such attempts – the attempt undertaken by Stephen Toulmin in his *Human Understanding*. This particular proposition deserves attention and recalling for at least two reasons. The first is that the British philosopher explicitly adopted and employed the conceptual apparatus of Darwin's theory to describe the development of scientific knowledge. The second is that he spectacularly failed in doing so. By focusing on Toulmin's attempt I want to uncover some of the reasons of this failure. In what follows, I will first discuss the key notions of Toulmin's theory and then move on to the critical analysis of some of them (therefore, those well acquainted with Toulmin's conception may want to skip the first part of the article and go straight to the second).

I. Toulmin's theory

I.1 The road to evolution

Because Toulmin's theory owes much to his general criticism of contemporary epistemology it would be best to start with a brief recollection of the key assumptions and ideas which led him to the evolutionary approach to science. For Toulmin, the most fundamental problem of contemporary epistemology was its attachment to the view which equals rationality with reasoning based on a logical-mathematical model. This view traces back to Plato's dialogues and Euclid's *Elements* and it was reinforced by the seventeenth century rationalism, which shaped the tradition that has shaped our basic epistemological problems since then. The inability of contemporary theories of knowledge to account for the problem of change and continuity is a testimony to their roots in this tradition. This very problem divides contemporary theorists of knowledge into two camps – the absolutists; with the followers of Gottlob Frege on one hand, and the relativists; with philosophers such as Robin G. Collingwood or Thomas S. Kuhn on the other. The main difference between these camps is that while absolutists claim that the whole discrepancy in seeing the world – whether among cultures or historical periods – is nothing more than the effect of errors of cognition, the relativists go to the other extreme and claim that every concept has value only in relation to its original context. Absolutists, under the surface of discrepant accounts of physical reality in which the history of science is abundant, see only atemporal principles reflecting clear and ideal forms of concepts. Relativists see this as a reflection of the fact that there are no universal criteria or principles which would allow us to go from one conceptual system to the other, what leads them to the conclusion that changes in the systems of knowledge are either not entirely conscious or not entirely rational. In short, the absolutists completely ignore the problem of scientific change and the relativists are overestimating it, but neither of them managed to resolve it in a satisfying way. This inability is due to the fact that, while seemingly extremely different, both these approaches are founded on the same view, according to which scientific knowledge should be treated as an ordered structure or system which is describable in logical terms, which in turn, implies that if it is to be judged as rational, the development of knowledge should be given universal intellectual

foundations. As Toulmin puts it, by identifying rationality with logicalness, both camps "subscribe to the philosophical cult of systematicity."[1]

The postulate to transform the conception of rationality together with the deep conviction about the rationality of the process of scientific change are the very things which pushed the British philosopher towards making an attempt at designing a theory for the development of scientific knowledge which would be founded on the theory of evolution and which would offer a third way between absolutism and relativism. It is in the analogy to the development of living organisms that the continuity of the processes which have been shaping our scientific knowledge becomes evident. Additionally, such an approach should offer an account of this process which would allow for admitting to historical relativity without falling into socio-historical relativism.

What interests Toulmin in the theory of evolution is primarily the sole explanatory schema. He explicitly assumes that Darwin's theory is only one of the specific cases of employing a method which could be equally successfully applied in describing phenomena from other domains.[2] Therefore, he focuses primarily on general relations which can be distinguished in historical process between long-term schemata of conceptual changes, everyday activity of users of concepts, and stable conditions determining the preservation of short-term decisions. Assuming that the development of knowledge can be described in a way analogical to the Darwinian description of the development of living organisms, Toulmin claims that scientific concepts should not be treated as logical systems but as populations. They come into being and disappear in response to intellectual conditions of the intellectual environment and are subject to the process of variation and natural selection.

I.2 Concepts as the subject of evolutionary theory

The very notion of science is as general as it is vague and it would be difficult to find a more unambiguous answer than that science is what people who call themselves scientists are employed in doing. Though trivial, this constatation is in fact quite close to the way in which Toulmin conceived science. He claimed that if we want to consider science as a process, we cannot separate the subject of science from the work of scientists. Therefore, we should search for the answer to the question of what, in the development of scientific knowledge, undergoes evolution somewhere on the borders of (the products of) scientific practices and the actual activities of scientists. And what we find is that there are concepts. They are the constituents of the subject and object of science; they are the products of scientific practices, but as such they belong to and are dependent upon those practices.

The very notion of concept is for Toulmin a complex one. He distinguishes three elements in it: language, representation techniques and application procedures.[3] The former two are related to the symbolic aspect of scientific explanation

[1] S.E. Toulmin, *Human Understanding*, Princeton 1972, p. 83.
[2] *Cf*. *ibidem*, p.135.
[3] *Cf*. *ibidem*, p. 161.

which is "the scientific activity that we call 'explaining"'.[4] They include the technical language in which the concepts, laws and generalizations are expressed,the laws described, as well as the products of the explanations. The third element is related to "the recognition of situations to which those symbolic activities are appropriate"[5] and as such it constitutes a reason for the existence of the first two because it determines the possibility and the scope of their application.

Concepts are therefore not merely abstract products of science but they have a concrete existence as the subjects of scientists' research. Therefore, *in abstracto* concepts are never exhaustive of any scientific discipline, but only present it a specific moment in time. As with institutions, we can understand a concept only by analysing how, over time, it allows scientists to achieve the stated goals. This brings us to the question of how the very process of conceptual evolution proceeds and what the criteria of selection are governing it.

I.3 The process of conceptual evolution

The notion of concept is closely related to the notion of a scientific discipline. When we focus on a scientific discipline in a certain historical moment, we are dealing only with temporary products or cross-sections of complex, historically developing enterprises. The basic question which Toulmin encourages us to ask, is: "What makes the later phases of science the 'legitimate heirs' of the earlier?"[6] As an evolving organism, scientific disciplines are developing over time and the continuity of this development becomes visible in that any later phase owes its legitimacy to the fact that it managed to solve at least some of the problems left unsolved by its predecessors. At the same time, the problems are not in any way a stable and invariable element of science but quite the contrary. According to Toulmin they are changeable, but the very process of their change is continuous. Scientific problems form genealogies of difficulties by explaining nature in which later generations follow the earlier. If, however, it is problems that constitute the basis for determining the continuity of scientific disciplines, the question is now about the very process in which they are growing one from another and, first of all, how do they come into being in the first place. As Toulmin points out, "the problems of science have never been determined by the nature of the world alone, but have arisen always from the fact that, in the field concerned, our ideas about the world are at variance either with Nature or with one another".[7] This being-at-variance is the source of scientific problems and the large-scale and long-term changes in science are the consequence of the gradual accumulation of smaller modifications. Those modifications appear always as responses to some specific problems and are preserved as a result of giving a successful answer to them. It is not the case, therefore, that the problems in science result simply from comparing our claims about nature with observations. In reality, they result from discrepancies between explanatory ideals, i.e. ambitions of a given discipline

[4] *Ibidem*.
[5] *Ibidem*.
[6] *Ibidem*, p. 146.
[7] *Ibidem*, p. 150.

concerning the explanation of phenomena, and its actual capacities to fulfill those ambitions. This claim Toulmin summarises in one formula: "Scientific Problems = Explanatory Ideas – Current Capacities".[8]

Within a given discipline, researchers aim to solve specific problems they have with explaining nature. However, their aims are never solely consequences of empirical observations nor the introduction of new mathematical models, but rather of adopting certain views about the problem's situation. For example, what, according to Toulmin, was the real achievement of Rutherford and Thomson – regarded as the inventors of atomic physics – was the creation of a new intellectual ideal of science.[9] In an effort of imagination, they managed to break away from settled conceptions and sketch a new problem situation. By doing so they determined the future character of the discipline, because only the adoption of a given intellectual ideal introduces a link between the disciple and the empirical world and, as such, is primary to the empirical investigations. On this account, "The chief explanatory patterns, forms of theory or 'themata' of science were all worked out in advance of any clear recognition of their empirical scope."[10]

The coming into being of a new concept is always preceded by the recognition of a new problem situation and is a consequence of introducing innovative procedures of dealing with difficulties. Concepts develop and evolve when they are applied to solving specific problems. Theories are thereby separated from considerations about truth – propositions of science do not apply to the real object directly. However it is not to their reference to the empirical domain that they owe their validity but rather to the fact that they may be applied to it – scientists are not asking whether a given proposition is true, but how and under what circumstances it may be applied.[11]

And so, as members of a population must constantly prove their value in order to maintain their position, so must concepts. The moment a given concept ceases to deliver what is expected of it in terms of explaining certain phenomena, the evolutionary mechanisms of change start to work. This eventually leads to the substitution of an older conception by a new, better-adapted one. However, particular problem situations differ highly from each other and that is why the very process of change may take different forms. Depending on the context, Toulmin introduces different classifications for the types of conceptual change – which correspond to the types of problems which they are supposed to solve. In order to discuss the actual process of conceptual change, I will focus on the distinction between, what we may call here, routine and extraordinary problem situations.

The first of the distinguished types of situations is defined by the problems which, in a way, impose the adoption of specific solutions,[12] i.e. their recognition determines the choice of the concepts which will serve as their solutions. This choice depends only on scientists' estimations about which, out of the available

[8] *Ibidem*, p.152.
[9] *Cf*. *ibidem*, p. 153.
[10] *Ibidem*, s. 152.
[11] *Cf*. *ibidem*, p. 170.
[12] *Cf*. *ibidem*, pp. 224–225.

conceptual variations, would best fit the problem. The scientists are thus comparing concepts from the perspective of their potential capabilities in explaining a given phenomenon. This is not to say that we can talk here about some universal formal criteria. It is rarely the case that scientists have a ready innovation which would solve the problem completely on the one hand, without it generating any new ones on the other. Moreover, even explicitly accepted at a given point, criteria my point in different ways – for example, when a potential innovation offers a higher degree of simplicity and coherence, but is, at the same time, less accurate. Instead of formal criteria, scientists follow their intuitions which reflect their disciplinary ideals and they always ask themselves whether a loss in light of one criteria is compensated by a gain in light of the others.

Despite lacking universal criteria, the ways of solving the problems of the discussed type are characterized in every case by scientists' collective agreement about the existence of *any* criteria; that is, that they can explain what they would count as the solution of the problem. Yet, there are also cases in which such collective agreement is lacking and then we are dealing with extraordinary problem situations which Toulmin describes as "intrinsically 'cloudy' ".[13] This cloudiness is a result of a lack of consensus among scientists within a given discipline as to what sort of strategy this discipline should employ and "is a direct consequence of the fact that our disciplines are in the course of historical change, *even in their deepest rational strategies*."[14]

To illustrate the nature of such problem situations, let us consider Toulmin's example concerning the dispute which took place in 1910–1911 between Ernst Mach and Max Planck.[15] The dispute concerned the situation in which the then contemporary physics found themselves when the program for this discipline outlined in Newton's *Optics* was at the point of becoming exhausted – and scientists were questioning and arguing about the very foundations of their discipline. In *Physikalische Zeitung,* Planck criticised Mach's sensationalism arguing that the historical development of physics makes it evident that any subjectivist elements, which Mach's theory wants to bring back, are and should be successively eliminated from it. In response, Mach argued that what should be eliminated from physics is metaphysics. Regardless of whose arguments prevailed, the character of this dispute clearly shows that it was exactly about the very disciplinary strategies. In particular, Planck's position deserves notice because he explicitly drew conclusion about the future of his disciplines by analysing its past development. As Toulmin emphasises, he acknowledged that

the new strategies appropriate to the problems of theoretical physics in his own day must make it the 'legitimate heir' of all previous physical investigations; they had, therefore, to be formulated and judged not in formal or abstract terms, but with an eye to the entire historical evolution of physics, and its ideals of 'physical explanation'.[16]

[13] *Ibidem*, p. 232.
[14] *Ibidem.*
[15] See *ibidem*, pp. 232–233.
[16] *Ibidem*, s. 233.

When there is no collective agreement in a discipline about a strategy of explanation, the central disciplinary question is not about what would be the best solution to the problem, but what is the problem which the discipline should seek to solve. In such situations we cannot talk about an established or universal criteria of selection, because these are cases in which all the previously latent disagreements among scientists are coming to the fore and there is no higher instance to which all or most scientists would agree. It is exactly the higher instance that becomes the subject of the dispute.[17] However, what Toulmin emphasises, even in such situations we cannot say that scientists simply give up rational argumentation for persuasion. The core of such disputes – as well as the thing that secures their rationality – is that they "call for appeal, not to the codified rubrics of an established theory, but to broader arguments involving the comparison of alternative intellectual strategies, in the light of historical experience and precedents".[18] Although initially these disputes may not be exactly substantial or to the point, they are becoming to be such, as scientists are starting to realize the position they found themselves in. From that point they eventually move from formal to historical argumentation and this argumentation is governed by rules resembling those of argumentation in a courtroom. Toulmin even compares their decision making process to the one of a Supreme Court of the United States in which it is reinterpreting constitution while taking always into account the function that a given law should serve in contemporary socio-historical circumstances. When the very foundations of science are being reconsidered, it is obvious that criteria and decision-making procedures will not be stated in an unequivocal manner. Even though, if only the scientists are able to consider their situation and aims in the context of the current situation of their discipline and appeal to its history, they will secure the rational foundations for their decision.[19]

Both described above types of processes of change can be – according to Toulmin – analysed in terms of reasons, that is rational, intra-scientific factors influencing scientists' decisions. Apart from them he distinguished also a type of change characterised by the crisis of rationally. Cases of this kind cannot be judged in terms of reasons, but rather in terms of causes, which are extra-scientific factors determining the decision-making processes within science. In reality, no decision within a discipline is free from such factors, but in both discussed situations, they play a marginal role. These very factors can, however, sometimes decide not only about the character of science, but about the very possibility of its coming into being. If the development of a concept is a matter of a collective work on a problem situation, the necessary condition for the appearance of variations is always the existence of a forum within which potential innovations can be discussed and modified. The development and survival of a concept is dependent on the environment, and specifically on such social conditions as the existence of scientific institutions – which are exactly what creates the "ecological niches". Barriers of those niches

[17] *Cf*. *ibidem*, p. 237.
[18] *Ibidem*.
[19] *Cf*. Toulmin's account of the dispute about the status of quantum mechanics, *ibidem*, pp. 236–242.

are places where professional factors meet the environmental ones. For the optimal course of the process of conceptual evolution, the relations between those factors must be optimal. The mentioned barriers cannot be too low, since in that case the new concepts would not be able to establish stable relations within some established edifice of science. As a result they would dissolve in endless debates and lose their unique characteristics or they would be eliminated precisely because of their innovative nature. This was the case, as Toulmin noticed, in ancient China, where despite the advancement of applied sciences, no pure sciences have emerged; i.e. there was not a programme of theoretical investigations, which would determine the scope of problem situations. On the other hand, when the barriers of an ecological niche are too high, new concepts, although enjoying recognition among the professionals, cannot enter into the broader market of ideas – as it was in Babylon, where highly professionalised astronomers isolated themselves from other social groups and protected their secrets, which led to the disappearance of their findings and methods with the downfall of the state.[20]

The sole existence of scientific institutions marks the fact that a given society has fulfilled the fundamental condition for conceptual evolution, i.e. that its people recognised that their system of knowledge is insufficient for them. Additionally, the coming into being of a scientific discipline depends often on the emergence of institutions which expect to benefit from providing scientists with conditions to conduct their research. But of course the sole existence of scientific institutions and professions does not necessary have to mean that there exist conditions for the unconstrained development of science,

For the life of science is embodied in the lives of these men: exchanging information, arguing, and presenting their results through a variety of publications and meetings, competing for professorships and presidencies of academies, seeking to excel while still vying for each other's esteem.[21]

Scientific disciplines evolve along the evolution of scientific institutions and professions and the scientists holding achievements are gaining authority and defining the institutional framework of their professions. The new concepts and people behind them have practically no chance of entering the forum if they are lacking support of the authorities. However, science in its social aspect does not function as homogenised whole with strict intra-relations. On the contrary, the populational analysis shows that within the domain of science various institutions are in a state of permanent competition for prestige and authority. The existence of authority groups is a result of the function they fulfill within the institutional structures of science which is the supervision of institutions in their achieving the disciplinary goals. When such groups fail to fulfill their function they may be "dethroned" – although they have an influence on the shape the domain of science, authority groups are always subject to criticism from the world of science which may boycott or even overthrow them.[22] Parallel to the existence of leading institutions, smaller associations of scientists are being formed which try to develop

[20] *Cf*. *ibidem*, pp. 215–220.
[21] *Ibidem*, p. 262.
[22] *Cf*. *ibidem*, pp. 274–275.

concepts alternative to the mainstream. The function they serve is twofold. First of all, the innovation proposed and discussed within them feed the mainstream. Second of all, the moment they manage to achieve successes they themselves enter the mainstream and start to shape it.[23] It is these types of processes that, to a significant extent, influence the development of scientific professions and institutions and, at the same time, of the intellectual content of disciplines. Analogically to the evolution of concepts, in the evolution of scientific institutions, scientists are passing authority to these groups which within their own enterprises managed to adapt to the changing ecological situation by solving previously unsolved problems. The professional development of science should be seen as a process in which new generations replace the older ones. The intellectual content of science cannot change independently of the changes within institutions which determine this content and the lack of institutional development goes in hand with intellectual stagnation. And vice versa, the institutional changes are always related to the intellectual ones. If conceptual evolution is to proceed optimally, these aspects of science must be interrelated. However, Toulmin emphasises that "The Social factors are *necessary*, but the intellectual ones are *crucial*."[24]

II. The failure of Toulmin's theory

The evolutionary theory of conceptual development outlined in *Human Understanding* has met with some serious criticism on the part of philosophers of different traditions. Although not all of this criticism has been equally apt nor justified, much of it has been. In the next paragraphs I will discuss some of the crucial arguments levied against Toulmin's conception. As we shall see, although some of them were missing the point, others were on the right track and, as I will be trying to show, what made them valid was the evident discrepancy between what Toulmin claims he is arguing for, and to where his arguments are actually leading. Among the most serious challenges to Toulmin's conception, Tomasz Zarbski lists Toulmin's unjustified use of the concepts of evolution and the charge of relativism.[25] Let me start my discussion of the reasons of the failure of Toulmin's project with a few words about the first one.

II.1 Scientific vs. biological evolution

It is usually the case that when a theoretical device designed for describing a specific class of phenomena is being used outside its intended domain, many question the justification of such a measure. It is therefore not surprising that such doubts were also cast on Toulmin's use of the concepts of the theory of biological evolution in order to describe the development of science.Most notably, in his review article *Is the progress of science evolutionary?* Jonathan L. Cohen argued that the fundamental problem of Toulmin's conception is the fact that it is founded on a misguided analogy between both types of development – of scientific knowledge and living organisms. Cohen pointed out many discrepancies between

[23] As Toulmin notices this was the path of the Royal Society. See *ibidem*, pp. 273–274.
[24] *Ibidem*, p. 221.
[25] *Cf*. T. Zarbski, *Od paradygmatu do kosmopolis*, Wrocaw 2005, p. 183.

the theory of biological evolution and the discussed conception and stated that "Toulmin's claim to be using the term 'evolutionary' in the precise and strict neo-Darwinian sense seems hardly more accurate than the claim of some cultural relativists to be generalising from relativity physics".[26]

Instead of discussing or contesting here the details of this line of criticism, I would rather like to suggest a way of defending Toulmin's conception against this and any similar arguments. As Zarbski aptly noticed, the conformity of Toulmin's conception with the theory of biological evolution is not a *sine qua non* condition for it and proving any disconformities does not compromise it.[27] A careful reading of *Human Understanding* confirms that this is also Toulmin's own view, since he often mentions that there are no necessary connections between both theories. And so, although at one point he writes that his "analysis should be an 'evolutionary' one, not just in the broad sense of being non-revolutionary, but in a quite precise and strict sense of the term",[28] he quickly adds, that "it will not be necessary to assume – as Ernst Mach unfortunately supposed – that intellectual evolution has something 'biological' about it, or even that the process of conceptual change in the sciences displays any substantial resemblance to the process of organic change."[29] The ambiguity of the first quotation is most likely due to the fact that Toulmin sought a rationale for his theory as something more than just a critique of Kuhn's conception while the latter clearly shows that Toulmin did not want to find his conception on any strict analogy to the actual theory of evolution. Hence, just proving that Toulmin's conception is not entirely consistent with the theory of biological evolution does not undermine it.

What, however, could undermine Toulmin's conception is an examination of how the development of scientific knowledge – accounted for in terms derived from the theory of evolution–fits the realties which it is supposed to explain and to what extent the adoption of Darwin's apparatus allows Toulmin to fulfill his main declared goal which was to defend the development of science as a rational process. This is the line of thought I shall follow in the oncoming paragraphs, and the mentioned charge of relativism offers here a good starting point.

II.2 Toulmin's relativism?

Let me recall from the previous sections, that for Toulmin the adoption of tools and concepts from Darwin's theory was connected with some precise goals which it was supposed to achieve. First of all, it was supposed to be helpful in designing a theory which would allow us to offer one consistent explanation of both, the changes in science and lack of them. To achieve this aim the theory would have to exclude relativism and at the same time admit for some sort of historical relativity. Secondly, such a theory should be compatible with the claim that the historical change within our systems of knowledge is a rational process. Both these issues

[26] L.J. Cohen, 'Is the Progress of Science Evolutionary?', *The British Journal for the Philosophy of Science* 24 (1973), p. 49.

[27] *Cf*. T. Zarbski, *Od paradygmatu...*, p. 139.

[28] S.E. Toulmin, *Human understanding*, p. 134.

[29] *Ibidem*, p. 135.

are for Toulmin closely connected. It is therefore not unreasonable to conclude that if it is not the conformity with Darwin's theory that secures the success of Toulmin's proposition, it is the extent to which the adoption of evolutionary tools allowed him to achieve his own stated goals.

With this conclusion in mind let us first examine how Toulmin's theory deals with the problem of relativism. There are at least two ways of doing this. We can either look at Toulmin's conception 'from the outside' and ask, whether it possess features typical of conceptions we usually deem relativistic, or we can look at it 'from the inside' and confront its aspects with Toulmin's own claims about relativism.

In respect of the first way, some reasons for giving a positive answer to the stated question can be found in the work of Alina Motycka.[30] The Polish author claims that the moment Toulmin's states that there are no universal criteria of choice and emphasises the influence of psychological and sociological factors on the shape of science, he faces the problem of relativism, "which cannot be resolved by changes in terminology."[31] As she notices in Toulmin's situation of competition, the conceptual change is influenced by the role of individual scientists, the socio-cultural context and the social background. As a result, in the case of the clash of values, we cannot say anything about the decision-making process if we will not see it as relative to the given situation. For Motycka it is futile to argue here that the adaptive capacity of concepts allows them to adapt to changing scientific environments and that change is a necessary condition for continuity, because "such argumentation is viciously circular–what is assumed by the theory of evolution is becoming an argument for evolutionary continuity of change."[32]

Zarbski tries to refute Motycka's arguments on the ground of Toulmin's conception by claiming that although Toulmin admits the role of extra-rational factors, he denies that scientific change can be caused only and solely by them and emphasises that it is the rational factors that always prevail. In a situation of competition, scientists appeal to their own experience and substantial arguments. Zarbski emphasises also that on the ground of Toulmin's conception, wrong decisions may always be verified in the light of empirical data. Although I am quite sympathetic with his claim that Motycka's arguments result from her "undue devotion to the logical model of rationality and are put forward from the absolutist position",[33] I also think that they are nonetheless quite accurate, while Zarbski's refutation amounts to the restating of the claims from *Human Understanding* (like the above claim about the prevalence of rational factors) and is passing over their insufficiency. The claim–that scientists in the situation of competition are appealing to their own experience–may well serve for as against Toulmin's relativism because they appeal to their disciplinary experience, which is what their decisions are relative to. As to the empirical verifiability of innovations, it is worth noting that

[30] A. Motycka, *Relatywistyczna wizja nauki: Analiza krytyczna koncepcji T.S. Kuhna i S.E. Toulmina*, Wrocaw 1980.

[31] *Ibidem*, p. 89.

[32] *Ibidem*.

[33] T. Zarbski, *Od paradygmatu...*, p. 142.

by introducing various concepts or levels of objectivity,[34] Toulmin is perhaps not entirely separating experimental practices from science as a methodological tool, but without a doubt he is diminishing their role in science. In fact, by introducing the notion of objectivity as a feature ascribed to science not only in light of the confrontation of theoretical claims with empirical data, but which also can be ascribed to strategies which can generate concepts, which can in turn, generate claims liable of empirical verification.[35] Toulmin trivialises the role of the empirical verification of scientific theories and experimental practices of science. It is, perhaps, even more striking if we realise that he simultaneously claims that his theory is supposed to do justice to the actual practices of scientists and to the history of scientific practices that shows the importance of empirical evidence at the point of adopting new concepts. As to Motycka's vicious circle argument, it is refuted by Zarbski by a repetition of this circle, because he claims that the lack of continuity of science is in Toulmin's theory only alleged, because "when the change of scientific strategy is well justified, and therefore rational, then the continuity of science is preserved".[36]

On the other hand, it is doubtful that Motycka's arguments simply prove that Toulmin's conception is unviable as non-relativistic. This is partly due to her "absolutist approach" which makes any sign of relativity tantamount to relativism, and partly because of a more general problem of stating a precise commonly accepted definition of relativism. The latter is an issue for any "external" criticisms of relativist conceptions and it is probably not something that cannot be overcome, but if we would manage to show that a given idea is not a form of relativism in light of some external criteria, but judged by its own terms, it would make all further arguments unnecessary.

Toulmin formulated his theory to a significant extent in opposition to T.S. Kuhn conception of the revolutionary development of science – which he explicitly deemed relativistic. If, therefore, by confronting his criticism of Kuhn with his own claims, we could show that certain aspects of his theory are consistent with certain aspects of Kuhn's conception, we might prove that Toulmin's conception is a form of relativism by Toulmin's own standards. Let us then proceed in this fashion.

Among other things, Toulmin criticises Kuhn's conceptionof revolutionary development – presented in *Human Understanding* in a very tendentious way – as not doing justice to the historical realities and for the fact that his conception of paradigms is independent of his conception of scientific revolutions.[37] The letter is supposedly due to the fact that for Kuhn, the notion of a paradigm is identical to the notion of a logical system. However, Toulmin himself *de facto* introduces two types of change in science – the change in response to the routine and to the extraordinary problem situations – and those types are also to a significant extent independent. Moreover, both these types seem to correspond with those of Kuhn.

[34] S.E. Toulmin, *Human Understanding*, p 242–244.
[35] Cf. *ibidem*, p. 243.
[36] T. Zarbski, *Od paradygmatu...*, p. 144.
[37] For Toulmin's discussion of Kuhn's conception see *Human Understanding*, pp. 96–130.

The development of science within a paradigm is supposed to be characterised by the fact that the paradigm organises the work of scientists to the extent that it influences the way they perceive phenomena and makes their work tantamount to "solving puzzles" that are, in a way, assigned by the paradigm.[38] But, as we have seen, the first distinguished types of change in Toulmin's conception involves the problems and solutions to them which are determined by the accepted intellectual or explanational ideals. Furthermore, according to Kuhn, the exhausting of the pool of puzzles to solve within a paradigm and the appearance of more and more cases which scientists are unable to account for, leads to a crisis in a discipline which eventually results in a revolution – i.e. substitution of old paradigm by a new one.[39] Toulmin, in turn, claimed that extraordinary problem situations arise when disciplinary ideals and concepts accepted within them are losing their capability to explain newly discovered phenomena. In other words, they are being exhausted and, as a result, the very goals of a discipline must be rethought and new ideals established. What is more, both authors emphasise the change in the nature of scientific reasoning and the role of extra-scientific factors in times of this type of change and the differences between their conceptions are to a large extent the differences of terminology – what one calls 'persuasion', the other calls 'argumentation' from the history of science.

Toulmin eagerly grasps at the letter of *The Structure of Scientific Revolutions* and accuses the conception outlined there of reducing the paradigmatic change to persuasion and of postulating complete incommensurability of paradigms which make it impossible for the adherence of two different paradigms to communicate with each other. He, in turn, claims that we are never dealing with communication breakdowns.[40] It does not occur to him, that even at this point his claims are quite compatible with Kuhn's and that the too literal reading and related criticism of Kuhn leads him to contradict himself. Consider, for example, the contrast between his claim about the substantial discussion throughout the time of Copernican revolution and what he said about the coming to being of atomic physics (i.e. the Thomson-Rutherford case mentioned in I.3). Both are cases of the second type of scientific change – of revolution for Kuhn and of dealing with extraordinary problem situations for Toulmin. However, in the first case Toulmin stresses – while opposing Kuhn – the existence of mutually communicable reasons throughout the whole extended process of change from Ptolemy's to Copernicus' astronomy[41] and in the second case, which is quite analogical but brought about in a different context, he emphasises the problems of communication between scientists resulting from their adherence to different intellectual ideals (he mentions how many scientists contemporary to Thomson and Rutherford were unable to accept their postulated ideal and thus understand them[42]).

[38] *Cf*. T.S. Kuhn, *The Structure of Scientific Revolutions*, Chicago–London 1996, pp. 35–42.
[39] *Cf*. *ibidem*, pp. 123-163.
[40] *Cf*. for example his account of the debate during Copernican revolution in *Does the Distinction between Normal and Revolutionary Science Hold Water?*, [in:] I. Lakatos, A. Musgrave (eds.), *Criticism and Growth of Knowledge*, Cambridge 1970, pp. 43–44.
[41] *Cf*. *ibidem*; as well as *Human Understanding*, p. 105.
[42] *Cf*. S.E. Toulmin, *Human Understanding*, p. 153.

The existence of a rational debate is for Toulmin a warrant of not falling into relativism, but considering the insufficient justification of this rationality (which will be yet discussed) it seems quite evident that many differences between his and Kuhn's conception have a rather verbal character. The fact that Kuhn is not clear on the issue whether the process of scientific change is rational or what sort of rationally would it be while Toulmin does not mean that if judged by some same criteria, both conceptions wouldn't be equally postulating rationality (or lack of it).

II.3 Rationality and logicality

Before moving on to discuss the difficulties with Toulmin's justification of his rationality claim, it is worth to consider a broader issue that to a significant extent is influencing the whole concept. The rejection of traditional notions of rationality – i.e. those which supposedly equal rationality with reasoning founded on the principles of formal logic – serves as peculiar *leitmotif* of much of Toulmin's philosophy and, while not questioning here the reasons of his criticism, I would like to focus on some problems of the positive or constructive side of this criticism. The most problematic issue here concerns the very postulate of separating rational reasoning and logic.

Although central to Toulmin's though, this postulate is far from being clear and it can be understood in two ways. We can conclude that what Toulmin is claiming is just that what is rational does not correspond completely to what is logical. Should this be the case however, his postulate would be barely novel. It is familiar to philosophers at least since times of David Hume and only a few would question the rationality of many of our behavioural patterns despite of their lack of logical rationale. The other way of making sense of Toulmin's claim is to conclude that according to it, the categories of rationality and those of logicality are mutually exclusive. Indeed, many of his comments suggest that while ignoring any options for a middle ground, Toulmin consents to the claim that being logical has nothing to do with being rational. Larry Briskman notices[43] that the reason why Toulmin leans to this view is his very limited view of logic as "concerned simply with the inner articulation of intellectual systems whose basic concepts are not currently in doubt".[44] On this account, logic cannot be anything more than a set of directives, which allow us to transfer the truth or acceptance of premises to the inferred conclusions. As such, it is not capable of accounting for the processes of accepting new concepts in science. This process – as described in I.3 – is not based on the pattern of deduction and confirmation, but rather it requires, from time to time, a break with the old standards and the introduction of new ones – which is supposed to confirm its rationality. Thus, Toulmin has to claim that that scientific rationality is not connected in any way with logicality, because only then can he postulate the rationality of the second distinguished types of change. The talk about ties between logic and rationality is possible only in relation to the method-

[43] *Cf*. L. Briskman, 'Toulmin's Evolutionary Epistemology', *The Philosophical Quarterly* 95(1974), pp. 160–169.

[44] S.E. Toulmin, *Human Understanding*, p. 84.

ological principles of science, whose counterpart in Toulmin's conception are the explanatory ideals stating not only their scope but also the ways of conducting research (another thing is the fact that Toulmin does not distinguish clearly the methodological principles of science from its subject and content). From the types of change distinguished by Toulmin, the first one occurs within an established ideal and the second involves a break with it and the introduction of a new one– which implies also the substitution of the methodological principles of a discipline. The acceptance of such a view of scientific change seems to be precisely what drives Toulmin to the negation of any ties between rationality and logicality.

The central problem of this approach derives from the fact that while it is not hard to see that with the development of science the methodological principles are changing, it is hard to see how it would be true about all of them – including the most fundamental principles of logic which are no more connected to scientific reasoning than to reasoning in general. Imagine a scientist who is not accepting the principle of non-contradiction – what form would his research take if he would be unwilling to see why from two mutually exclusive propositions only one may possibly be true? This would be of course reducing Toulmin's postulate to the absurd and it is doubtful that what he had in mind while insisting on making a distinction between logicality and rationality was the complete separation of these two domains. But this also shows the fundamental ambiguity of Toulmin's own view of rationality and this ambiguity is something that influences his whole conception as a proposition for the account of scientific development as rational.

II.4 Rationality and the development of science

Although the world "rational" appears in *Human Understanding* in all possible cases, it seems, as I have already suggested, that Toulmin finds it much easier to talk about what rationality is not, than about what it actually is. All in all, his argumentation is lacking any clear explication, definition or criteria that would allow one to judge certain behaviours or views as rational (or not). At best, he offers us an intuitive notion of rationality as the capability for critical analysis of one's beliefs. The motto opening *Human Understanding* declares that "A man demonstrates his rationality, not by a commitment to fixed ideas, stereotyped procedures, or immutable concepts, but by the manner in which, and the occasions on which, he changes those ideas, procedures and concepts."[45] Unfortunately this is the most complete insight to the positive side of Toulmin's conception of rationality we can possibly find in *Human Understanding*. The insufficiency of the positive description of the notion of rationality becomes even more problematic when Toulmin explicitly postulates the break with *a priori* in deciding whether a given enterprise is scientific or not[46] (which is, in this context, tantamount to it being rational or not). He claims that only apparently we are forced to choose between accepting fixed definitions of what it is to be scientific and accepting that each historical period has its own standards of assessing intellectual enterprises. The first choice would mean a return to absolutism and the second to relativism

[45] *Ibidem*, p. X.
[46] *Cf*. *ibidem*, pp. 495–503.

and for Toulmin absolutism and relativisms are extreme positions between which a middle ground is possible. To achieve this middle ground we must liberate ourselves from any demarcation criteria whatsoever. Only then would we be able to compare– from the perspective of being scientific–intellectual enterprises of different epochs and cultures.[47] And to compare them we would simply have to take into account the extent to which alternative strategies of different epochs managed to achieve the developing goals of given enterprises.

If, however, only those actions within intellectual enterprises are rational which result in the realisation of stated goals, we cannot, on the grounds of this conception, talk about the rationality of any enterprise until we find out in what way it has managed to achieve its goals. So on this view, rationality can be, in practice, attributed to a given action only *ex post* and, as a result, we are lacking any tools for analysing the rationality of the choice of goals and, to some respect, also the means employed to achieve them. When rationality is made completely relative to the aims of an enterprise, we are forced to judge many enterprises as rational only because they managed to achieve their stated goals – even if we otherwise would consider either their goals or their methods irrational.

Consider, for example, the case of beliefs of the proponents of Copernican astronomy. In his classic *The Copernican Revolution,* Kuhn distinguishes two factors driving the work of Johannes Kepler.[48] First of all, he came into possession of the very accurate data collected by Tycho Brache; and second of all – and what will interest us here – he was an ardent supporter of Neo-Platonism – convinced that the whole Universe is governed by simple mathematical laws and that the Sun is the only possible cause of the motions of the heavenly bodies. There can be little doubt that the Neo-Platonic worldview of the German astronomer played a crucial role in his discovery of the famous three laws of planetary motion, which we commonly judge as scientific today. However, the same system of beliefs and methods induced him to derive conclusions, whose scientific character is widely contested today – to say the least.[49] How should we therefore account, on the grounds of Toulmin's conception, for Kepler's investigations in respect of their rationality, if his research was to a large extent determined by unscientific (and, as such, irrational) factors? Since Toulmin's conception offers us no practical tools for assessing the rationality of methods and goals, we can only conclude that it was rational of Kepler to accept a certain metaphysical worldview, as it was rational of him to pursue an account of the structure of solar system in terms of harmony between "cosmic" solids. Of course, this would go far beyond of what Toulmin would like to include under the label of rationality. It seems, though, that while Toulmin emphasised the difference between scientific and extrascientific factors influencing the development of science, this distinction breaks down when we try

[47] Provided that in this epochs and cultures there exist collectively established problems which are the subject of collective investigation and that these problems are sufficiently similar. *Cf*. *ibidem*, p. 498.

[48] *Cf*. T.S. Kuhn, *The Copernican Revolution*, Cambridge–London 1995, pp. 209–219 (212–214).

[49] With the so-called Kepler's Fourth Law as a prime example.

to apply his conception to the actual historical cases. The apparent reason for this is Toulmin's tacit reversal of reasoning, in which he starts with the assumption that science is rational and only afterwards tries to make his conception of rationality to agree with this assumption. That's probably also the reason why he never discusses cases of, in the long run, ineffective scientific enterprises – if genuine scientific enterprises are thoroughly rational in character and rational enterprises are successful we cannot talk about rational and, at the same time, unsuccessful scientific enterprises. Such reversed reasoning also accounts for his insistence on the rational character of scientific change – if science is rational and science is changing than this change is rational.[50]

As I have already suggested in the previous paragraph, Toulmin needed to break with what for him was the paradigm of rationality as logicality in order to maintain the claim that rationality can be attributed to any type of conceptual change, including one involving the substitution of whole disciplinary ideals. I have already tried to show that the introduction of this particular kind of scientific change caused Toulmin's falling into relativism. The same may now serve as an argument for the irrationality of science in Toulmin's conception. After accepting the claim that in the face of problems which call for the transformation of the established scientific approach in a given discipline, scientist are changing more formal argumentation for arguments appealing to analogical situations in the history of their discipline Toulmin cannot avoid the assertion that their arguments are to a significant extent persuasive in nature. Motycka points out to this very moment when she notices that "the irrationality of science in Toulmin's conception is that in the case of change of disciplinary and theoretic principles, science does not provide sufficient reasons and arguments for the selected choice; it is motivated by causes (i.e. external factors)".[51] That this is truly the case we can easily convince ourselves by reflecting on Toulmin's claim concerning the nature of argumentation in the face of this type of change. His postulate of rationality is founded here on the analogy between scientific and legal argumentation, in which assumed rationality of the latter warrants rationality of the former. But in what way is legal argumentation rational? By claiming that the rationality of courts of law and legal argumentation is reflected in their objectivity because the sentences of justice oriented on interpreting law in the context of a given socio-historical situation, Toulmin contradicts himself. To reinterpret law and adjust it to the current context is nothing else but to relativise it to the socio-cultural circumstances. The key problem here is that it does not spring to Toulmin's mind that court verdicts are, like nothing else, reflecting not the collective agreement on what is best for society, but of the interests and beliefs of ruling or dominating classes or simply of (as it is usually the case) socio-cultural beliefs of a given historical epoch. To

[50] Motycka reaches similar conclusions when she writes: "Toulmin's reasoning appears to be, in short, of this form: since the changeability of science (a historical fact) is at odds with the accepted notion of rationality, then if we want to maintain the rationality of science and can't deny the fact that it is changeable, we have to take the changeability of science for its rationality", A. Motycka, *Relatywistyczna...*, p. 90.

[51] *Ibidem*, p. 101.

insist on something else is to claim that it is equally rational to sentence convicts to death in the past as it is to sentence them to life imprisonment today and if we are willing to agree that this is the case, we are falling into relativism, what, on the ground of Toulmin's conception, is tantamount to falling into irrationalism.[52]

Conclusions

In the second part of the article I have tried to show that the adoption of conceptual apparatus of the theory of biological evolution did not allow Toulmin to fulfill the hopes he had for such a measure. Although I dismissed the charge of Toulmin's unjustified use of the theory of evolution, I think that what is actually problematic is the very choice of Darwinian concepts for designing a conception of the development of science as a rational process in the first place. It is worth noting that although he accounts for the development of science in categories of evolutionary change, he cannot avoid the claim that this development is progressive. It is evident in many moments of *Human Understanding* because Toulmin does not limit himself to the claims that the conceptual change is tantamount to substituting old concepts for new ones but he says that this process "involves replacing one set of concepts by another *improved* set."[53] By claiming this, Toulmin tacitly departs from Darwin's theory in order to defend the rationality of his image of science. To create a foundation for such a defense he must, at least implicitly, accept that the development of science is a goal-oriented process, whereas the theory of evolution accounts for biological change as not directed towards any particular end. One of the main achievements of Darwin's theory was exactly that it accounted for the process of the development of biological organisms as "blind" and "accidental." It means that, firstly, accidental is the very way in which variations are coming into being and, secondly, their preservation is an effect of their better, from the point of survival, adjustment of their contingent features to the conditions in certain ecological niches in certain periods (the very process of preservation is of course not accidental).[54] And it is precisely survival by adaptation that is the only goal about which we can talk (and still with some oversimplification) in the case of Darwin's conception of the development of biological organisms. The process of achieving this "goal" cannot be in any case called progress, because there is no fixed direction or end point of the process of evolution. However, Toulmin, while talking about the historical process of the development of science which he sees as rational, does not avoid deeming it progressive. And it cannot be the case that what he had in mind was progress in a purely adaptive sense – which is connected with another problem of Toulmin's adoption of Darwin's explanatory schemata. Namely, Darwin's theory of evolution talks about the adaptation of organisms to ecological niches which are given. In Toulmin's conception, evolving concepts do not only adapt to the ecological niches, but are also – to a large extent – creating

[52] One could, perhaps, point out here that I am ignoring the seemingly extended discussions of rationality and models of rational reasoning presented by Toulmin primarily in his *The Uses of Argument* (Cambridge 1958); however, as

[53] *Ibidem*, p. 486; emphasis mine.

[54] *Cf*. for example Ch. Darwin, *On the Origin of Species*, London 1859, pp. 80–87.

them. Toulmin's theory cannot therefore limit itself to, as it is the case, accounting for changes in our systems of knowledge by appealing to the process of adaptation to the ecological conditions of the environment, because science itself is shaping these conditions. If so, then Toulmin's argument in defense of the rationality of scientific change as adaptation seems to be completely losing its ground.

Ethics and Philosophy of Politics

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

RICHARD KEARNEY*
Boston College

Forgiveness at the Limit: Impossible or Possible?

Abstract

In the paper, the author analyses the concept and the phenomenon of forgiveness as discussed by several contemporary philosophers. He focuses his attention upon the contemporary debate on forgiveness at the limit, with particular reference to the question of pardon as a secret gift. Several contemporary thinkers have responded to the question of the limits of forgiveness. Jankelevitch and Primo Levi have affirmed the impossibility of forgiving those who do not ask for forgiveness. Hannah Arendt talked of the impossibility of forgiving radical evil; Derrida has written of the impossibility of pure forgiveness tout court. Paul Ricoeur seeks an alternative response to the limit of forgiveness. He attempts to give due credence to the strong arguments of Derrida, Jankelevitch and Arendt, while seeking to shift the final emphasis from 'impossible' to 'difficult'. In author's discussion of forgiveness of the radical evil, he argues that it calls for an answering power of radical good. He stresses that the possibility of forgiveness is a "marvel" precisely because it surpasses the limits of rational calculation and explanation; gratuitousness of the pardon is due to the very fact that the evil it addresses is not part of some dialectical necessity: the pardon is something that makes little sense before we give it, but much sense once we do.

* Richard Kearney holds the Charles B. Seelig Chair of Philosophy at Boston College and has served as a Visiting Professor at University College Dublin, the University of Paris (Sorbonne) and the University of Nice. He is the author of over 20 books on European philosophy and literature (including two novels and a volume of poetry) and has edited or co-edited 14 more. He was formerly a member of the Arts Council of Ireland, the Higher Education Authority of Ireland and chairman of the Irish School of Film at University College Dublin. As a public intellectual in Ireland, he was involved in drafting a number of proposals for a Northern Irish peace agreement (1983, 1993, 1995) and in speechwriting for the Irish President, Mary Robinson. He has presented five series on culture and philosophy for Irish and/or British television and broadcast extensively on the European media. His most recent work in philosophy comprises a trilogy entitled *Philosophy at the Limit*. The three volumes are *On Stories* (Abingdon–New York 2002), *The God Who May Be* (Bloomington, 2001) and *Strangers, Gods, and Monsters. Interpreting Otherness* (London–New York, 2003).

J. Derrida, *On Cosmopolitanism and Forgiveness*, London–New York 2001; and *'On Forgiveness' (in dialogue with Richard Kearney et al)*, [in:] J. Derrida, *On Cosmopolitanism and Forgiveness*, pp. 52–72. See also J. Derrida, 'Hospitality, Justice and Responsibility', [in:] *Questioning Ethics: Contemporary Debates in Philosophy*, R. Kearney, M. Dooley (eds.), London–New York 1998, pp. 65–84.

Several contemporary thinkers have responded to the question of the limits of forgiveness. Vladimir Jankelevitch and Primo Levi have both affirmed the impossibility of forgiving those who do not ask for forgiveness. Hannah Arendt talked of the impossibility of forgiving radical evil; and more recently Derrida has written of the impossibility of pure forgiveness *tout court.*

Paul Ricoeur seeks an alternative response to the limit of forgiveness. In an essay entitled 'Difficult Forgiveness' – which serves as Epilogue to his last major work, *Memory History, Forgetting* – he attempts to give due credence to the strong arguments of Derrida, Jankelevitch and Arendt, while seeking to shift the final emphasis from 'impossible' to 'difficult'. (As he confessed, the key word separating him from Derrida is the word impossible). In what follows I will address this contemporary debate on forgiveness at the limit, with particular reference to the question of pardon as a secret gift.

I

Let me begin with a short account of Derrida's approach to forgiveness before looking to Ricoeur's alternative reading. I believe this crucial debate serves to illustrate the different moral positions adopted by hermeneutics and deconstruction at the close of the 20th century.

Why is pardon impossible for Derrida? We can only forgive the unforgiveable, he says, and that is precisely what cannot humanly be forgiven. If someone asks for forgiveness that person has already atoned and so does not require forgiveness. Only radical evil and hatred, the imprescriptible crime, the irreparable effect, the inexpiable act, are matters for forgiveness. Such forgiveness is therefore, for Derrida, unconditional, undeserved and ultimately impossible. But if it were possible it, and it alone, would be true.

How does Derrida come to this conclusion? Pure forgiveness, if it existed, would be beyond repentance, atonement or any account of the crime. It would include the pardoning of radical evil and have nothing to do with reconciliation, healing, remorse or repentance. It would be forgiveness of the 'guilty as guilty';[1] and, as such, it would not be applicable to those who have repented or apologized (and are therefore no longer guilty). Conditional forgiveness is not forgiveness, argues Derrida, because it is 'corrupted' by calculations of the weight of crime and punishment. Unconditional forgiveness, by contrast, would involve forgiving the unforgivable (pace Arendt and Jankelevitch) and is impossible. It has nothing to do with judgment, punition or recompense. It is beyond laws, norms and obligations. Even the Abrahamic account of forgiveness is ultimately compromised, Derrida suggests, in that it introduces the notion of pardon in proportion to repentance; and, so doing, it limits its own ostensible message of pure gratuity and generosity. True unconditional forgiveness is *madness*, a private and inaccessible event, never a matter of public or political action. It lies beyond the logic of rights or duties.

Unconditional and conditional forgiveness are, Derrida concludes, irreducibly

[1] J. Derrida, *On Cosmopolitanism and Forgiveness*, p. 3.

heterogeneous and irreconcilable.[2] Forgiveness calls for a 'hyperbolic ethics' beyond ethics. And in this sense Derrida holds out forgiveness as an impossible ideal, even as he admits that in everyday life and history we have to engage in acts of pardon 'in a series of conditions of all kinds ('psycho-sociological, political' etc)'.[3] But the problem, as I see it, is that there is no way for Derrida to *transit* or *translate* between the conditional and unconditional. There are no criteria, mediations or orientations. Pardon is, at best, a leap in the dark, a form of insane guesswork or indiscriminate decision. All we know is that we can only forgive the unforgivable, except perhaps for the unforgiving, namely those who refuse to forgive. And this, of course, places a heavy burden to forgive on the victims of radical evil while affirming that all perpetrators of radical evil must be unconditionally forgiven. This seems unjust, to say the least; but we must remember that we are not talking here of what is possible. Maybe pure forgiveness has little to do with real human beings, since it is unrealisable in any case?[4] Who knows?

II

Ricoeur takes Derrida's account on board while moving from the impossible to the possible. From the outset, Ricoeur confesses that his analysis will be formulated in the 'optative' mood. It will operate under the sign of a certain 'eschatology' of memory.[5] In other words, he lets us know that he is going to discuss the possibility of 'difficult' forgiveness in terms of a projection of an act of unbinding – an act which goes beyond the limits of law and prescription, crime and punishment, fault and reparation. (Limits to be respected and recognized as necessary in the order of politics and justice). But unlike Derrida, who sees such forgiveness as a hyperbolic and impossible ideal, Ricoeur wants to inscribe it under the sign of an 'anthropology of capable being': an anthropology grafted onto a philosophy of religion which says '*you can forgive*'.[6]

How does he propose to do this? Let me briefly trace Ricoeur's argument. Just as the voice of evil, fault and guilt proceed from the unfathomable depths of human selfhood, the voice of forgiveness is a 'voice from above'.[7] To the abyss of radical evil responds the vertical height of forgiveness. There is a radical disproportion between this polar dichotomy of depth and height which, Ricoeur concedes, constitutes the 'torment' of his analysis. But while he is prepared to agree with Derrida that forgiveness is indeed directed towards the unforgivable (it is without condition, exception or restriction), he refuses to conclude that it is therefore *impossible*. Suggesting how the seeming impossibility of forgiveness gives way to possibility is the difficult task of his reflection.

[2] J. Derrida, *On Forgiveness...*, p. 44.

[3] *Ibidem*, p. 49

[4] J. Derrida, *To Forgive*, [in:] J. Caputo, M. Dooley, M. Scanlon (eds.), *Questioning God*, Bloomington 2001, pp. 21–51. See also the excellent commentary by Marguerite La Caze, *Wonder and Generosity: Their Role in Ethics and Politics*, Albany 2013. See in particular chapter 6.

[5] P. Ricoeur, *Memory, History, Forgetting*, Chicago–London 2004, p. 459.

[6] *Ibidem*, p. 463.

[7] *Ibidem*, p. 467.

First, Ricoeur insists we separate the unforgiveable and the imprescriptible, for while the imprescriptible – e.g crimes against humanity, genocide – require justice to be done, pardon operates at a level of surplus love beyond the limits of justice. 'To forgive (genocide) would be to ratify impunity, which would be a grave injustice committed at the expense of the law, and even more so, of the victims'.[8] This does not mean of course that forgiveness dispenses with justice, only that it supplements it with a logic of excess and gift beyond the economy of exchange, and outside the circle of accusation and punishment. His solution to this dilemma will ultimately be an unbinding of the agent from the act. (Or as Augustine might have put it, of the sinner from the sin). But I shall return to this in a moment.

Ricoeur, unlike Derrida, accepts that a certain stage of exchange is part of the odyssey of the 'spirit of forgiveness'.[9] Ricoeur believes (again pace Derrida) that at the level of practice, there does exist a correlation between forgiveness requested and forgiveness granted. And he cites the example of certain exceptional public gestures like Chancellor Brandt kneeling in Warsaw or the Pope during his visit to Jerusalem. Ricoeur agrees that while only the victims can forgive (no one can do it for them), there is still a possibility of verticality which can supplement, without dispensing with, this limit of forgiveness. This is where Ricoeur rejoins the question of forgiveness as gift (*par-don, ver-geben*). The difficulty with gift as a model of exchange is, Derrida and other critics argue, that it can place the beneficiary under a debt he/she cannot repay. But this is to remain within the economic model of market exchange. And that is precisely what the commandment to love one's enemy as oneself contests in so far as it breaks the rule of reciprocity and 'requires the extraordinary'. Proposing a non-market form of gift as love or 'extravagance', Ricoeur proposes that 'faithful to the gospel rhetoric of hyperbole, according to this commandment the only gift that is justified is the one given to the enemy, from whom, by hypothesis, one expects nothing in return. But precisely, the hypothesis is false: what one expects from love is that it will convert the enemy into a friend according to a vertical event of surplus. And this surplus implies an unfathomable enigma of asymmetry between the height of forgiveness and the abyss of guilt.[10] For Ricoeur, forgiveness is difficult indeed – but, again, not impossible!

Ricoeur cites the Truth and Reconciliation Commission in South Africa (1996–1999), established by Mandela and Desmond Tutu, as a model of exchange which seeks to purge a violent past. As a public political process, Ricoeur commends it, while recognizing its limits. The purpose of the commission was, in its own words, to 'collect testimony, console the injured, indemnify the victims and amnesty those who confessed to committing political crimes'.[11] The aim of this process was not in fact pardon as such but reconciliation, in a political sense. And the positive benefits were clear in therapeutic, moral and political terms. 'In offering a public space for complaints and the recounting of suffering, the commission certainly gave

[8] *Ibidem*, p. 473.
[9] *Ibidem*, p. 478.
[10] *Ibidem*, p. 482–483.
[11] *Ibidem*, p. 483.

rise to a shared *katharsis'*[12] the 'amnesty granted by the competent committee did not amount to forgiveness on the part of victims'.[13] It was a matter of the victims having their memories and stories of suffering told and recognized as true by the perpetrators and the committee. While acknowledging the clear limits of this process of 'understanding without revenge' or recompense (the victims were deprived the satisfaction of any normal sanction of a trial – punishment and judgment of perpetrators), Ricoeur nonetheless celebrates the Commission as an 'historic opportunity for a public form of the work of memory and mourning in the service of public peace'.[14] But Ricoeur goes further, for he dares suggest the possibility of seeing under the figure of such a 'public exercise of political reconciliation something like an *incognito* of forgiveness': something which can only occur, at the 'most secret level of selfhood' and personhood. Pardon is not a universal rule or law to be prescribed or imposed; it is an act of surprising gratuity which may emerge through reconciliation but is by no means necessitated or even implicated by it. In short, the exchange model of reconciliation may be inspired or informed by some secret spirit of forgiveness though it is by no means its equivalent. Pardon and reconciliation operate at different levels; but they may nonetheless interanimate each other in secret, non-prescriptive ways.

But here again Ricoeur is faced with the vexed question: how does one overcome the ostensible incommensurability between the unconditionality of forgiveness and the conditionality of the request for forgiveness. Ricoeur proposes a non-market model of exchange of gift and receptivity which nonetheless preserves the polarity of the extremes – of conditional and unconditional. (Asking for pardon must be open to receiving a negative response from the other: I cannot forgive you). But the ultimate question is: 'what force makes one capable of asking, giving or receiving the word of forgiveness?'.[15] In short, what power do we appeal to ask for forgiveness?

Ricoeur looks first to the capacities of unbinding (forgiveness) and binding (promising) to suggest a way of mastering the course of time and giving a continuity to the present by giving a future to the past. He borrows here from Hannah Arendt's notion of the continuation and renovation of action (natality) outlined in *The Human Condition* (a response , in part, to Heidegger's preference for Dasein's mortality and rupture). What is crucial for both Ricoeur and Arendt is the notion that forgiving and promising are capacities which depend on human plurality – that is the idea of persons relating to each other in an intersubjective context. Acknowledging that forgiveness has a religious aura that promising does not, Arendt nonetheless wants to argue that forgiveness, which opposes vengeance, is a *human* power. Even the Gospels, she notes, require that humans forgive each other before they seek forgiveness from God. And this act of unbinding is the token of human freedom, of the ability to find some release from the evils and errors of the past in order to be able to start all over again: what she famously calls the event of

[12] *Ibidem*, p. 484.
[13] *Ibidem*.
[14] *Ibidem*, p. 485.
[15] *Ibidem*, p. 486.

natality. Only through a mutual release from what they have done can humans remain free agents.[16] But while promising represents the possibility of a political act of will (treaties, accords, pacts between governments and peoples), forgiveness is, concedes Arendt, an act of love which keeps a distance from the political.

Ricoeur agrees with much of this; but he goes further than Arendt in insisting that forgiveness needs to be understood not only as the unbinding of debt but , at the very 'heart of selfhood', as the unbinding of the *agent from the act*. But how, we may ask, do we move from the unforgiveable fault to the miracle of forgiveness? Ricoeur responds that forgiveness renders the guilty person *able to begin again* by unbinding the person as agent from the act which, qua act, remains condemned and unforgiveable. He also goes further than Derrida here who argued that to forgive a person but condemn their act is like pardoning a subject other than the one who committed the act:[17] in other words, one would be talking about two different people. But Ricoeur takes a decisive step from impossibility to possibility by appealing to his fundamental notion of *l'homme capable.* This is crucial. The person who committed the crime is *also* a person *capable* of doing otherwise, that is, of committing good acts (including those, post hoc, of repentance and remorse). Here Ricoeur speaks of the radical uncoupling 'at the heart of the very power to act – of agency – namely, between the effectuation and the capacity that it actualizes. This intimate dissociation signifies that the *capacity* of commitment belonging to the moral subject is not exhausted by its various inscriptions in the affairs of the world. This dissociation expresses an act of faith, a *credit* addressed to the resources of self-regeneration'.[18]

It is telling that at this pivotal point in his analysis, Ricoeur speaks of an 'ultimate act of trust' , an act based on an 'intimate' pairing proposed by the Abrahamic memory of the Religions of the Book – namely the pair forgiveness and repentance. This forms a paradox in that the response to forgiveness is implied in the gift itself, 'while the antecedence of the gift is recognized at the very heart of the inaugural gesture of repentance'.[19] And he goes further to suggest that if forgiveness is indeed the supreme height – responding to the abyss of fault – it lasts 'forever' beyond notions of before and after; and this in contrast to repentance which occurs in historical time (whether sudden or protracted). So the paradox relates to a circle – namely, the circle between the gift of forgiveness that remains *forever* and what comes to be *in each instance.* Is this not pardon as the entry of eternity into history?

Rather than engaging here in standard theological arguments about grace and nature, divine or human initiative, Ricoeur prefers to remain within the limits of a philosophy of religion grafted onto 1) an anthropology of human persons as 'capable beings', 2) a fundamental ontology of being as act and power (*dynamis*), to be traced from Aristotle to Leibniz, Spinoza and Bergson; and finally 3) a moral philosophy, as in Kant, which recognizes that the 'predisposition to good' is more

[16] *Ibidem*, p. 487.
[17] *Ibidem*, p. 490.
[18] *Ibidem*.
[19] *Ibidem*, p. 491.

original than the radical propensity to evil (*Religion within the Boundaries of Mere Reason*).[20] Moroever, Ricoeur extends his plea for the primacy of goodness, capacity and natality – evinced in the circle of forgiveness and repentance – to a hermeneutic analysis of the great myths of creation, previously analysed in his *Symbolism of Evil* (Paris 1960). Referring specifically to the Adamic myth, he speaks of the narrative of the Fall as symbolizing something irremediable but in no way inevitable in its consequences[21]. This is a pivotal point for Ricoeur – the excess of capacity over the past. 'The gap with respect to creation holds in reserve the possibility of another history inaugurated in each case by the act of repentance and punctuated by all the irruptions of goodness and of innocence over the course of time'.[22] Indeed, Ricoeur goes on to add that this 'immense project of restoration' can in turn be serviced by a philosophical reading of the Jewish and Christian 'imagination' of the suffering servant. (The terms 'philosophical' and 'imagination' are telling).

Refusing recourse to speculative or transcendental solutions to the paradox of forgiveness and repentance, Ricoeur returns once again to his insistence on a practical philosophy of action uttered in the 'optative mood'. He endorses, in the final analysis, a discreet eschatology whose ultimate word is happiness. 'Under the sign of forgiveness', concludes Ricoeur, 'the guilty person is to be considered capable of something other than his offences and his faults. He is held to be restored to his capacity for acting, and action restored to its capacity for continuing. This capacity is signaled in the small acts of consideration in which we recognized the incognito of forgiveness played out on the public stage. And finally, this restored capacity is enlisted by promising as it projects action toward the future. The formula for this liberating word, reduced to the bareness of its utterance, would be: you are better than your actions'.[23] In short, the power that enables us to give and receive forgiveness is the word: *you are able*! In spite of the ostensible impossibility of forgiveness, you *can* forgive and be forgiven. You can be restored to the world of action and the hope of happiness.

It is significant, I think, that in spite of his insistence on the philosophical nature of his analysis, Ricoeur signs off with the suggestion that under the sign of the ultimate *incognito* of forgiveness can be found an echo of the word of wisdom uttered in the Song of Songs, 'Love is as strong as death'.[24] The terms 'Incognito' and 'echo' are safety nets here, but one senses that the sacred is not far off.

Part II

So how does Ricoeur make the final leap from impossible to possible forgiveness? How does he surmount the claim by Derrida, Arendt and Jankelevitch that forgiveness of radical evil is impossible? Acknowledging that such forgiveness is extremely 'difficult' (the title of his essay), Ricoeur ultimately seems to point to

[20] *Ibidem*.
[21] P. Ricoeur, *The Symbolism of Evil*, transl. E. Buchanan, Boston 1967.
[22] P. Ricoeur, *Memory, History, Forgetting*, p. 492.
[23] *Ibidem*, p. 493.
[24] *Ibidem*, p. 506.

a super-human origin of gift and capacity which belongs to the order of spirit and love, an order which observes a logic of surplus and superabundance. In short, what is impossible to humans – as Derrida rightly notes – is not impossible to God or, by extension, the divine capacity for renovation and rebirth which is the mark of the 'gap of creation', the miracle of origin, in each human being. Derrida too admitted that forgiveness is only possible for something or someone beyond the human, but he does not name a tradition of memory, faith or love to which one might adhere. He leaves the space of the 'inhuman' empty, without hermeneutic or practical bridge back to the human. Derrida does not sign off by citing the Song of Songs or giving the last word to love over death. Nor finally, does he give primacy to the origin of good over evil, restoration over rupture, reconciliation over aporia, happiness over angst. Perhaps it is a similar miracle of love that Derrida privately intends in his call for a messianic 'democracy to come'? But he does not say it and it is impossible to know.

Ricoeur, by contrast, makes his intentions clear even if he acknowledges the huge difficulties involved in moving from the impossible to the possible. First, he openly if gently confesses his adherence to the Jewish and Christian imagination of the suffering servant and the vertical height of forgiveness (it comes 'from above'). This is somewhat analogous, I would suggest, to the crucial move in AA where adherents incapable of controlling their lives hand over to 'a higher power' who in turns empowers them to do the impossible – unbind themselves as agents from the past acts of addiction, and thereby to realize that they are more than their past history and can be restored to a new capacity to begin.

Ricoeur also differs from Derrida, it seems to me, in acknowledging numerous ways in which the leap towards forgiveness can be prepared for, though never guaranteed or demanded as a law or method. One of these ways is the narrative power of exchanging memories and stories with one's enemies, those we cannot forgive.

In conclusion, let me say a few words about this hermeneutic of narrative pre-forgiveness. In an essay entitled 'Reflections on a New Ethos for Europe', Ricoeur outlines an *ethic of narrative hospitality* which may nurture a predispositon but by no means a guarantee of forgiveness. Forgiveness comes from beyond us, as Ricoeur insists, but humans may be more inclined to receive and offer this gift if they learn to love their enemies by exchanging narrative memories with them. This involves 'taking responsibility in imagination and in sympathy for the story of the other, through the life narratives which concern the other'.[25] In the case of genocide or famine memorials (I am thinking, for example, of the Holocaust and Irish Famine memorials side by side in Battersy Park, New York), this takes the form of an exchange between different people's histories such that we practice an art of transference and translation which allows us to welcome the story of the other – the memory of the stranger, the victim, the forgotten one.

This practice of narrative hospitality poses a particular problem in the limit case of hereditary hatred. Here, Ricoeur insists, there is no quick therapeautic fix

[25] P. Ricoeur, *Reflections on a New Ethos for Europe*, [in:] R. Kearney (ed.), *Paul Ricoeur: The Hermeneutics of Praxis*, London 1996, p. 7.

or exoneration but a difficult labor of attending to founding events that are not my own and, at times, to life stories that belong to my long sworn adversary. We are faced with the difficult task of learning to 'recount otherwise'[26]. But the best that such narrative hospitality can achieve is to serve as a 'secret alchemy' which may induce a certain 'disposition to consideration'. Such gestures of narrative imagination and empathy can sometimes lead to an exchange between a request and an offer of forgiveness. But this can never be institutionalized as a political right or duty. And questions of guilt and accountability are not suspended. At best, translating the stories of the other challenge of resisting the reification of an historical event into a fixed obsession by showing how each event may be told in different ways by narrators other than ourselves. Not that everything becomes relative and arbitrary. On the contrary, acts of trauma and suffering call out for justice, and the best way of achieving this is often to invite empathy with strangers and adversaries by allowing for a plurality of narrative perspectives. The resulting overlap may thus lead to what Gadamer calls a 'fusion of horizons' where diverse horizons of consciousness may at last find some common ground.[27] A reciprocal transfer between opposite minds. 'The identity of a group, culture, people or nation, is not that of an immutable substance', writes Ricoeur, 'nor that of a fixed structure, but that, rather, of a recounted story'. A hermeneutic exchange of stories effectively resists an arrogant or rigid conception of cultural identity which prevents us from perceiving the radical implications of narrative hospitality – namely, the possibility of 'revising every story which has been handed down and of carving out a place for several stories directed towards the same past'.[28] Of course while this model of narrative hospitality may work in historical conflicts like Northern Ireland, The Balkans or South Africa, it is not easily applied to limit situations like the Holocaust. For while a plurality of narratives by the victims is desirable (as Primo Levi says, the story must be told again and again so that the holocaust never be repeated), a plurality of narratives by the perpetrators – unless explicitly expressing apology, guilt and remorse – can easily lead to relativism or revisionism. And there are other cases of genocide where a reciprocal exchange of memories is equally difficult. One thinks of the Armenian genocide in Turkey? Might it ever be possible for an open exchange of memories between Turks and Armenians to bring about some kind of reconciliation, preparing eventually for the miraculous 'incognito of forgiveness'? Or for a narrative hospitality between Jews and Arabs?

A plurality of narratives should increase not diminish respect for the singularity of the events narrated through the various acts of remembering. It might even be said to increase our sense of awareness of such events, especially if it is foreign to us in time, space or cultural provenance. '*Recounting differently* is not inimical to a certain historical reverence to the extent that the inexhaustible richness of the event is honored by the diversity of stories which are made of it, and

[26] P. Ricoeur, *Memory, History, Forgetting*, p. 477.
[27] H.-G. Gadamer, *Truth and Method*, transl. G. Barden, J. Cumming, London 1975.
[28] P. Ricoeur, *Reflections...*, p. 7.

by the competition to which that diversity gives rise'.[29] And Ricoeur adds this critical point: 'The ability to recount the founding events of our national history in different ways is reinforced by the exchange of cultural memories. This ability to exchange has as a touchstone the will to share symbolically and respectfully in the commemoration of the founding events of other national cultures, as well as those of their ethnic minorities and their minority religious denominations'.[30] When it comes to the question of reconciliation and forgiveness this point applies particularly to events of pain and trauma (as in Famine or war memorials). And here again it is not a question of guaranteeing pardon but, as Ricoeur reminds us, of 'an exchange between a request and an offer, in which the unforgivable begins to be chipped away'.[31] I think the term 'chipped away' is critical here. It is a matter of a long working-through not some cheap therapeutic magic.

Narrative hospitality may also prepare for forgiveness in so far as it allows for a retrieval of the betrayed promises of the past, so that we may respond to our 'debt to the dead' and endeavor to give them a voice. The goal of narrative retrieval is, therefore, to try to give a future to the past by remembering it in the right way, ethically and poetically. A crucial aspect of reinterpreting transmitted traditions is the task of discerning past promises which have not been honored. For 'the past is not only what is bygone – that which has taken place and can no longer be changed – it also lives in the memory thanks to arrows of futurity which have not been fired or whose trajectory has been interrupted'.[32] In other words, the unfulfilled future of the past may well signal the richest part of a tradition – its unactualized *possibilities*; and the emancipation of 'this unfulfilled future of the past is the major benefit that we can expect from the crossing of memories and the exchange of narratives'.[33] It is especially the founding events of a community – traumatic or dramatic – which require to be reread in this critical manner in order to unlock the potencies and expectancies which the subsequent unfolding of history may have forgotten or travestied. This is why narrative hospitality often involves a recovery of some seminal moment of suffering or hope, of the repressed traumas or impeded promises which are all too often occluded by Official History. 'The past is a cemetery of promises which have not been kept', notes Ricoeur. And narrative hospitality can, at best, offer ways of 'bringing them back to life like the dry bones in the valley described in the prophecy of Ezekiel'.[34] And, for Arendt as for Ricoeur, promising is the other side of forgiving, as it opens history to natality and enables agents to begin again.

One of the ultimate goals of narrative hospitality between enemies is *pardon.* Though the goal is not of the order of teleology but of eschatology, nor to necessity but of surprise. And here again we encounter the boundary situation of unforgivable guilt and the possibility of 'something other' which might make impossible

[29] *Ibidem*, p. 8.
[30] *Ibidem*, p. 9.
[31] P. Ricoeur, *Memory, History, Forgetting*, pp. 477–478.
[32] *Ibidem*, p. 8.
[33] *Ibidem.*
[34] *Ibidem*, p. 9.

forgiveness possible. If empathy and hospitality towards others are crucial steps in the ethics of remembrance there is something *more* – something which entails moving beyond narrative imagination to forgiveness. In short, the exchange of memories of suffering demands more than sympathy and duty (though these are essential for any kind of justice). And this something 'extra' involves pardon in so far as pardon means 'shattering the debt'. Here the order of justice and reciprocity can be supplemented, but not replaced, by that of 'charity and gift'. Such forgiveness demands huge patience, an enduring practice of 'working-through', mourning and letting go. But it is not a forgetful forgiveness. Amnesty can never be based on amnesia. It remembers our debt to the dead while at the same time introducing something other, something difficult almost to the point of impossibility, but something all the more important for that reason. One thinks of Brandt kneeling at Warsaw, Havel's apology to the Sudeten Germans, Hume's preparedness to speak with the IRA, Sadat's visit to Jerusalem, Hillesum's refusal to hate her hateful persecutioners. All miraculous moments where an ethics of reciprocity is touched and transfigured by a poetics of pardon. The leap made. But I repeat: one does not replace the other – *both* justice *and* pardon are equally important in the act of remembering past trauma. Ricoeur insists on this point. 'To the degree that charity exceeds justice we must guard against substituting it for justice. Charity remains a surplus; this surplus of compassion and tenderness is capable of giving the exchange of memories its profound motivation, its daring and its momentum'.[35]

When we dare to listen to the stories of enemies or strangers, to other peoples and communities not our own, are we not suddenly all famine sufferers, all genocide victims, casualties of the Vietnam war – at least for a special, fleeting moment? A moment, out of time yet also in time, that bears the trace of the incognito of forgiveness?

CONCLUSION

We return finally to the limit situation of evil which serves as abyssal opposite to the gift of forgiveness. Unforgivable Evil is not just something we struggle against. It is also something we undergo. To ignore this passivity of evil suffered is, Ricoeur concludes, to ignore the extent to which evil strikes us as shockingly strange and disempowering. One of the wisest responses to evil is, on this count, to acknowledge its traumatizing effects and work-them-through *(durcharbeiten)* as best we can. Practical understanding can only redirect us toward action if it has already recognized that some element of estrangement almost always attaches to evil, especially when it concerns illness, horror, catastrophe, or death. No matter how prepared we are to make sense of evil, we are never prepared enough. That is why the "work of mourning" is so important as a way of not allowing the inhuman nature of suffering to result in a complete "loss of self" (what Freud called "melancholia"). For without selfhood no pardon could be possible. Some kind of catharsis is necessary to prevent the slide into fatalism that all too often issues in

[35] *Ibidem*, p. 11.

despairing self-destruction. The critical detachment brought about by cathartic mourning elicits a wisdom that may turn *passive lament* into the possibility of *active complaint*, that is, *protest*.[36] Though protest is, of course, not yet pardon.

Here narrative testimonies, mentioned above, may help the victim to escape the alienation of evil, that is, to move from a position of mute helplessness to a form of self-renewal. Some kind of narrative working-through is necessary, it seems, for survivors of evil not to feel crippled by grief or guilt (about the death of others and their own survival) nor to succumb to the game of the "expiatory victim" which makes pardon impossible. What the catharsis of mourning-narrative allows is that new actions and responses – including pardon – are still possible *in spite of evil suffered.* It detaches us from the obsessional repetitions and repressions of the past and frees us for a future. For only in unleashing the agent from the act and the victim from the evil – in the miracle of secret pardon – can one escape the disabling cycles of retribution, fate, and destiny: cycles which alienate us from the possibility to forgive by instilling the view that evil is overpoweringly alien – that is, irresistible.

Working-through the experience of evil – narratively, practically, cathartically – helps us to take the paralyzing allure out of evil. And so doing it enables us to remain open to the incognito gift of pardon. Working-through is central to an anthropology of capability and an ontology of act and potency in that makes evil *resistible*. In sum, by transforming the alienation and victimization of lament into 1) a moral response of just struggle, 2) opening the possibility of a spiritual response of forgiveness, we refuse victory to evil, declaring love as strong as death.. But while narrative working-through, testimony and catharsis may bring us to the threshold of pardon, they cannot cross it of their own momentum. It can predispose us to the gift of forgiveness but it cannot deliver it.

Something 'more' is required. Radical evil calls for an answering power of radical good. Against the 'never' of evil, which makes pardon impossible, we are asked to embrace what Ricoeur calls the 'marvel of a once again' which makes it possible.[37] But the possibility of forgiveness is a 'marvel', we noted, precisely because it surpasses the limits of rational calculation and explanation. There is a certain gratuitousness about pardon due to the very fact that the evil it addresses is not part of some dialectical necessity. Pardon is something that makes little sense before we give it but much sense once we do. Before it occurs it seems impossible, unpredictable, incalculable in terms of an economy of exchange. There is no science of forgiveness. And yet this is precisely where hermeneutic sensibility, attentive to the particularity of specific evil events, joins forces with the practice

[36] P. Ricoeur, *Memory and Forgetting*, [in:] *Questioning Ethics*, pp. 5–12. See also P. Ricoeur, *Evil: A Challenge to Philosophy and Theology*, [in:] P. Ricoeur, *Figuring he Sacred: Religion, Narrative and Imagination*, Indianapolis 1995, pp. 250 f. See also our own analysis of this theme in *Evil, Monstrosity and the Sublime*, [in:] R. Kearney, *Strangers, Gods and Monsters*, pp. 83 f.

[37] P. Ricoeur, cited in R. Kearney *Evil, Monstrosity and the Sublime*, pp. 105–106. See also W. Desmond, *Beyond Hegel and Dialectic*, Albany 1992, pp. 238–239. And for a comparative and contrasting 'Eastern' perspective on the topic of pardon as it relates to a number of contemporary political situations of violence and war, see J.S. O'Leary, 'Buddhism and Forgiveness', *Japan Mission Journal* 56 (2002), pp. 37–49.

of patient working-through – their joint aim being to ensure that past evils might be prevented from recurring. Such prevention calls for pardon as well as protest in order that the cycles of repetition and revenge give way to future possibilities of non-evil. This is a good example of Ricoeur's claim that pardon gives a future to the past.

Cathartic narration can, Ricoeur concludes, help to make the impossible task of pardon that bit more possible without ever allowing amnesty to fall into amnesia. The past must be recollected and worked-through so that we can identify, what it is that we are forgiving. For if pardon is beyond reason, it is never as blind or mad as Derrida suggests. And if it is mobilized by the gratuity of love – which calls for that element of extra – it is never insensitive to the logic of justice. Or to put it in Pascal's terms, pardon has its reasons that reason cannot comprehend. Perhaps only a divinity could forgive indiscriminately. And there may indeed be some crimes that a God alone is able to pardon. Even Christ, as Ricoeur notes, had to ask his Father to forgive his crucifiers: 'Father forgive them for they know not what they do'. As man alone he could not do it. Impossible for us, possible for God. But here an ethics of pardon approaches the threshold of a religious hermeneutics.

But, finally, what kind of religious hermeneutics are we talking about? In his essay on evil and in the essay on pardon in *Memory, History, Forgetting*, Ricoeur seems to work within an exclusively Judeo-Christian tradition. But in his last testament, *Vivant jusqu'à la mort*, Ricoeur extends the horizon of the sacred that makes possible (God as *Posse* as he puts it) to all great wisdom traditions, amounting to a call for a radically interconfessional hospitality. Here too there is need for pardon, to forgive the great crimes committed by one religion against another in history. And so in this confessional testimony, which uncharacteristically bridges the divide between the philosophical and the theological, Ricoeur speaks of a 'grace' which takes the form of an 'intimate transcendence which rips through the veils of confessional religious codes'.[38] Some might suggest that Ricoeur is approximating here to Derrida's anonymous structure of messianicity, of a religion without religion, an Other without face, tradition or voice. But I think not. For while the advent of such an Other is impossible for Derrida, for Ricoeur it is a sacred marvel that makes the impossible possible in each lived moment that pardon is given or received.

[38] P. Ricoeur, *Vivant jusqu' à la mort*, Paris, 2007, p. 45.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

URSZULA LISOWSKA
University of Wrocław

How to appear to be perceived ethically – Martha Nussbaum's ethical conception from the perspective of political aesthetics

Abstract

The paper discusses Martha Nussbaum's ethical project from the perspective of political aesthetics, i.e. the reflection on the political significance of perception as such and the perception of beauty in particular. The article attempts to demonstrate that a full account of this type should allow for the relationship between perception and certain acts of appearing. It analyses the strengths and weaknesses of Nussbaum's project considered in such a context. This involves the reconstruction of Nussbaum's conceptions of human good and practical rationality and the search for an interpretation of society compatible with these assumptions.

Introduction – perceiving and appearing

Martha Nussbaum's contribution to the contemporary socio-political philosophy is unquestionable, the philosopher being chiefly associated with what has become known as the capabilities approach. This paradigm has been worked out by Nussbaum and Amartya Sen as an alternative to the dominating methods of assessing the quality of life, such as utilitarianism, human rights' perspectives or the Rawlsian conception of primary goods. The two thinkers argued that, instead of analysing the sum total of utility, legal guarantees or the distribution of goods, the research should be focused on people's actual possibilities of functioning – i.e. capabilities. It is this comparative use of the notion that is of main interest for Sen, an awardee of the Noble Prize in Economic Sciences. Nussbaum, however, places the idea of capabilities in a broad philosophical context, infusing it with a deep anthropological meaning. For her, capabilities first and foremost constitute a concept for human good, which, in turn, can be employed on the socio-political level. In other words, she starts with an account of a good human life, which comprises the set of basic entitlements expressed in the language of capabilities.[39]

[39] M.C. Nussbaum, *Women and Human Development. The Capabilities Approach*, Cambridge 2001, pp. 11–15.

Thus, Nussbaum forges an intricate project which spreads from the realm of anthropology towards social and political philosophy. Such a scope of reflection may seem intimidating. Yet, the philosopher manages to construe a coherent concept, whose unity, I suggest, can best be grasped by interpreting it from the perspective of political aesthetics. This is taken to mean the attribution of political significance to perception as such (from Greek: *aisthēsis* – perception) and, more specifically, to the perception of beauty. In other words, it is the analysis of the manner in which human beings as members of society perceive each other and the relevance of the notion of beauty for this process. Nussbaum's project can be regarded as an example of this approach. Drawing on her account of human good, Nussbaum claims that our practical rationality rests on the capability of the respectful perception of others. As such, perception is supposed to provide the link between anthropological and socio-political levels of her project. It is the manner in which humans, as beings equipped with a certain type of rationality, should attempt to approach each other in order to create a just community.

It could be objected, however, that perception itself is not an independent process, since the completing side of the act of seeing is the act of appearing. From an epistemological point of view, the degree of the activity of an appearing "object" could be disputed, it is reasonable, though, to assume that in the social realm perception involves an interaction between two active beings. Thus, my perception of another person is partly determined by the manner in which she appears to me, the appearing itself being irreducible to my own manner of approaching her. A comprehensive politico-aesthetic concept should comprise both of these elements.

I shall argue that, although Nussbaum's project offers a good background for such analysis, the philosopher focuses too much on the issue of perception at the cost of appearing. This turns out to pose certain limitations on the intended socio-political application of perception as a method of public reasoning. I suggest that the success of this philosophical enterprise depends on finding an account of society which would do justice to both perception and appearing. In order to meet this objective, I propose that we proceed in the following manner. First, Nussbaum's account of human good has to be presented. This will help us place perception and the problem of appearing within Nussbaum's anthropological considerations. Due to the specificity of Nussbaum's conception of human good, this will at the same time point to the socio-political implications of these phenomena. Next, an outline of a conception of society compatible with Nussbaum's assumptions will be provided. I will conclude with a few comments on an exemplification of this approach, which can be found in Wrocław's successful application for the title of European Capital of Culture.

Capabilities approach – humans as social and rational animals

As I have suggested, Nussbaum's concept rests on certain anthropological ideas. That these should have socio-political implications is clear from the outset since the philosopher underlines their value-laden character. Reluctant to rely on any metaphysical – i.e. objective and unchangeable – notion of human good, she opts

for what she describes as internal essentalism (as opposed to external essentialism).[40] This approach requires tracing the concept of 'human good' back to our own interpretations of what it means to be human. Nussbaum models it on the Aristotelian method of *phainomena*-based inquiry, *phainomena* (appearances) being not so much "pure" experiential data (as most of the translations would have it) as common sense knowledge. i.e. widespread interpretations of reality.[41] Thus, we find the first hint at the aesthetic dimension of Nussbaum's conception. The notion of human good comprises all the elements which *appear* to us as necessary to lead a good, fully human life. It is anthropocentric – based on our own *perception* of humanity – and social – imbedded in common sense beliefs.[42]

In order to find out where this quality of Nussbaum's thought stems from, we have to analyse the conception of human good in greater detail. Two levels could be distinguished within it – the features of "the shape of the human form of life" which we deem essential and the desirable ways of their development, i.e. human functional capabilities.[43] The former function as facts about the human condition (their choice, however, is already a matter of interpretation), which then are subjected to evaluative reflection. What stands out in this account of humanity is probably its focus on our animality.[44] Significantly, on Nussbaum's list, mortality and human body with its basic needs precede cognitive, social and cultural capacities. She also includes in this list the specificity of human infant development and the relationship with nature in general. Such insistence on human essential bodiliness is the characteristic mark of Nussbaum's philosophy and the core of her project. Including animality in the notion of human good, the philosopher underlines that we are needy, vulnerable creatures. These aspects of our condition are present in all spheres of our functioning, which is partly what the notion of capabilities stands for.

This takes us to the second level of the conception of human good. The set of human functional capabilities represents the criteria of a life which seems worth living.[45] Nussbaum has chosen to construe it in the language of capabilities, basing on the description of the human condition.[46] Capabilities are, as we have

[40] See M.C. Nussbaum, 'Human Functioning and Social Justice. In Defense of Aristotelian Essentialism', *Political Theory* 20 [2] (1992), pp. 205–214; *The Therapy of Desire. Theory and Practice in Hellenistic Ethics*, Princeton 2009, pp. 29–32.

[41] See: M.C. Nussbaum, *The Fragility of Goodness. Luck and Ethics in Greek Tragedy and Philosophy*, Cambridge 2001, pp. 240–245

[42] *Ibidem*, pp. 290–294.

[43] M.C. Nussbaum, *Human Functioning...*, pp. 216–222.

[44] Nussbaum characterises the shape of the human form of life by the following elements: mortality, the human body, capacity for pleasure and pain, cognitive capability: perceiving, imagining, thinking, early infant development, practical reason, affiliation with other animals, relatedness to other species and to nature, humor and play, separateness (*ibidem*, pp. 216–220).

[45] Nussbaum describes the two levels of her notion of human nature as two thresholds. The lower threshold of the shape of the human form of life expresses "the bare minimum" which has to be met in order for a life to be considered as human at all. The second threshold provides a higher standard of a desirable life in which all the elements characteristics of a human life have the chance to flourish (*ibidem*, p. 221).

[46] The list includes the following items: life, bodily health, bodily integrity, senses, imagination, and thought, emotions, practical reason, affiliation (this includes the ability to form

said, defined as the possibilities of acting. When applied to the concept of human good, the concept denotes specifically human potentialities for functioning.[47] Thus, referring to Aristotle,[48] Nussbaum presents a teleological account of humanity, according to which human good is characterised not by a set of qualities but by certain developmental tendencies. This means that we are not complete, self-contained beings. On the contrary, human existence is a constant process aimed at the realisation of one's idea of a good life.

Significantly, such development requires not only maturity on the side of an individual – i.e. what Nussbaum calls "internal capabilities" to use an inborn equipment ("basic capabilities") – but also a facilitating environment. For, as the brief introductory remarks about Sen's and Nussbaum's socio-political contributions have already suggested, the notion of capabilities represents the actual life opportunities of individuals. I am capable of living healthily thanks to my innate constitution and good habits but also thanks to medical care; I am capable of making political choices if I am mature enough to form my own beliefs and if the surrounding political *regime* respects my opinion, *etc.* Thus, the capabilities which are of the greatest interest for Nussbaum (capabilities *par excellence*, we might say) are the so-called "combined capabilities", that is, internal capabilities coupled with external conditions necessary to exercise them. It is this type of capabilities which constitutes the list referred to.[49]

As we can see, then, the concept of capabilities is based on the assumption of human animality and the neediness inherent in it. It expresses an individual's reliance on external support for her flourishing, thereby pointing to one of the two distinctive features of the human type of animality, namely sociability. For Nussbaum, our social nature is therefore a reflection and extension of our lack of self-sufficiency and in this sense humans can be defined, Aristotelian-wise, as political animals.[50] We are naturally inclined to form interpersonal relationships, in which we seek the completion of our internal capabilities. However, at this point it could be objected that deriving sociability from the state of animal neediness can hardly deliver a distinctively human feature. For in this respect we are basically similar to other animals and the long period of infancy alone would not make a qualitative difference. What is crucial, though, is the other distinctive feature of the human type of animality which Nussbaum presents, namely rationality. Her understanding of this capacity is informed by the Kantian tradition with its synthesis of rationality and dignity. Thus, on the one hand, Nussbaum underscores

relationships with other people and protection from discrimination), relation to other species, play, control over one's political and material environment (see: M.C. Nussbaum, *Women and Human Development. The Capabilities Approach*, Cambridge 2000, pp. 78–80; *Upheavals of Thought. The Intelligence of Emotions*, Cambridge 2001, pp. 416–418.

[47] M.C. Nussbaum, *Women...*, pp. 71–72.

[48] M.C. Nussbaum, 'Nature, Function, and Capability', *Oxford Studies in Ancient Philosophy*, suppl. vol. 1 (1988), pp. 145–84.

[49] *Ibidem*, pp. 20–25; *Women...*, pp. 84–85.

[50] M.C. Nussbaum, *Frontiers of Justice. Disability, Nationality, Species Membership*, Harvard 2007, p. 159.

the autonomous value of each person as an autonomous end.[51] Dignity functions for her as the basic evaluative category, the consistency with it being the criterion of a good life.[52] On the other hand, she rejects the Kantian opposition of rationality-dignity versus bodiliness. Nussbaum believes that we are rational, and therefore dignified, *as* animals. Far from being a reductionist, she merely denies the existence of an ontological gap between our cognitive capacities and our animality.[53]

Thus, Nussbaum adds another element to her picture of a human animal. It is the combination of rationality and sociability that ultimately defines the specificity of our type of bodiliness. For a "truly human way" of living requires that we exercise all animal functions in a rationally planned manner and in cooperation with others.[54] At the same time, being essentially the features of animality, rationality and sociability are marked by the neediness characteristic of the bodily creatures that we are.[55] They are both expressive of an animal lack of sufficiency and, as such, interrelated. This means that, on the one hand, our rationality is essentially social. By defining rationality in terms of capabilities, Nussbaum suggests that our reflective capabilities develop in the social context in which we are embedded. Importantly, the list of combined capabilities includes not *theoretical* but *practical* rationality, i.e. the ability to define one's own life goal. This suggests that it is not abstract, theoretical thinking that defines humanity. The rationality of human animals is "garden-variety" and practice-oriented, scientific deliberations being based on everyday knowledge contained in *phainomena*.[56] Therefore, on the other hand, we are sociable as practically rational beings. Our mutual relations are imbued with ethical considerations – the questions of rightness and wrongness, goodness and evil.

Perception-based practical rationality

This lengthy introduction has, hopefully, helped us place the issue of practical rationality within the framework of Nussbaum's project. Rational and social animals are capable of reflection, which expresses their neediness and sociability, as well as their dignity. At this point, we may become aware of certain difficulty. For, on the one hand, practical rationality is, as we have said, defined as the ability to form one's own conception of a good life. Nussbaum's teleological account of human good is then rooted in the tradition of individualism and respectful of each person's right of self-determination. On the other hand, due to our lack of self-sufficiency, this right can be exercised only with some degree of external support (which is why Nussbaum rephrases it in terms of a capability). We might,

[51] *Ibidem*, pp. 70–71.
[52] M.C. Nussbaum, *Women...*, p.73.
[53] *Ibidem*, pp. 72–74; *Frontiers...*, pp. 159–160.
[54] M.C. Nussbaum, *Human Functioning...*, pp. 222–223.
[55] "Bodily need, including the need for care, is a feature of our rationality and our sociability; it is one aspect of our dignity, then, rather than something to be contrasted with it" (M.C. Nussbaum*Frontiers...*, p. 160).
[56] On the method of construing knowledge on the basis of *phainomena* see M.C. Nussbaum, *Fragility...*, pp. 245–263.

therefore, ask where the ethical character of our sociability stems from. Are not interpersonal bonds mere relations of the mutual exchange of services, likely to degenerate into exploitation?

This is where the conception of perception steps in. A proper manner of seeing is what it takes for a human animal to be practically rational, argues Nussbaum. In the first place, it requires that we perceive other persons as more than instruments to our own flourishing (that is, that we respect their dignity). This attitude then determines our ability to assess the situation in which an ethical judgment is made, such ability being what I shall call ethical perception.[57] Nussbaum's conception is a philosophical one, grounded in her account of human good and, as we shall see, yet more Aristotelian solutions. However, in order to present and exemplify some of its elements, Nussbaum turns to psychoanalytical narratives about the human maturational process. I suggest that we follow her along this path for awhile.

The narratives in question belong to the object relations theories of the development of the self, such as the works of D.W. Winnicott, W.R.D. Fairbairn and J. Bowlby.[58] Such choice complies with Nussbaum's insistence on the significance of interpersonal relations for practical rationality and helps to shed light on its specificity. The starting point of the object relations accounts is one of the elements featured in "the shape of the human form of life" list, namely early infant development. They begin with the state of extreme neediness "more or less unparalleled in any other animal species."[59] The weakness of the human bodily constitution – the lack of innate equipment comparable to that of other animals – results in a specific combination of helplessness and omnipotence.[60] For an infant narcisticly expects the world to revolve around her needs as it did in her mother's womb, this demand being coupled with an utter inability to cater for herself on her own.[61] At this stage, the self has not emerged yet and the infant does not differentiate herself from the world.[62] For this to happen, the sense of neediness is crucial. The infant gradually learns which of her needs she is capable of meeting on her own and which require external support. These experiences give her the sense of interior and exterior, initiating the development of the self.[63]

What is particularly important for practical rationality is that these recognitions take the form of emotions.[64] At this point, Nussbaum introduces her "neo-Stoic"[65] – as she describes it – philosophical conception of emotions. Drawing on the Stoic heritage, she presents a cognitive account of emotions as judgments about

[57] Nussbaum uses the notion of perception in the latter context, referring to the ability of grasping a given situation. However, the former dimension is, as I will argue, a more basic one. Nussbaum describes it in the language of seeing and appearing, which justifies the extension of the concept of perception onto these issues. In this context, perception stands for a more general ability to understand reality from a given perspective.

[58] M.C. Nussbaum, *Upheavals*..., p. 180.

[59] *Ibidem*, p. 181.

[60] *Ibidem*, p. 196.

[61] *Ibidem*, pp. 184–185.

[62] *Ibidem*, p. 190.

[63] *Ibidem*.

[64] *Ibidem*.

[65] *Ibidem*, p. 27.

the eudaimonistic value of external goods. That is to say, emotions are assertions that certain factors uncontrolled by an individual are crucial for her well-being. As such, they express the sense of dependency on the world and record the awareness of one's separateness from the surroundings. Developing simultaneously to the emerging of individual consciousness, emotions constitute the most fundamental ways of approaching reality.

This basic cognitive capability is, however, completed by another one, as the emotions' status of judgments suggests. Following Stoics, Nussbaum conceives judgments as appraisals of appearances (*phantasmata*)[66] suggested by the senses, in the case of emotions the criterion of the assessment being the eudaimonistic value. This means that emotions consist in taking the stance on (assenting to or rejecting) what the world looks to be like, depending on whether a given appearance has bearing on one's flourishing or not.[67] Thus, we once again encounter the category of appearances. Both *phainomena* mentioned above and *phantasmata* derive from the verb *phainesthai* – 'to appear'.[68] The cognitive capability behind them is what Aristotle called *phantasia*, usually translated as 'imagination'.[69] Imagination, then, is the general ability to make the world *appear* to us in a certain way, and therefore – to *see* things *as* other things.[70]

The italicised words point to three important features of the said capability. To start from the very last, imagination is inherently interpretational as it always approaches reality from a certain perspective. It transcends raw experiential data and attaches meaning to it. Emotions represent one of the possible angles of perceiving reality[71]. They provide "the map of the world", the landmarks being items crucial for our well-being[72]. As such, they involve the exercise of imagination since they are essentially acts of interpreting reality in terms of its eudaimonistic value. Such interpretation consists in perceiving reality through the lenses of the appearances suggested by our imagination. That is to say, we approach reality with certain prejudices (in this case an idea of a good life) and, as a result, the world appears to us in a given manner and is perceived accordingly.

[66] *Cf.* V. Caston, *Intentionality in Ancient Philosophy*, [in:] E.N. Zalta (ed.), *Stanford Encyclopedia in Philosophy*, Stanford 2007; http://plato.stanford.edu/entries/intentionality-ancient/#6, [retrieved: 2.10.2013].

[67] M.C. Nussbaum, *Upheavals*... pp. 37–38, *Therapy*..., pp. 374–375.

[68] *Ibidem*, p. 85; *Fragility*..., p. 240.

[69] M.C. Nussbaum, *The Discernment of Perception: An Aristotelian Conception of Rationality* [in:] M.C. Nussbaum, *Love's Knowledge. Essays on Philosophy and Literature*, Oxford 1992, p. 77 [hereinafter referred to as *An Aristotelian*..,]. In the context of Stoic epistemology, *phantasmata* are also translated as "apparitions" or "impressions" (see: V. Caston, *Intentionality in Ancient Philosophy* and N.J.T. Thomas, *Mental Imaginary. From the Hellenistic to the Early Modern Era*, [in:] *Stanford Encyclopedia in Philosophy*, http://plato.stanford.edu/entries/mental-imagery/hellenistic-modern.html [retrieved: 2.10.2013]). Nussbaum's choice of translation points to the common root of common sense knowledge and individual interpretations of reality. They are both the expressions of our ability to imagine the world to be such and such. *Phainomena* have an intersubjective character but they employ the same mechanism and *phantasmata*.

[70] On Nussbaum's interpretation of the Aristotelian idea of imagination see M.C. Nussbaum, *Essay V* [in:] M.C. Nussbaum, *Aristotle's De Motu Animalium*, Princeton 1985, pp. 221–269.

[71] M.C. Nussbaum, *Upheavals*..., p. 27.

[72] *Ibidem*, pp. 206–207.

Thus, Nussbaum conceives appearances as the derivatives of the human ability to interpret reality. Rather than external stimuli affecting the process of perception, they are its parts, rooted in the perceiver's perspective of approaching reality. Such stance reveals Nussbaum's tendency to focus on perception alone, which, as I have suggested at the beginning, is problematic from the socio-political point of view. Before we discuss this issue however, we need to find out what makes this aesthetic – perception-based – account of practical rationality relevant to political issues.

From perception to sociability

In order to settle these questions, we have to turn to the narrower of the two meanings of aesthetics mentioned in the beginning, namely the question of the perception of beauty. This requires that we go back to our infant, whom we have left on the threshold of personal identity. We already know that the basic sense of her separateness from the world is recorded in emotions, which are grounded in the experience of neediness. Emotions are the acknowledgments of her dependency on certain external goods requisite for her flourishing. What Nussbaum finds particularly inspiring in object relations theories is that their representatives recognise a distinct need for security (Bowlby) or a "facilitating environment" (Winnicot), irreducible to bodily needs.[73] This, in the light of Nussbaum's belief in human inherent sociability, could be interpreted as a protosocial drive. The infant feels the need for stable, secure surroundings and these are guaranteed by the permanent presence of caretakers.[74] Thus, the lack of sufficiency motivates the first interpersonal bonds, which complies with Nussbaum's insistence on the interrelation between practical rationality and sociability. Furthermore, the relationship with caretakers gives the child the sense of security, which encourages her to reflexively "turn inward, discovering her own personal life", "inner depth or creativity."[75] Therefore, her sense of identity is always a reaction to somebody else's presence and care. As such, it involves the experience of being in relationships with others.

Since the child's reliance on other people is the expression of her neediness and she values them as external goods, her emotional attachments involve the sense of a certain independence from the other person. The child realises that her caretakers are separate entities with lives of their own which she cannot fully control.[76] For this reason there always seems to be an element of wonder in interpersonal bonds. Wonder stands out in Nussbaum's concept because it is "as non-eudaimonistic as an emotion can be,"[77] being the recognition of the intrinsic value of an object without direct reference to one's idea of a good life.[78] Although in the case of intimate relationships, it is accompanied by other emotions which assert the importance of a person for the perceiver's well-being, wonder itself is

[73] *Ibidem*, pp. 185–186.
[74] *Ibidem*, p. 187.
[75] *Ibidem*, p. 208.
[76] *Ibidem*, p. 209.
[77] *Ibidem*, p. 55.
[78] *Ibidem*, pp. 54–55.

the grasping of her autonomous worth, i.e. dignity. Thus, the acknowledgment of another person's dignity could be regarded as an act of aesthetic contemplation. She is perceived as valuable "in her own right,"[79] endowed with non-instrumental value, in other words – beautiful.[80] This also involves imagination, which helps us interpret other people's behavior as the expression of their activity and gradually recognise unique, active persons behind externalized behaviour.[81]

We discover that, on the one hand, the ethical and not merely exploitative character of social bonds is connected with the aesthetic dimension of our cognition. Individuals approach each other with reverence – characteristic for the contemplation of beauty, employing their interpretative capabilities in the attempt at reconstructing each other's inner lives. They are ends in themselves, who perceive their ends (that is – themselves) as essentially related to (the ends of) other individuals. The ability to imagine co-constitutes interpersonal bonds of mutual respect, which can later become the basis for common sense knowledge. Therefore, on the other hand, the social character of practical rationality is essentially connected with its perceptional dimension. Perception is where sociability and rationality meet.

Ethical perception – literature and public life

So far we have concentrated only on intimate relationships, such as those between a child and her caretakers. To complete the presentation of Nussbaum's project as a politico-aesthetic one, we need to demonstrate that the above account of practical rationality can be applied on a more general scale. In order to achieve this, Nussbaum asks what it means for rational animals to perform ethical considerations. She underscores that our reflection always expresses the condition of beings both dignified and needy. We are concrete individuals, who make ethical judgments in a particular situation determined by our eudaimonistic projects and the surrounding network of interpersonal relationships. Therefore, Nussbaum advocates the principle of "the priority of the particular", which she again associates with Aristotle.[82] In this situation, an ethical judgment cannot simply be deduced from general principles. These are of course useful, but they function as the Lesbian rule (a form of measurement used on Lesbos), which "¿¿bends to the shape of stone¡¡."[83] What is important is that "the bending" takes the form of perceiving, since, as Nussbaum often repeats after Aristotle, "the discernment rests with perception."[84] This is what Nussbaum calls (ethical) perception – the ability to grasp a situation in its intricacy.[85] It involves forming a judgment, after

[79] *Ibidem*, p. 237

[80] *Ibidem*, p. 54.

[81] M.C. Nussbaum, *Poetic Justice: The Literary Imagination and Public Life*, Boston 1995, pp. 36–46.

[82] M.C. Nussbaum, *An Aristotelian...*, pp. 66–75; *Introduction: Form and Content, Philosophy and Literature*, [in:] M.C. Nussbaum, *Love's Knowledge. Essays on Philosophy and Literature*, Oxford 1992, pp. 37–40.

[83] M.C. Nussbaum, *An Aristotelian...*, p. 70.

[84] *Ibidem*, pp. 55, 66.

[85] See, for example, M.C. Nussbaum, *Perceptive Equilibrium: Literary Theory and Ethical*

taking into account multiple (and often conflicting) elements which are involved in this particular situation: the people concerned, the history of the relationships with them, proper general principles, *etc*. It is a kind of tact – as some translators have suggested[86] – or sensitivity to the complexity of the context.

As such, it rests on the cognitive capabilities just mentioned – emotions and imagination. Emotions point to the individual important elements of the situation, helping her differentiate between neutral and valuable issues. When emotionally involved, she is more responsive to other people,[87] whose behavior she attempts to interpret imaginatively. Imagination, as the capability of "producing" appearances, is also concrete and synthesising. Due to this capacity, the situation appears to the individual as a unique combination of various elements.[88] Thus, emotions and imagination first shape our idea of a good life as inherently connected with the good of others, and then help us form ethical judgments in accordance with the specificity of our eudaimonistic projects. Ethical perception provides the model of practical considerations compatible with this condition. It takes into account both human dignity (reflected in each person's separateness and right to self-definition) and neediness (expressed in relationships between people). This involves viewing people as unique, flesh and blood but also as interdependent beings.

Importantly, Nussbaum finds the best exemplifications of ethical perception in classical realist and psychological novels, such as the works of Charles Dickens or Henry James. Such means of novelistic expression as the focus on concrete individuals and their inner life, diachronicity and the appeal to imagination enable these texts to represent our perspective of practical reasoning.[89] For Nussbaum, then, novels have a significant ethical dimension. This claim would probably shock only postmodern literary theorists[90] if it were not for the original implications which Nussbaum draws from it – namely, the philosopher presents literary texts and, consequently, the method of ethical perception displayed in them as the models of socio-political deliberations.[91] Nussbaum believes that considerations in the sphere of public policy should respect the values reflected in ethical perception, so vividly portrayed in novels. This means that each person ought to be treated as a unique, dignified if needy being, not only in intimate relationships but also on the socio-political level. In this manner, ethical perception can have the function mentioned at the very beginning, i.e. that of providing the link between different

Theory, [in:] M.C. Nussbaum, *Love's Knowledge. Essays on Philosophy and Literature*, Oxford 1992, pp. 176–186. Although Nussbaum does not use the adjective "ethical", I propose this addition in order to underline that it is the application of a more general capability in the process of forming a particular ethical judgment.

[86] See the translator's note in the Polish edition of Aristotle's *Politics* (Arystoteles, *Polityka*, tłum. D. Gromska, Warszawa 2007, p. 121).

[87] M.C. Nussbaum, *An Aristotelian...*, pp. 78–79.

[88] *Ibidem*, pp. 77–78.

[89] See, for example, *ibidem*, pp. 84–93.

[90] On the relationship between ethical and literary theory see: M.C. Nussbaum, *Perceptive Equilibrium...*, pp. 169–172, 190–193.

[91] See: M.C. Nussbaum, *Poetic Justice...*, pp. 1–12; *Perception and Revolution: "The Princess Casamassima" and the Political Imagination*, [in:] M.C. Nussbaum, *Love's Knowledge. Essays on Philosophy and Literature*, Oxford 1992, pp. 195–219.

dimensions of Nussbaum's concept. It is a method of applying an account of human good in the socio-political sphere.

Compassion and wonder – the priority of appearing

Does, however, perception alone suffice to cultivate the mutual respect on a socio-political scale? This might seem problematic. Ethical perception reflects the conception of rationality according to which each person is entitled to define her own idea of a good life, which also comprises the good of other people. They are perceived as important for an individual's well-being due to their inherent worth. Still, the judgment of intrinsic value and the eudaimonistic judgment constitute two sides of the same coin. In other words, it is only in connection with my own flourishing that other people enter the into the circle of my concern. Thus, the basic question to be settled if Nussbaum's concept is to be truly politico-aesthetic, is whether our perception can be so extended as to transcend intimate commitments. Are we capable of caring for people outside the most direct relationships in which we are embedded? If so, what can make us sensitive to their well-being?

All these difficulties are reflected in Nussbaum's discussion of the emotion of compassion.[92] It is a reaction to harm experienced by another person, which is judged to be serious, undeserved and of eudaimonistic value for the judging individual.[93] This suggests that compassion has political relevance since it involves some sense of the community of human condition. The judgment of seriousness presupposes a universal idea of 'human flourishing' (expressed in Nussbaum's list of capabilities), which helps us appreciate the gravity of somebody else's suffering. As such, compassion seems to be potentially addressed to all humanity. At the same time, however, due to its eudaimonistic character, it is prone to the limitations just mentioned. For Nussbaum, then, compassion constitutes the crucial socio-political issue. The success of her project depends on our capability to compassionately approach people outside our intimate attachments.[94]

Since the condition of compassion is the sense of "a common form of life,"[95] what seems to be at stake here is the notion of society. The borders of compassion are the borders of our conception of society, therefore we need to find an account of society which would meet the universal claims of Nussbaum's project. Nussbaum is inclined to reason in terms of a type of the Rawlsian Original Position argument. Namely – to recapitulate this construction in a simplified manner – she assumes that people are more likely to care for individuals who they conceive as co-members of a group governed by the same rules. Since they can never be sure in which social position they may one day find themselves, they feel concern for people of all walks of life, any of which is potentially theirs.[96] Therefore, society is envisaged

[92] *Ibidem*, pp. 36–38.

[93] M.C. Nussbaum, *Upheavals...*, pp. 306–321. This definition could also be expressed in a more general, hypothetical manner: compassion is the judgment that, were certain harm to befall somebody, it would be undeserved deprivation of goods crucial to her well-being.

[94] *Ibidem*, pp. 420–421.

[95] *Ibidem*, p. 422.

[96] *Ibidem*, p. 320

as a group integrated by common interest – the search for just principles. The sense of a shared goal of this type, however, is not an autonomous argument. The Original Position experiment is based on the principle of equality before law and this is granted only to those individuals who are already considered rightful members of the group. Thus, we discover that the idea of "a common form of life" is a politico-aesthetic notion. The question of social inclusion (that is – the broadening of compassion) is the question of our capability to perceive other people as our fellows.

This turns out to be dependent on certain extra-perceptional stimuli. Here Nussbaum's concept seems to shift from appearances understood as the constructs of our imagination towards what I propose to describe as acts of appearing. Nussbaum touches upon this problem when she reflects upon two compassion-aiding factors. The first of them is the judgment of similar possibilities, which reinforces the eudaimonistic judgment. If another person appears to me as subject to the same vulnerabilities as myself, I am likely to sympathise with her. The apprehension of our common neediness steers concern in me, thereby extending my eudaimonistic judgement towards her.[97] In this way, the manner in which she appears to me inspires compassion, which is a type of perception. Thus, perception is posterior to appearing, the latter being not so much my means of interpreting reality as an expression of another person's activity. Her ability to project an image of herself influences my attitude towards her.

On Nussbaum's account, however, individuals are not only vulnerable but also dignified. This assumption motivates the second compassion-aiding factor, namely the emotion of wonder. As we have seen, it has a special and important place in Nussbaum's concept. And it is its non-eudaimonistic character that creates its socio-political relevance. For, although in the context of intimate bonds, wonder is always completed by the awareness of the eudaimonistic importance of another person – this is not part of the emotion itself. This suggests that wonder can operate independently of our eudaimonistic prejudices or rather – reverse the order of the process of making a eudaimonistic judgment. Wonder is the enchantment with an object, which in this case is another person. It is a response to her act of appearing as beautiful, i.e. inherently worthy. As such, she becomes included in the perceiver's vision of a good life.[98] In this way, wonder can inspire compassion on a broader scale. We recognise the dignity of others and start to care for them for the sake of their intrinsic value. As in the case of similar possibilities judgment, another person presents herself to the perceiver in a certain manner (as a dignified being), which inspires wonder and this emotion, in turn, aids compassion.[99] Adding both of the elements together, we can say that compassion requires that people appear to one another as both vulnerable and dignified. That is to say, they should present themselves as endowed with capabilities, i.e. as beings whose rationality involves fragility.

[97] *Ibidem*, pp. 318–319.
[98] *Ibidem*, pp. 54–55.
[99] *Ibidem*, pp. 321–322.

Novels versus tragedies

If Nussbaum's aesthetics (the conception of perception in general and the perception of beauty) is to be political, then, the self-projecting activity of individuals should be allowed for. In order to define the basic features of the account of society compatible with this requirement, certain revision of Nussbaum's literary inspirations might be helpful. As we have seen, the philosopher gladly points to the socio-political relevance of realist and psychological novels, which according to her, provide the best examples of ethical perception at work. Doubtlessly, they can also teach its readers compassion and wonder. After all, reading is essentially a disinterested participation in the beautiful realm of art. Books display the beauty of their protagonists, who thereby inspire wonder in readers.[100] Furthermore, novels can uncover the familiarity of the lot of individuals belonging to underprivileged groups. When readers recognise their own vulnerabilities in the lives of excluded members of society, they are likely to feel compassion towards them. This is how the novels such as Richard Wright's *Native Sun*, devoted to the problem of racial hatred, or E.M. Forster's *Maurice*, picturing a male homosexual couple, can work.[101]

The importance of novels in this respect, however, should not be overestimated. A novel is a certain object (incomplete and interpretation-dependent though it may be), and interaction with it is an intellectual play between the novel and the reader. As such, it involves the operation of imagination and appearances rather than an encounter with acts of appearing. In other words, it does not leave space for the self-projecting activity of individuals. Consequently, it does not prepare readers to open themselves up to the behaviour of the real-life counterparts of fictional characters. And since it is this type of interaction between people that can inspire the broadening of compassion, it seems that a novel's significance in this regard is limited.

However, there is another literary genre which Nussbaum likes to turn to. Although the philosopher usually refers to it in connection with her idea of ethical theory, it is relevant to socio-political issues as well. Ancient tragedies – for this is the genre in question – display the vulnerability of human *goodness* by revealing how dependent on external happenings our capability to maintain moral worth is. Consequently, they picture the fragility of human *good*, showing that uncontrolled events can easily and through no fault of ours deprive us of things which we deem most valuable, including the ability to act on our moral principles (*i.e.* practical rationality).[102] In so doing, tragedies present human life in a compassion-inspiring manner. They appeal to spectators by unveiling the familiarity of the characters' lot. As works of art, they do it beautifully, in a wonder*ful* and wonder-*inspiring* way. Unlike novels, however, tragedies break with the solitariness of reading. Intended for stage renditions, they introduce an element of dialogue, an interaction between appearing actors and perceiving viewers. Notably, it is tragedies that serve

[100] *Ibidem*, p. 237.
[101] M.C. Nussbaum, *Poetic Justice*..., pp. 93–99.
[102] *Cf.* Nussbaum's analyses of ancient tragedies in *Part I. Tragedy: fragility and ambition*, [in:] M.C. Nussbaum, *Fragility*..., pp. 23–84.

as the basic material for philosophical interpretations of compassion, including Nussbaum's Aristotle-based account.[103] This allows us to treat a theatrical tragic performance as the paradigmatic situation in which acts of appearing are perceived with compassion. And since the limits of compassion are the limits of our sense of community, it seems that an inclusive concept of society should likewise be founded on an image of a theatrical hall, where appearances are staged and viewed.

This line of thinking may be developed with the help of the comments which Nussbaum makes in a different context. When discussing one of the most fundamental of *phainomena*, (i.e. appearances understood as our common sense beliefs), the principle of non-contradiction, Nussbaum recalls Aristotle's claim that the validity of this principle can only be demonstrated by reference to linguistic practices. It cannot be discursively refuted because it is assumed in every speech act. Thus, its denial would be what we today call performative contradiction.[104] This epistemological remark is relevant to our current considerations not only because it shows the inherently social character of our rationality (which we have already discussed) but also because it explores the performative potential of language. The very act of making an utterance shapes our perception of reality, in this case forcing us to acknowledge the principle of non-contradiction.

What is important, theatrical performances employ the same mechanism. The words uttered by actors contribute to the creation of stage reality, which evokes certain emotions, for example compassion, in response. This suggests that the account of society required to complete Nussbaum's concept is one of a realm, where people are granted the possibility to present themselves to each other. Only when there is enough room for such activity left, can the limitations and prejudices characteristic of compassion be transcended and its universal potential met.

This proposition ought to be understood as both the improvement on Nussbaum's assumptions and their consequence. For we have seen that Nussbaum's conception of practical rationality is one-sided in that it privileges the perceiver's capability to produce appearances and her account of society focuses chiefly on the community of interests.[105] However, her idea of compassion calls for an inclusive interpretation of such community, which can be delivered only by recognising the value of the acts of appearing. Whereas "the discernment rests with perception", perception rests with the acts of appearing. To become a method of public reasoning, ethical perception has to operate in response to this self-projecting agency of

[103] M.C. Nussbaum, *Upheavals*..., pp. 304–327.

[104] M.C. Nussbaum, *Fragility*..., pp. 251–258.

[105] It is worth adding, however, that Nussbaum sympathises with Winnicott's interpretation of art as the adult expression of childhood play. Following the psychologist, Nussbaum observes that play is crucial for the development of the capability of imagination. When playing, children learn to take roles and react to the behaviour of other participants of the game, all of which requires imaginative reconstruction of different points of view (M.C. Nussbaum, *Not for Profit. Why Democracy Needs Humanities*, Princeton 2010, pp. 97–101). But play is essentially interactive, which suggests that imagination, as a form of perception, involves the ability to respond to the acts of appearing of another person. Unfortunately, this "playful", "interactive" aspect of imagination is not given sufficient attention in Nussbaum's writings on ethical perception, whence the above objection. At the same time, this aspect of Nussbaum's conception shows that its extension offered in this paper is suitable.

co-members of society. A politico-aesthetic concept, then, should envisage society as the space where individuals can present themselves to each other.

Urban spaces of appearing

Naturally, this approach to society has a very rich tradition. The theatrical dimension of Nussbaum's project could be developed in terms of Goffmanian dramaturgical sociology, whereas the notion of space points to Hannah Arendt's idea of the space of appearance. It is with this latter idea that I would like to conclude the paper. Rather than draw a comparison between the two philosophical projects – Nussbaum's and Arendt's – I propose to focus on the very notion of space, interpreted quite literally. That is to say, I suggest that we examine how concrete spaces can function as the exemplifications of Nussbaum's politico-aesthetic ideas. If society is conceived as space, particular spaces can be interpreted as models of society. Such reversal of the society-space comparison will hopefully shed some more light on the proposed account of society, at the same time presenting its practical relevance.

The material for such analysis can be found in the concept of urban environment conveyed in Wrocław's successful application for the title of European Capital of Culture. Significantly, the city functions here as a social microcosm, a small model of society conceived as a sphere for the acts of appearing. This is clear from the outset as the document presents the city as an intersection of various spaces – natural, social, public, private, intimate and cyber. These are interpreted dynamically, as the realms of human interaction. On this account then, a city is essentially a set of spheres, where people appear to each other on different levels.[106] At the same time, it is also a place of encounters, which facilitates the conditions for mutual understanding,[107] that is – the conception of what might count as Nussbaum's "common form of life". It brings people together so that they appear to each other as members of one community.

Significantly, the application is worded in explicitly politico-aesthetic terms. The program undertakes to regard "human moral faculty" as "the ability arising from the capabilities to perceive things and evaluate them as ones to be desired, or to be rejected."[108] Furthermore, it proposes to rediscover the ancient idea of the affinity between good and beauty, expressed in the notion of *kalokagatia*.[109] Using Nussbaum's categories, we could say that this account assumes the inherent element of wonder in our moral capabilities. An ethical judgment is an act of perceiving which involves the delight in its object, the recognition of its intrinsic value. Thus the appearing of the object stimulates its appreciative perception.

The program's appeal is for the creation of spaces where people could appear to each other beautifully, as the title slogan *Spaces for Beauty* suggests. This involves

[106] *Spaces for Beauty. Revisited. Wrocław's Application for the title of European Capital of Culture 2016*, Wrocław 2011, p. 14.

[107] This has been deftly captured in Wrocław's promotional slogan: "a city of encounters, a city that unites" (*ibidem*, p. 9).

[108] *Ibidem*, p. 15

[109] *Ibidem*, p. 14.

the fight against "public agoraphobia" – the reluctance to enter the public space. It is assumed that people should be given the chance to shape the urban space so as to reclaim it as the real place of encounters.[110] They should also gain broader access to the experience of beauty through art and culture, whose *spaces* need to be *opened up*.[111] Furthermore, presenting *Intimate Beauty* as one of its themes, the program encourages the discovery of the beauty of human body, thereby embracing our vulnerability.[112]

I believe that what all these propositions have in common is the belief in the empowering character of the experience of human creativity. It entails the sense of agency and dignity, whence its socio-political significance. To appear beautifully to oneself and to each other, then, means to appear as a creative being. Creativity is the expression of our inherent value, of subjectivity irreducible to the status of an object in somebody else's eudaimonistic project. Hence the call for the provision of space for creativity in urban environment. People should be given the chance to explore their own and each other's agency in the process of joint determination of public space, while exercising interpretative capabilities in the experience of art and also by discovering the beauty in bodily human vulnerability. Thus, a city, and therefore society in general, is envisaged as a meeting place of beauty bearers, beauty makers and beauty perceivers. It is "a common form of life", where human creativity is the uniting element.

Conclusion

The 2016 application draws on the account of society which I suggest is compatible with Nussbaum's politico-aesthetic project. As such, it both presents the possible practical implications of Nussbaum's concept (in this case – for urban politics) and provides a model of society as the space of appearing. Based on the themes of creativity and beauty, it projects a vision of society as the realm where people can discover their own and each other's worth, drawing the sense of mutual respect from the feeling of delight.

Such an account of society allows for the values which Nussbaum's concept of ethical perception is intended to protect. We have seen that the critical point of the philosopher's project is the transition from the circle of intimate bonds into the broader socio-political sphere. Here the idea of perception, based on the concept of human beings as rational and inherently social animals, does not suffice to justify the extension of our concern outside our most immediate relationships. This is why the importance of an individual's self-projecting agency has to be recognised. I have presented the shift from perception to appearing as a line of interpretation suggested by Nussbaum's writings, which, nevertheless, require supplementation with a proper interpretation of society. Only after such completion do Nussbaum's aesthetics become political.

Finally, it has to be remembered that the scope of Nussbaum's project is global, as her universalistic approach to the notion of human good suggests. From this it follows that spaces of appearing on a larger scale should be searched. If an account

[110] *Ibidem*, p. 11.
[111] *Opening Up Spaces* is the title of one of the programmatic themes (*ibidem*, pp. 40–41)
[112] *Ibidem*, pp. 43–44.

of global justice is to be worked out, citizens of different parts of the world could gain the possibility to present to each other the similarity of their vulnerabilities and their beauty. "Places of encounters" of various cultures and nations have to be provided in order to transcend local geopolitical commitments.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

DAMIAN LESZCZYŃSKI
University of Wrocław

Positive basis of negative freedom. Remarks on Isaiah Berlin's concept*

Abstract

In this article I examine some aspects of the conception of negative liberty, mainly in view presented by Isaiah Berlin in his famous *Four Essays on Liberty*. I try to prove that his dualistic conception of liberty is unsatisfactory and incomplete. It results from his one-sided and in fact incorrect analysis of positive liberty. Berlin wrongly identified positive liberty only with the source of oppression neglecting the fact that this kind of liberty is a necessary condition for liberal negative liberty he wanted to defend.

The distinction between negative and positive freedom proposed by Isaiah Berlin has received a number of comments and criticisms. Some of these comments (formulated by, among others, Gerald MacCalluna, John Gray and Charles Taylor) are very insightful and accurate. They mainly point out the lack of precision in this distinction (as was acknowledged by Berlin himself), the lack of understanding of the nature of freedom as a tripartite relationship (rather than between two parties) or too radical a juxtaposition of the two types of freedom and excessive depreciation of one of them. Nevertheless, this distinction settled into political philosophy and is still in use (functioning) but sometimes distant from the original sense (such as in Erich Fromm), usually fulfilling the role of a general or even a common sense introduction to the discussion of freedom. This is especially true in the texts dedicated to the emergence of specificity of the liberal approach, a typical negative freedom as the opposite of positive freedom, which is the possible source of unjustified tyranny, is often mentioned. On this general and common level the concept of positive freedom has gained (nomen omen) some negative characteristics, which is probably the intention of the author himself, becoming an exemplification of the essence of enslavement from which we can only be released by – positively evaluated – negative freedom.

Coming back to Berlin's source text and considering once again the distinction that he proposes, it turns out that this issue is neither as simple nor clear as it is

* Polish version of the paper was published in *Studia Philosophica Wratislaviensia* 7 [1] (2012), pp. 121–139.

often presented today. It results not only from Berlin's imprecise and somewhat chaotic (though charming) style, but primarily from the fact that focusing mainly on the criticism of positive freedom, he did not carefully consider the complex relationships between the two concepts of freedom, and above all he did not take into account the issue of the basis, which would require the realization – both in theory and in practice – of the concept, (which he accepted) of negative liberty as a freedom typical for liberalism. In this article I would like to address this last issue, trying at the same time once again to look at the structure and consequences of the idea of positive freedom put forward by Berlin.

Positive liberty as autonomy

The general understanding of positive freedom is more or less known but let us recall it here, based on the Berlin's paper. Positive freedom is connected with the question "What, or who, is the source of control or interference that can determine someone to do, or be, this rather than that?"[1] While the measure of negative freedom would be the level of interference in the activities of the subject (the lower the level, the greater freedom), the range of positive freedom can be measured by the degree to which an entity has an impact on his decision, and so is their conscious author. Positive freedom is associated with the "desire of an individual to become the master of his own fate" and the desire to "rule himself." Berlin puts it in this way: "I wish my life and decisions to depend on myself, not on external forces of whatever kind. I wish to be the instrument of my own, not of other men's acts of will. I want to be a subject and not an object, to be moved by reasons, by conscious purposes, which are my own, not by causes which affect me, as it were, from outside. I wish to be somebody, not nobody; a doer-deciding, not being decided for, self-directed and not acted upon by external nature or by other people."[2] This characteristic is quite chaotic and allows the construction of different and sometimes conflicting definitions of what freedom would be (the freedom from other people for example, is different to the freedom from the "laws of nature"), but so far it can be summed up as follows: an individual is free in a positive way when he is conscious and a sovereign author of his actions. "Awareness" can be understood here both epistemically, as knowledge about what you are doing, as well as teleologically, as intentionality. Thus some deliberate targeting of specific actions heading towards some specific results: "sovereignty" would be understood as autonomy, that is *independence from the factors which are transcendent to the subject in decision making process* (as opposed to heteronomy). The term "author" would suggest self-rule, which to some extent would lead to autonomy, whereas "the actions" would, generally speaking, define any possible

[1] I. Berlin, *Two Concepts of Liberty*, [in:], I. Berlin, *Liberty*, H. Hardy (ed.), Oxford–New York 2002, p. 169. On the subject of positive freedom as autonomy *cf.* i.a: J. Gray, *Liberalism*, Minneapolis 1986, pp. 56–60; *idem*, *On Negative and Positive Liberty*, [in:] *Conceptions of Liberty in Political Philosophy*, J. Gray, Z. Pelczynski (eds.), London 1984; L.J. McFarlane, 'On Two Concepts of Liberty', *Political Studies* 14 (1966); G. Dworkin, *Theory and Practice of Autonomy*, Cambridge 1988; G. MacCallum, *Negative and Positive Freedom*, [in:] D. Miller (ed.), *Liberty*, Oxford 1991.

[2] I. Berlin, *Two Concepts of Liberty*, p. 178.

effects of the subject's activities, both in material, and mental, terms (for example, loyalty to the nation as a conscious and autonomous choice). Simply speaking, if I am free (in a positive sense), then I do what I want and I am enslaved when I have to do what someone (or something) tells me to do. To describe such a situation Berlin uses the psychological term "self-direction". He also talks about the power over oneself, controlling yourself. Following Kant, Berlin refers to the concept of an autonomous self as a creator of values and goals. Such a situation can be schematically recognized as follows:

$$S \rightarrow A$$

This indicates the existence of a relation of determination between a subject (S) and her actions (A), which can be described as asymmetric and irreflexive (which is not of special importance here). However, the case of slavery (lack of positive freedom) would look something like this:

$$E \rightarrow S \rightarrow A$$

In this case we would talk about the relation of determining the subject by an external factor (E), which would shape the relationship between S and A. We would deal here with a transitive relationship, in the sense that S would be an intermediate element used by E to obtain A. Therefore we cannot talk about S as autonomous, sovereign, and self-directed. What is more, in principle, S cannot be called a subject in the strict sense, because from this perspective it would be treated as an object (tool or means) to achieve a goal, not as a goal in itself (in the language of Kant). Let us add one more important thing: in the first case the subject as a conscious and sovereign author of his actions would be responsible for them in the sense that he would be their intentional culprit. However, in the second situation, the issue of responsibility would become more problematic: the subject (in a limited sense, given its predetermination) would be, from the material point of view, a doer but we could think of him as an unintentional doer, and therefore an active element in terms of performance, but not their author. What is more, in a very general sense of an external determination (which will be expanded further on in the text) we could demonstrate that even the authorship and intentionality of an action does not determine the responsibility of the subject for them if you can show them as a result of transcendent causes (in extreme cases it can lead to a situation in which the subject is never responsible for his actions; it is here to explain to what extent he remains "the subject", and also how we should understand this relationship of indirect "agency").

Freedom, split, tyranny

Writing about positive liberty Berlin attempts to demonstrate how it becomes the basis for tyranny and a threat to liberal negative freedom. His reasoning is as follows: the desire to be a conscious and sovereign author of his actions causes a subject to become divided into the "true self" (rational, higher) and the "false self" (empirical, lower). In the case of freedom there is a compatibility between

these levels of consciousness, that is, the empirical self follows the guidelines of the rational self. However, what may happen is the case when the first escapes the former and, under the influence of external factors, acts against it. What is worse, it may lead to a situation when the rational self disappears completely or is dominated by the empirical self, giving some false justifications imposed by transcendent causes. We are dealing here with a classic enslavement combined with the lack of personal awareness of this enslavement. According to Berlin, this may result in a situation when someone claims the right to knowledge about the true goals and intentions of the rational self, determined by the false goals and intentions of the empirical self. Thus, referring to this knowledge about the real needs and interests of individuals, we can, at the political level, coerce them to perform certain actions, referring not only to their well-being but most of all to the compatibility of these measures with the content of the rational self and therefore including real goals, desires and interests. Therefore we can reach a paradoxical situation where individuals enslaved by determining transcendent factors inconsistent with their rational self will be freed by obedience to other external determinants allegedly consistent with their rational self. In short, a man is freed from coercion by the use of coercion of another kind.

Before analyzing these ideas, I would like to focus our attention on several things. Firstly, the reasoning of Berlin, based on the transition from positive freedom by splitting the subject to transform his freedom into a kind of slavery, does not describe the only possible and necessary consequence of positive freedom. On the contrary, it seems that the ontological problem of the relationship between the subject, his actions and external factors does not necessarily entail the adoption of a psycho-epistemological concept of the "self" split. Rather his concept, whose origins will be mentioned later on, seems to be something rather added to the concept of positive freedom, and at the same time, because of its consequences, difficult to integrate without falling into some serious problems (for example, is the concept of the split of the subject into the real and false the result of the real or false self ?). In short, it seems possible to accept the concept of positive freedom without accepting the concept of the split subject.

Secondly, a similar problem relates to the transition from the concept of the split "self" to the political issue of justifying coercion by knowledge inaccessible to the subject of his own rational self. Berlin himself demonstrates that the problem of the split can be solved within the subject, for example, based on a scheme of asceticism and therefore does not necessarily entail launching political, or any other external coercion mechanisms. Additionally, there is a serious epistemological problem here, and therefore the question of the origin and validation of the knowledge possessed by someone about the subject which is inaccessible to the subject himself. Due to the difficulties associated with this idea, it appears that the transition from positive freedom and the subject split to the described herein form of a "releasing constraint" is not inevitable. And similarly due to a combination of two prime elements, requires some theoretical effort, much bigger here because it is easy to encounter difficulties (ex. Is my knowledge of the content of the rational self unattainable to a given subject the result of my rational self or

the empirical one?). Regardless of these comments, it must be stated that such a relationship between the three elements discussed, though not necessary, does exist and what is more important, we find numerous practical confirmations of it, which are very well pointed out by Berlin in his essay. Since the transition from positive freedom to tyranny is based on coercion justified by the knowledge of the real aims of individuals, let us have a close look at this very issue.

The Epistemic dimension of freedom

In this context the following remark by Berlin appears to be very significant: "I feel free to the degree that I believe this to be true, and enslaved to the degree that I am made to realise that it is not".[3] We are dealing here with two important issues. Firstly, what is being addressed here is the "notion" of being free, not the reality of being free, which will play a significant role in attempts to impose on individuals certain actions which are against their will, justifying them by saying that that their subjective "notion" does not matter in comparison with the objective facts. Secondly, freedom, or being free, is described here in terms of truth, doubt and scepticism. Let us focus on the latter problem.

Admittedly, Berlin clearly writes here about a situation in which it is believed that the statement "I am free" is true. However, his reasoning leads us to a conclusion that first of all we have to consider a state where the truth of this assertion is known. Berlin, referring to Enlightenment thinkers, as well as Hegel and Marx, tries to demonstrate that the consequence of the doctrine of positive freedom is the thesis according to which "to understand the world is to be freed",[4] which can be paraphrased as "knowledge will set you free". In order for the same knowledge to become a justified basis for coercion, it must have a specific character. Let us now consider its two dimensions – subjective and objective.

The subjective dimension corresponds roughly to what is included in Berlin's statement quoted at the beginning of this section. What does it mean that subjectively I know I am free? Identifying this claim of having a comprehensive knowledge of the reasons for my actions will not be enough because I can have a thorough knowledge of the reasons for my actions, which are heteronomous. There is in fact no contradiction in the idea that I may know about my own enslavement. So such a knowledge is not enough but it must be a knowledge about me being the cause of my actions and that there are no external reasons for my actions. In short, I know that I am free in a positive way when I know that I am a conscious and sovereign author of my actions and their only cause. What does the subjective nature of this knowledge rely on? On the fact that the verification of the legitimacy of this information about being a conscious and sovereign author consists in referring to my own sense of freedom. I am free because I feel free, I do not feel anyone's interference, I feel I am in control of myself. In the context of the aforementioned split, I could say that my rational self is perfectly clear and completely controls my empirical self, namely, that I am convinced that this is so.

[3] *Ibidem.*
[4] *Ibidem*, p. 189.

There is of course a problem indicated at some point by Bachelard as a problem of a controlling meta-cogito. The point is that the verification of one's own thinking is done on a meta-level which, in order to be verified and analysed, also needs to be treated from the next meta-level up, and so on ad infinitum. Therefore, the binary distinction into the empirical and rational self would not seem sufficient: the rational self controls the empirical self, but to make sure that this control is performed properly, it would be necessary to control the rational self from the perspective of the rational meta-self, etc.

If we put the matter in such a way, it would appear that if the knowledge of our freedom is subjectively verified by referring it to the feeling of being free (we may quote Descartes and say "clear and explicit"), that mere verification must also be verified. Since this process is infinite, the final decision, whether our subjective knowledge of freedom/slavery is legally valid, would be impossible. An attempt to appeal to existing information about being the conscious and sovereign author of one's own actions will not change much here, because it would also require some verification that, on the subjective level, would probably always proceed according to the above scheme (i.e. to avoid regress, it would ultimately require some version of the evidentialist theory of truth). But this infinite regress is not the most serious problem. The fundamental difficulty lies in the fact that, by accepting the split of the subject into the real, rational self and the false, empirical self, we assume that some of our beliefs are false. This is inadequate in terms of non-compliance with our "real" nature. However, the feature that attests to their falsity is not immanent, because in this respect they do not differ from true beliefs. Besides, this is what the specific character of their falsehood that "pretends" to be real and "pretends" to be them relies on. We take false desires as true because we do not find in them a criterion which would enable us to classify them into one category or the other: this criterion would have to be external, but we do not have it on a subjective level, as the ultimate criterion here is our deep sense of obviousness. This situation slightly resembles the Cartesian problem with the demon deceiver: the excellence of deception lies in the fact that the falsehood resembles the truth in every respect (e.g. a perfectly deceptive dream is the one which is identical to reality, since it does not include any immanent characteristics which would distinguish them). In this situation our "clear and explicit" feelings are not any criterion. According to Descartes, the demon can deceive us even about obvious things such as the fact that $2 + 3 = 5$, i.e. that deep sense of obviousness linked to this equation, and the inability to imagine that it could be otherwise are also the result of deception. What is more, as shown by Father Bourdain, who pointed out some errors and inconsistencies in Descartes' thinking, even the conjectural certitude of the principle of contradiction, which is used in the formulation of considerations, may be based on falsehood. In short, assuming the split of a subject we know only that each of our beliefs may be an expression of either the real or false self. However, we do not know which is which. Moreover, we have here one more difficulty, mentioned earlier, which consists in checking whether the distinction between these two selves is not the result of the false self.

Therefore, the situation is as follows: some (or all) of the subject's beliefs re-

lated to the sovereignty of his decisions are wrong, and he, even being aware of this fact, does not have adequate criteria to identify them as such. Generally speaking, we can say that all the beliefs of the same subject concerning himself – and therefore also of his objectives, desires, interests and all of his self-knowledge – may be wrong, and from the immanent (subjective) point of view, it is impossible to state that. It is easy to notice here a classical theory of the the so-called false consciousness. Besides, Berlin examining these issues, makes an explicit reference to Hegel and Marx (it may be appropriate to recall here also Freudian psychoanalysis and Nietzsche's genealogy as related sources). However, the statement that a subject, as Freud said, is "a stranger to himself", i.e. has a false understanding of himself or lacks the criteria to determine which part of his self-knowledge is true and which is false, is one thing, and the other is the recognition that because of that somebody from the outside can direct him in such a way that he will implement his "real" goals and interests.

In this situation, from the subjective knowledge of personal freedom, which, as we have shown, based on the considered theory, can be totally misleading, we should move on to some form of objective knowledge. Thus it would not only be possible to ignore out insufficient and confusing subjective criteria, but also to find a basis on which we could present and explain their imperfection. The solution here would be a knowledge of freedom which would be of a scientific nature in a general and common sense of the word, i.e. general knowledge, universally valid, independent of individuals and concerning the true – independent of subjects – reality. This knowledge will not only allow us to recognize our own situation in the context of freedom/slavery, but also, as a result, the ability to control reality through anticipating and constructing conscious strategies for our actions. This approach is determined by Berlin as "rationalism" (Oakeshott and Hayek use this term in a similar way), associating it with, among many, the doctrines of Marx and Comte, noticing the assumptions underlying the "many modern ideologies: nationalist, communist, authoritarian and totalitarian"[5] and recognizing them as an inevitable consequence of the doctrine of positive liberty. However, as I mentioned at the beginning, from the point of view of the logic of the theory, this relationship is neither necessary nor the only one possible. In conjunction with the concept of positive liberty, and especially the concept of the duality of the subject and false consciousness, the concept of the possibility of objective knowledge as the basis for controlling people is not easy to validate, though, admittedly, it is the only reasonable way, from the standpoint of the theory, which is supposed to be the ground for political practice. Besides, from this perspective, the whole structure – and therefore the theory of false consciousness connected with the doctrine of objective knowledge – is an extremely powerful tool by means of which it is very easy to justify totalitarian action as heading towards liberation. And in this respect Berlin is right, but the fact that positive freedom has been connected with morally reprehensible doctrines, (which he considers to be the main argument for the rejection of the former) is, as I shall show later, not only unreasonable, but

[5] *Ibidem*, p. 191.

theoretically impossible if you want to defend negative freedom; and that is what Berlin wants to do. For the time being, I suggest we examine this "objective knowledge about freedom," and primarily how we can implement it in the context of the doctrine of false consciousness.

Freedom and gnosis

First of all, it should be noted that the very wording of the theory of false consciousness has some contradictions. This theory states that all human actions, whether practical or theoretical, and also their justifications, are determined by external interests which do not allow us to treat them as real. The fallacy is defined here as a result of external determination: whatever is heteronomous, is false. This theory, however, says that all human creations are somehow determined which raises a question about the status of this claim. For if it is true, it is also determined, so it is false, and if it is not determined, then the theory is false, since a large quantifier has been used in this theory. It can be formulated as follows: all human creations are false as heteronomous except this very claim, but is hard to take it seriously. Therefore this concept should be reformulated in such a way so that it could include conditions which would allow it to be treated, as opposed to all the others, as objective and autonomous. In order to do this we can, for example, make use of some version of historicism that incorporates a teleological theory of the development of human thought, whose culmination would be gaining some self-knowledge of the theory defended by us. We can also form some philosophical anthropology, within which some conditions the human "enlightenment" could be determined and thus a situation in which a man would have a chance to get to know the "truth" about reality, hidden until now behind the veil of Maya.

As I said, all these problems are related to the already formulated theory of false consciousness. It is easy to notice that they are also associated with an attempt to formulate and validate any other theory which would accept the assumption of a subject split, and especially one whose aim is to define the "appropriate" human desires and verify subjective beliefs about freedom/slavery, and thus whether one is, or not a conscious and sovereign author of one's own actions. I have mentioned that from a theoretical point of view it is a very difficult task – especially if we try to remain loyal to the principles of traditional logic – but a close look at the history of ideas shows that there have been many attempts to implement them, referring to the scheme, which, though based on religious and mystical sources, fitted quite well into the western intellectual tradition. What is in question here is a model of Gnostic enlightenment, consisting in the capture of knowledge (*gnosis*) about reality inaccessible to others, which is supposed to be one step closer to salvation. It just so happens that both the myth of the hermetic knowledge for the initiated, as well as the eschatological perspective associated with it, would perfectly fit into the doctrines designed not only to explain the reality, but also to change it. An interesting description of numerous contemporary social movements as a form of political gnosis can be found in the writings of many contemporary

political philosophers, including Eric Voegelin and Alain Besancon.[6] Here I will try to focus only on those elements that are related to the issue of freedom.

Let's start with this: what would an objective knowledge about freedom rely on? First of all, it could not be justified on the basis of our sense of being a conscious and sovereign author of our own actions. Besides the very feeling could not be any argument in favor of being free – it would be easier to present it as the effect of false consciousness, that is, in fact, the consequence of the enslavement. Thus, the subject cannot have an objective knowledge about himself. So can he have this type of knowledge about anyone else? Such knowledge would be related to, firstly, the relationship between the subject and the external determinants, and secondly, the same mental state of a subject, distorted by false consciousness. So in the latter case, it is knowledge of the third-person perspective (in the words of Searle) about the first – person perspective. In short, this knowledge would have to include information about whether a given subject (group of subjects) is or is not a conscious and sovereign author of actions, and what is his own self-image. Obviously, this would imply an opportunity to gain knowledge of the relationship existing in the sphere of being and mental states of others which would not determined by external factors (i.e. true), which, if we accept the theory of false consciousness, is indeed possible, but cannot be recognized as such (let me repeat it once again: this theory does not have to assume that all knowledge is false, only that there is no criterion to distinguish false knowledge from the true knowledge, and to be more precise, we cannot not get to know it because it does not exist. The mere knowledge of the criterion would necessarily imply that we had it before, because what I take as a criterion, may yet prove to be false). Thus, as you can see, even though we can formulate conditions which must be met by objective knowledge about freedom, and therefore relating to the relationship between the subject, his actions and external determinants. However, due to the adoption of the concept of duality (false consciousness), its correct formulation becomes impossible. We should now consider how it happened that despite these serious theoretical difficulties this type of such knowledge was formulated.

As I mentioned before, the whole structure of the concept presented here is clearly Gnostic. The drama unfolds on three levels – the ontological, epistemological and ethical (political). At the ontological level there is a close predetermination of human actions by external factors over which people have no control and do not even suspect their existence. So we have to deal with the widespread enslavement, which in typical Gnosticism takes the form of binding the soul with matter by an evil god, and in philosophical and political doctrines can be shown as the domination of an individual by an evil "system" or imperfect social relations (relations of production). At the epistemological level, we can talk about the lack of knowledge of individuals about their situation or having some kind of false knowledge, and thus showing a different picture of reality (usually false knowledge contains some information which convinces a man that he is free, and the factors "really" acting against him are shown as beneficial for him). Also here

[6] *Cf*. eg. E. Voegelin, *From Enlightenment to Revlution*, Durham 1975; A. Besancon, *Les sources intellectuelles du léninisme* Paris 1996.

there is a demand for real knowledge, tearing the veil of false pretences. At the ethical level (political), demands for the release from the current situation are being formulated, and thus salvation (classical religious gnosticism), or emancipation (political gnosticism). The last two levels are closely related: salvation can only be gained through some secret knowledge. However, as I have already indicated, a problem emerges at this point: how do we gain this "real" knowledge, since it is assumed that false consciousness is the natural feature of subjects and thus creating false knowledge inspired by external factors and serving to justify their enslavement? Despite the rationalistic form of this eschatology, there is no rational transition which would allow us to justify the validity of such knowledge (the condition of salvation) assuming the theory of false consciousness (the condition of slavery). The only transition has a mystical character and consists in referring to some kind of enlightenment, so that at a given moment in history there is an entity that is given is a privilege to know the truth, and liberate humanity. Obviously, the appearance of such an entity (or a group of entities) may be justified by the occurrence of "relevant socio-historical conditions" but it does not alter the fact that after the stripping the usually pseudo-scientific phraseology we are faced here with a scheme reminiscent of the plot of the popular movie The Matrix .

Paternalism and coercion in the name of freedom

It is not my intention to provide a detailed analysis of this phenomenon but because it is a possible (though not necessary) consequence of the positive concept of freedom which is being examined here, let us take a look at it on the basis of the examples from the Marxist philosophy. It is known that both in classical Marxism and its subsequent mutations, the concept of false consciousness played an important role because it explained the reasons for the acceptance of the status quo by people. This, according to marxian theory, indicates universal slavery but also allows to reject any spontaneous attempts to liberate as false, because they are based on a false vision of reality (the "spontaneous" actions were generally understood as an attempt to change capitalism in an evolutionary way and through reforms suggested by trade unions). In the early writings of Marx we can find attempts to show that negative freedom which is offered to individuals in the framework of liberal democracies, is a deliberate action aimed to arouse in them a false sense of being free, while in fact these individuals do not have this freedom and their decisions are determined by the ownership and family relations, religion, etc. In this situation, the only solution – both theoretical and political – would be emancipation or salvation through the knowledge about what the world "really" is and drawing practical conclusion in the form of a revolutionary reform. Here we are faced with a clearly explained doctrine of false consciousness and a demand for an objective study of reality as a means to liberation. Obviously the latter is to be dialectical materialism in the form presented by Marx and Engels which would in this case play the role of political gnosticism.[7]

[7] About the concept of freedom in the works of Marx *cf.* J. Gray, 'Marxian freedom, Individual Liberty, and the End of Alienation, *Social Philosophy and Policy* 4 (1986); R. Aron, *Essai sur les libertés* Paris 1965, ch. I–II.

As I showed above there are serious difficulties involved in an attempt to formulate and justify an objective theory assuming the mechanism of false consciousness, which are also related to Marx's theory discussed here. In his writings – as well as the writings of his commentators – it is hard to find a decent methodological justification of how this theory can be validated. However, if we acknowledge that we are simply dealing here with is a secular version of some religious and mystical reasoning, perhaps such a validation will be able to take a purely declarative form. And indeed it does. In the Communist Manifesto, the work of Marx and Engels, we find the following excerpt: "When the class struggle nears the decisive hour [...] a portion of the bourgeoisie goes over to the proletariat"[8] to act as its guides, on the grounds that they "raised themselves to the level of comprehending theoretically the historical movement as a whole".[9] The validation that is being referred to, may also rely on the fact that what is possible in theory (the transition from the doctrine of false consciousness to objective science), becomes possible in practice, in action (more or less as in the discussion of Zeno with Diogenes: theoretically impossible movement becomes self-evident and possible in practice). "The Communists, therefore, are on the one hand, practically, the most advanced and resolute section of the working-class parties of every country, that section which pushes forward all others; on the other hand, theoretically, they have over the great mass of the proletariat the advantage of clearly understanding the line of march, the conditions, and the ultimate general results of the proletarian movement."[10] It can be said that according to the doctrine of dialectical materialism, we all, except its founders, have false consciousness. Proclaiming this doctrine is a proof of having the true and objective knowledge (and therefore any polemics with opponents seems to be so easy: anyone who does not agree with this theory can be simply accused of having false consciousness which is evident in the fact that he does not agree with this theory). For Lenin this kind of idea took the form of "the Party" doctrine as having a real and unmediated knowledge : "The role of vanguard fighter can be fulfilled only by a party that is guided by the most advanced theory."[11] From this perspective, party members know better than the workers, what are their real goals, needs and desires are : "The workers were not, and could not be, conscious of the irreconcilable antagonism of their interests to the whole of the modern political and social system [...] It can be brought to the workers only from without, that is, only from outside [...] The working class, exclusively by its own efforts, is able to develop only trade-union consciousness [...] The theory of socialism, however, grew out of the philosophic, historical, and economic theories elaborated by educated representatives of the propertied classes, by intellectuals."[12]

[8] K. Marx, F. Engels, *The Communist Manifesto*: http://www.marxists.org/archive/marx/works/1848/communist-manifesto/ch01.htm (1.12.2013).

[9] *Ibidem*.

[10] *Ibidem*: http://www.marxists.org/archive/marx/works/1848/communist-manifesto/ch02.htm (1.12.2013).

[11] W.I. Lenin, *What Is To Be Done?* https://www.marxists.org/archive/lenin/works/1901/witbd/i.htm (1.12.2013).

[12] *Ibidem*: http://www.marxists.org/archive/lenin/works/1901/witbd/iii.htm (1.12.2013).

Such examples are numerous and we do not have to reach out to Marxist philosophy, but also to psychoanalysis, Nietzscheanism or certain trends of postmodernism. As Berlin rightly points out, we are dealing here with a kind of paternalism, which consists of the fact that the use of coercion against people is justified as being good for them and their freedom so it helps them to become conscious and sovereign authors of their own actions.[13] This paternalism has two basic forms: epistemic and political. In the first case it looks like this: since it is acknowledged that all cognitive activities of a subject and their results in the form of statements are considered to be worthless because of their external determination and deformation, the subject must base all his knowledge on the knowledge of another subject who escaped universal determination. In other words, he has to acknowledge external criteria of truth and falsehood, however not in the form of facts as the basis for verification, but the authority of the "enlightened" or "initiated" individual. He must therefore not only recognise the other person to be always right, but also the fact of his "enlightenment", that is someone's ability to gain knowledge that he does not have because of some reasons. Political paternalism is based on this epistemic paternalism and thus a position according to which individuals are not able to make right decisions (that are beneficial for them), because they do not know their real objectives, needs and interests, and must yield to the leadership of some enlightened entities that would take appropriate decisions for them and for their benefit. Historically, the doctrine of paternalism was usually associated with monarchy and some conservative trends, where the opportunity to act as a "teacher" was justified by referring to class background or some traditional hierarchy. Nevertheless, as Berlin implies, since the Enlightenment it has become one of the most common characteristics of trends that can be described as progressive, anti – traditional and collective. Interestingly, the classic paternalism rarely referred to human freedom: controlling people and using coercion against their will was rather justified with the concern for their salvation, morals, or welfare. These movements, on the other hand, began to justify the concern of progressive paternalism for the "real" freedom of a man who could not gain it by himself, but only under coercion. Here the concept of coercion changed its meaning as it was claimed that it was only to make a man, as the result of external pressure, act as if he was acting of his own accord if he had a true knowledge of his situation.

Liberalism and negative freedom

Let us summarize our considerations so far. According to Berlin's statement, there is a string of consequences connecting the demand for positive freedom, understood as being a sovereign and conscious author of one's own actions with the claim of the subject split and the theory of false consciousness which is based on it, as well as the epistemic and political paternalism, manifesting itself in the use of coercion against people justified by the pursuit of their liberation. As I have tried to show, this type of reasoning can actually be carried out but it is not necessary and even if there is a fairly close relation between the theory of false

[13] Berlin also writes about paternalism, *cf*. I. Berlin, *Two concepts of Liberty*, p. 203.

consciousness and paternalism (especially in politics), it is difficult to demonstrate that the concept of positive freedom as autonomy must inevitably lead to the idea of the subject split and the double "I". Let us assume, however, that the mere possibility of carrying out such reasoning can be seen as so dangerous in practice that we should as a precaution eliminate the theoretical concept of positive freedom as a potential source of tyranny. Concluding his considerations, Berlin writes: "Pluralism, with the measure of 'negative' liberty that it entails, seems to me a truer and more humane ideal than the goals of those who seek in the great, disciplined, authoritarian structures the ideal of 'positive' self-mastery by classes, or peoples, or the whole of mankind."[14] Therefore, we should abandon the desire to pursue the idea of an autonomous and completely transparent subject and focus on the minimalist conception of freedom defined as "freedom from" and proposed by liberalism. Let us take a look at the fundamental features of the concept.

"Freedom – writes Berlin when characterising its negative concept – is usually measured by the scope where nobody interferes with my activities. Political freedom in this sense is the area within which a man can act unobstructed by others. If I am prevented by others from doing what I could otherwise do, I am to that degree unfree; and if this area is contracted by other men beyond a certain minimum, I can be described as being coerced , or, it may be, enslaved. Coercion is not, however, a term that covers every form of inability. If I say I am not able to jump more than ten feet in the air [...] it would be eccentric to say that I am to that degree enslaved or coerced [...] Mere incapacity to attain a goal in not lack of political freedom."[15] According to this position "defence of liberty consists in the 'negative' goal of warding off interference."[16] Obviously, Berlin notices that in some situations interference and limiting individual liberty would be justified, which would entail the necessity to define some criterion which would allow us to state whether such interference is justifiable or not, that is to define the degree of freedom which every individual is entitled to, and more precisely, because it is about negative freedom, freedom from, define an area within which individual freedom can be limited.[17] In order to define such a criterion we could refer to e.g. the concept of freedom proposed by J.S. Mill, who writes that "the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of *action* of any of their number is *self*-protection [...] the only purpose for which power can be rightfully exercised over a member of a civilized community, against his will, is to prevent harm to others."[18] Let us ignore here the highly ambiguous notion of "harm" that allows various interpretations (including those

[14] *Ibidem*, p. 216.
[15] *Ibidem*, p. 169.
[16] *Ibidem*, p. 174.
[17] This distinction is very important because determining the possible area of freedom entails an attempt to codify the rights available to man, which contradicts one of the fundamental principles of classical liberalism, according to which "what is not prohibited is permitted". The area of negative freedom is not determined by (from the inside) rules, but by (from the outside) prohibitions.
[18] J.S. Mill, *On Liberty*, [in:] J.S. Mill, *Utilitrianism and On Liberty*, M. Warnock (ed.), Malden–Oxford 2003, p. 94.

which would be incompatible with the presented here concept of negative liberty) and let us consider the rest of the definition, in which Mill negatively refers to paternalism and attempts to limit freedom of the human justified by being good for them. "His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise or even right. These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him or visiting him with any evil in case he do otherwise [...] Individuals are accountable to society only for actions and behaviour that affect others".[19] We are dealing here with a typical liberal concept related to individualistic anthropology. Using a metaphor we can say that the freedom of individuals resembles force fields whose edges are connected. The activities undertaken within these fields are the sole issue of individuals because they concern only them, and those that go beyond the edges automatically start to apply to other individuals. If they are undesirable and harmful, they cause intervention of some institution which was set up to serve this purpose. The area of negative liberty is therefore determined by the freedom of others. From this point of view a man can hurt himself (by risking his freedom, possessions, health and life), but he cannot hurt others (and thus, by analogy, risk their liberty, possessions, health and life). Obviously, this creates a lot of problems of axiological and institutional nature: we have to, as I have mentioned before, determine the definition of harm, thanks to which we will be able to establish the scope of this "force field", we should also address the issue of power and the way we appoint it etc. In this case, however, these problems are not of interest to us. According to Berlin's suggestion, what is important is that in a situation when positive freedom can quite easily turn into tyranny, it would be safer and more reasonable to take care of negative freedom, i.e. the non-interference and independence.

Freedom and responsibility

Let us focus here on the following issue. What exactly do we mean when we say that a person can harm himself and as long as he does it, it should not, from the point of view of liberal, negatively understood sense of freedom, involve any external intervention in the form of limiting his independence? Well, it means, firstly, that if he operates within the mentioned above "force field", and thus his actions, or the consequences of his decisions concern only him (despite various criticisms let us assume here that such a closed system is possible), all possible choices are available for him, including those that for various reasons we do not approve of and which are generally considered harmful and dangerous. Secondly, an approval of this situation and our restraint related to non-interference is based on the recognition that an individual can and must bear the consequences of his actions, both good and bad, beneficial and harmful, concerning him and others. By not establishing a set of rules, but merely defining which acts are punishable,

[19] *Ibidem*, p. 95.

we allow the latter as possible, but undesirable, taking into account the fact that because of free choice an individual can make a wrong choice. "Liberty is an opportunity for doing good, but this is only so when it is also an opportunity for doing wrong".[20] So why can we agree with Mill that a man while exercising his freedom could hurt himself? Because we respect his choice. The possibility of bad conduct is as important as the possibility of good conduct and bearing the consequences: when the choice of evil is not available, opting for good is no longer a choice but it becomes a duty. If I act right and become rich, the consequence will be my wealth, and, in accordance with the principle of negative freedom, no one should interfere. If I lose money – as a result of bad investments or gambling in a casino – its consequence will be my poverty, which, according to negative freedom, is only my business, and nobody should interfere.

The existence of such consequences of our own decisions and an approval of a man being fully responsible for them, regardless of their nature, means that the concept of negative liberty must be complemented with the notion of responsibility. A number of philosophers, including Locke, drew their attention to this fact, stressing out that without responsibility freedom becomes lawlessness. Thus, negative freedom is not only the question of the limits of legitimate intervention but also the problem of the scope of legitimate responsibility. Let us recall one passage from Hayek's *The Constitution of Liberty*: "Liberty not only means that the individual has both the opportunity and the burden of choice; it also means that he must bear the consequences of his actions and will receive praise or blame for them."[21] So if "If we allow men freedom because we presume them to be reasonable beings, we also must make it worth their while to act as reasonable beings by letting them bear the consequences of their decisions."[22]

In short, in order to implement the liberal negative freedom, one must at the same time recognize that people are responsible for their actions. In practice it means recognizing that they should bear their consequences. However, we can ask about the justification of this position. Why do we think that people are accountable for their decisions and should bear their consequences? Because we consider them to be conscious and sovereign authors of their decisions.

Here we have reached a fundamental problem which, it seems, escaped the attention of Berlin and his commentators. The easiest way to formulate it reads as follows: negative freedom is possible only when we assume that subjects are free in a positive way. Therefore, if we think that "a man can act without hindrance from other people" and that "he is responsible to the public only for the part of his conduct, which applies to others," we have to at the same time assume that he is an independent entity, and thus conscious and sole author of his actions. For if we acknowledge that the decisions of individuals are heteronomous and that their actions are determined by external factors, then the concept of responsibility becomes blurred and disappears or moves towards these factors: if I am not the author of my success or failure, I am not responsible for them but also I am not

[20] F.A. von Hayek, *The Constitution of Liberty*, Chicago 2011, p. 142.
[21] *Ibidem*, p. 133
[22] *Ibidem*, p. 139.

free. Responsibility and consequences can only be attributed to someone who makes his decisions independently. Therefore, if we recognize that liberal negative freedom can only function when it is linked with responsibility, we have to recognize that the individual seeking non-interference in his actions is, in the words Berlin "a subject, not an object" who is moved by his own reasons and purposes, decides for himself, determines his own goals and the means to achieve them. Thus, only be an entity who is positively free can be free in a negative way.

This situation, however, requires going beyond the Gnostic consequences of positive freedom. An example of such an approach may be the classic dispute of St. Augustine with Manichaeism, defending the thesis of heteronomous character of human acts determined by physically understood and transcendent "evil" or "forces of darkness." One of the main subjects of the dispute is the problem of sin, understood in the context of guilt, redemption and salvation. Well, the possibility to consider one deed as sinful (an evil act that should be condemned but which also can be forgiven) depends on whether it was committed knowingly, intentionally and independently. "Whoever has done anything evil by means of one unconscious or unable to resist, the latter can by no means be justly condemned."[23] In other words, you cannot draw the consequences and make the man responsible for actions committed under duress, when he lacks positive freedom. "He who is forced by necessity to do something, does not sin."[24] He who does not sin, cannot bear consequences for his deeds. However, it applies not only to works qualified as evil ("Why a penance was imposed on us if we do not commit evil acts?"),[25] but also good, because in this case the question of consequences looks the same: if my actions are heteronomous in the sense discussed here, I do not bear any consequences for them and therefore I am not responsible for them. In this case, however, as Augustine noticed, just as it is pointless to talk about crime and punishment for sin, we can neither talk about their forgiveness and remission, nor a reward for good deeds. In the absence of positive freedom the whole doctrine of salvation becomes meaningless. Putting it into modern categories , we can say that in the absence of a positive freedom, the concept of liberal negative freedom does not make sense: if in fact we are not independent in our actions, of what significance, from the point of view of our freedom, is the fact whether anyone interfered or not? But even if we agree that it is relevant (because, let's say, the question of our autonomy or heteronomy cannot be decided), a problem still remains how to legitimize the principle of non-interference based on the concept of responsibility. Let us repeat once again: since I am not positively free, I am not responsible for what I do (in the sense of authorship). However, setting out the limits of negative freedom (freedom from unjustified interference) is linked to the issue of bearing the consequences of our own decisions

[23] St. Augustine, *Concerning Two Souls*, http://gnosis.org/library/dedua.htm (1.12.2013).

[24] St. Augustine, *Acts or Disputation Against Fortunatus* 16 (Polish translation: Św. Augustyn, *Sprawozdanie z dyskusji z Fortunatem*, [in:] *Pisma przeciw manichejczykom*, transl. J. Sulowski, Warszawa 1990 p. 97.

[25] *Ibidem.*

Positive freedom and free will

Agreeing that the negative freedom of an individual may be based only on positive freedom being its predecessor, we are faced with the fundamental question of free will. There is no room here to deal with this problem; I will only try to indicate what possibilities should be, in my opinion, taken into account here. First of all, we should consider what status a statement about positive freedom of an individual would have. That is, if we take the first earlier described scheme, namely $S \rightarrow A$ (an individual being a conscious and sovereign author of his actions), we must consider whether we are dealing here with a classic analytical statement, or if is a claim based on facts or speculative ontological thesis, etc. To answer this question, it should be pointed out that there are numerous arguments – both empirical and rational – implicating the weakness of the above statement and rather reaffirming a thesis expressed by the following scheme: $E \rightarrow S \rightarrow A$, thus accepting that all human behaviour is determined by transcendent factors, independent of human will and consciousness. Let us have a look at two ways in which we could defend the thesis.

Firstly, as highlighted by, among others, Hayek, modern natural science (physics, biology) and social science (sociology, psychology) have strengthened the belief that the concept of free will is archaic and incompatible with their results, since they point to the fact of a permanent condition or mediation of human activities by biological, social or psychological factors. Hayek himself, rejecting this argument proposes to treat the concept of free will in a functionalist way, which would mean that "statement that a person is responsible for what he does aims at making his actions different from what they would be if he did not believe it to be true."[26] This approach has some advantages as it allows us to defend the claim of autonomy as a performative rather than a factual statement, but Hayek is trying to defend it in such a way that he shows a dependency of the concept of responsibility on the concept of determinism, and in an attempt to reject responsibility, he connects with the metaphysical concept of the self being beyond any cause and effect sequences. The latter attempts seem to be unnecessary or even ineffective because if we acknowledge that the idea of positive freedom is not a typical factual statement, there is no longer any need to verify it and agree (accord) with various versions of determinism or indeterminism. Anyway, we would have to show that the need to accept a statement about positive freedom would result not from the fact that it correctly describes a certain state of affairs, but from the fact that it would fulfill some positive function and in some way would organize the sphere of social activities. In that way we could avoid a dispute about "free will" at least at the level of the arguments that refer to real science.

Secondly, the problem of free will is also analyzed at the level of philosophy. Apart from reductionist arguments which refer mostly to the reasoning which is based on the achievement of one of the real sciences (which leads us to the above mentioned problems), we should first of all focus our attention to the fact that the problem of free will as a metaphysical dispute between determinism and

[26] F.A. von Hayek, *The Constitution of Liberty*, p. 138.

indeterminism was criticized by Kant as the third of the antinomy of rational cosmology. Without going into details here, let us remember that Kant points to the fact that, depending on whether one accepts the perspective of phenomena or things, each party of the dispute may be right, which leads to the conclusion that its positive and unambiguous solution exceeds the capacity of our knowledge. However, what is significant from the point of view of these considerations is that in spite of the implied skepticism, Kant finds here some positive solution. Freedom, he writes, Therefore freedom does not impede the natural law of appearances, any more than this law interferes with the freedom of the practical use of reason, a use that stands in connection with things in themselves as determining grounds. In this way practical freedom – namely, that freedom in which reason has causality in accordance with objective determining grounds – is rescued [...] so that in rational beings (or in general in any beings, provided that their causality is determined in them as things in themselves) one can conceiveof a faculty for beginning a series of states spontaneously without falling into contradiction with the laws of nature."[27] As Hayek writes, "Rather, the statement that a person is responsible for what he does aims at making his actions different from what they would be if he did not believe it to be true."[28]

Let us notice that the type of freedom that Kant writes about is understood in the same way that Berlin defines positive freedom (besides in this context, there is a reference to Kant in the works of Berlin), except that if the latter subjects it to criticism as a virtually unattainable goal of human endeavour (utopian and therefore entailing coercion which is a sign of helplessness in the face of a conflict between reality and ideas), the first treats it only as a specific demand which is not an end to political actions, but their beginning and condition. The postulate of positive freedom therefore enables morality in the sense that it allows one to define the limits of possible interference in one's actions, and thus determine the extent of negative freedom. So we can say that Berlin rightly struggles with freedom as a positive ideal which realization would determine the objectives of politics, but at the same time he does not notice that this type of positive freedom as a presumption is a condition of politics as such.

What I have described above are merely outlines of some ideas of how to justify and interpret a statement (assumption) of positive freedom, which, as I showed earlier, is a necessary condition for the liberal concept of negative freedom. To see the extent to which the paradigm of Kant's practical philosophy can provide a satisfactory solution here, we should closely examine not only its relationship with the political philosophy of classical liberalism (which has been done many times), but also their relations at the epistemological level.[29] A further analysis of

[27] I. Kant, *Prolegomena to Any Future Metaphysics*, 346 (transl. G. Hatfield, Cambridge 2004, pp. 97–98)

[28] F.A. von Hayek, *The Constitution of Liberty*, p. 138.

[29] In this respect it would be interesting to juxtapose Kant's transcendental philosophy with Hayek's anti – rational and somewhat skeptical approach and also the epistemology of the Austrian School of Economics, which he (Hayek) was associated with (e.g, the question of the limits of knowledge and the statement about the impossibility of efficient central planning in the economy).

the concept of negative freedom provided by Berlin would be of no less importance, which, regardless of some shortcomings and deficiencies, some of which I have tried to indicate, still constitute inspiring material for philosophical considerations.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

JON MILLS*
Adler Graduate Professional School
Toronto, Canada

Civilization and its Fate

Abstract

We may be generally suspicious of global speculations regarding the future of humanity; but in our contemporary socio-political climate of aggression, violence, and hate, *pathos* and its derivatives continue to grip world attention, thus subjugating any hope of their abolition to the bleak forecast of pessimism. This essay addresses the role of pathology and destruction in the process of civilization and explores the degree to which the positive significance of the negative may inform new valuation practices that in turn improve human relations and world accord. Juxtaposed to psychoanalytic anthropology, Hegel's dialectic becomes a logical model for examining the possibility of global amelioration of the pernicious forces that beset the fate of humankind. We must seriously question whether mankind's aggressive essentialism will eventually lead to the end of the human race.

When Einstein approached Freud on behalf of the League of Nations and asked the question: "Is there any way of delivering mankind from the curse of war?,"[1] Freud responded with reservation suggesting that perhaps it may only be mitigated. This is the general tenor of his anthropological treatment of humanity: until base instinct (*Trieb*) is sufficiently harnessed and transformed in the service of reason, our world communities will continue to be plagued by the dark marauders of our own insidious nature. Why war? – because hate and violence are "a piece of unconquerable nature ... a piece of our own psychical constitution."[2] With this dismal portrait of human relations, we may never come to throw our hatred down.

People are slaughtering one another all over the world in the name of religion, ethnic purity, and nationalism under the guise of freedom, justice, ethical duty,

[1] A. Einstein, *Letter to Freud, July 30, 1932, "Why War?"*, [in:] *Standard Edition of the Complete Psychological Works of Sigmund Freud*, transl. J. Strachey, Vol. 22, London 2001, p. 199. Hereafter, all references to the *Standard Edition* will refer to *SE* followed by the volume and page number.

[2] S. Freud, *Civilization of its Discontents*, 1930, *SE*, 21, p. 86.

and social reform. Within the past few decades alone, contemporary ethnopolitical warfare has raged throughout the strife-torn territories of the Middle East, Central Africa, Eastern Europe, South and East Asia, and Central America where civilian populations are the primary targets of terror – marked by sadism and butchery, while women and children comprise a large percentage of the incurred human rights atrocities. Those close to the front lines of ethnic and religious conflicts are oppressed by political violence, whether they are refugees who have lost their families in ethnic cleansing campaigns, to civilians who must dodge sniper fire everyday to run to the market to fetch a loaf of bread. When the constancy of violence, terror, and war continue to saturate our daily consciousness, we can only anticipate where it will emerge next.

To what degree will our disparate cultures be able to rise above this mode of existence, where violence becomes the right of a community, either chosen or impugned? This is further compounded by the historical fact that brutality was the driving force behind the emergence of law, which still requires the use of violence to be enforced. Can actual force be replaced by the force of ideas, or are we condemned to the perversions of *pathos*? Given Freud's ontological treatise on the structure of the psyche, "there is no use in trying to get rid of men's aggressive inclinations,"[3] they are as natural as breathing; for we can never escape from the fact that our minds are primitive. *Homo homini lupus est* – "Man is a wolf to man."[4]

The history of the human race is forged on traumatization, resentment, and the need for revenge, which preoccupies collective human consciousness and fuels pathological enactments. Aggression and violence directed toward others is part of human nature, an insidious derivative of our *pathos*. For the Greeks, to be human is to suffer, to be susceptible to pain (*pathētos*), to endure illness, in short, our accruing pathology. Our *pathos* may even become fused with desire, as in 'antipathy,' a passion (*patheia*) against (*anti*) another. Mental illness stems from this basic constituency of mind. This is why Freud observed that we are all neurotic,[5] that is, ill, whereby the human aspect is saturated with anxiety, suffering, and despair – it's just a matter degree. We are all deeply affected by our *pathos* to the point that what truly differentiates individuals and societies from one another is our level of functionality and adaptation to psychic pain. In other words, human pathology is normative throughout all cultures and all times. Being "normal" is merely another word for *pathos*.[6]

[3] S. Freud, *Freud's reply to Einstein*, 1932, *SE*, 22, p. 211.

[4] Derived from Plautus, *Asinaria* II, iv, 88; *SE*, 21, p. 111.

[5] S. Freud, *Introductory Lectures on Psycho-Analysis*, 1916–1917, *SE*, 16, p. 358.

[6] I realize that the antipode between normativity and pathology poses numerous definitional and epistemological problems, as does my attempt to obviate the issue by collapsing their differences into a universal category that would apply to all human beings. This is equally difficult when offering speculations on a collective psychopathology because we lack a clear referent or criteria on what exactly constitutes health and illness to begin with, which is likely to shift based on different cultural norms and social practices. I do not wish to engage this complex issue here, which deserves serious attention in another forum. I only ask the reader to entertain the notion that, although we may not agree about the hermeneutics, scope, and breadth of the phenomenology of pathology inherent in individual and collective social life, let alone the issue of the type or

Psychopathology (πάθος) is the essence of man,[7] and it is from this standpoint that all else shall be measured. Desire precedes and supersedes reason, for primitive forces govern the psyche, which are arguably responsible as well for the exalted achievements of reason itself. Irrationality – *pathos* – is our primordial being, and it is from this ontological ground that all else materializes and makes itself known through various forms of human enactments. Reason always remains a tool, if not a slave, of desire.

Throughout this essay, I will endeavor to provide a speculative account of the future of humanity based on a discernable pattern of violence and exploitation of the Other that characterizes human motivation and deed. I must confess that I can hardly do justice to this topic in the limited scope of this project, which would take volumes to address. At best I hope to frame the issue and the inherent problematics it poses, and certainly not pretend to offer any viable solutions, for I am unable to resolve the dilemma. Instead, I shall be concerned with a narrow scope of questions that investigate whether our pathological propensities as a human race will likely bring about our extinction, or whether we can transmogrify our destructive impulses through the relational negotiation of collective valuation practices that transcend our more primal constitutions. I hope the reader will forgive me for raising more conundrums rather than furnishing practical answers. Will the fate of civilization succumb to sordid desire inspiring our demise, or will

form and the degree of their instantiation, I wish to stay focused on the psychoanalytic premise that internal conflict is intrinsic to the human psyche and social relations, which manifests itself in both individuals and groups. A return to the ancient notion of *pathos* helps us locate a common shared, lived experience where anxiety, despair, and emotional anguish are acknowledged as a universal ontological dimension to the development of civilization and humanity. It is within the confines of this context that I wish to situate my arguments. Here the question does not become whether there is a core of health in our subjective or collective strivings that stand in relation to our pathologies, only that any discussion of such pockets of health or flourishment is to be situated within the psychological predicament of our thrownness as being in relation to *pathos*.

[7] Within this context, I specifically use the masculine gender to emphasize the notion that men are usually deemed to be the chief instruments of power, aggression, and violence, whom have historically and primarily inflicted suffering on others over that of women. Here I do not wish to assign principle responsibility for collective prejudice and aggression to only one sex, only to highlight the particularly identified phenomena of male dominance. However, an equally plausible case can be made for how women, the first and original love objects for both sexes (see S. Freud, *An Outline of Psycho-Analysis*, 1940, *SE*, 23, p. 188), and usually the crucial attachment figure that dominates the early childrearing scene, can easily foist their personalities and will on their children, sometimes quite insidiously, to the degree that either gendered child is equally and potentially exposed to such power differentials. The cycle of relatedness from parents of both gender, further fortified within various familial and cultural practices, ensures that pathological accommodations and manifestations, which arise within each individual, have overdetermined sources. Just as females who are born within an oppressive patriarchy inevitably suffer in various ways, whether directly or indirectly, so do males who receive an austere maternal factor during their upbringing due to cyclical patterns of real or perceived modes of relatedness from parental authority. These patterns may, of course, become entrenched within personality structure and form the basis for a transgenerational transmission of developmental trauma that have cultural specificity leading to further repetitions and pathological enactments. What is essential, hence necessary and non-accidental, is that the human species has an intimate relation to *pathos* that inevitably saturates our being.

human accord triumph in the end? The real issue involves: To what degree will the will toward violence be sublimated into the higher tiers of self-conscious ethical reflection that reason can afford?[8] We are a world divided by race, religion, ethnicity, economics, politics, and culture, where strong emotional bonds fuel and sustain separation and difference among our communities. I do not wish to express platitudes, illusory ideals, or provide false hope – the evidence, the brute facticity of impoverishment, suffering, cruelty, and murder – points to the most archaic configurations of psychic development that permeate our valuation practices.

Within today's multicultural world community, differences and prejudices continue to divide and polarize human relations into firm oppositions that become fortified within rigid group identifications that inform collectively shared value systems. What I mean by "prejudice" is that human beings are inclined toward the preferential self-expression of valuation based on self-interest and self-valuation. Ethnic, religious, cultural, and national identities are forged through prejudicial valuation practices that in some cases even legitimate heinous forms of injustice such as genocide, terrorism, human enslavement, and child trafficking. When collective identity is so firmly established in bipolar relation to the Other, is it possible for such valuation practices to abate under the rubric of peace? Prejudice, hate, and violence are no more likely to disappear than the reality of the external world, therefore the question becomes one of amelioration.

The Positive Significance of the Negative

As Hegel completed the final installments of the *Phenomenology of Spirit*, Napoleon was outside the city walls of Jena ushering in a new age – history was being transformed once again by the revolutionary currents of the dialectic. The battle of Jena may be said to parallel the very negative character of the dialectic itself, as conflict and violence pave the path toward progression. The self-generative process of the dialectic may provide us with a logical model for addressing the problem of *pathos*; but unlike Einstein's bane of war, the dialectic may also be the boon for its solution, one that nevertheless retains its destructive features as it wages combat against itself.

Both Hegel and Freud offer a view of the human condition that is characterized by destruction, negation, and conflict; yet it is paradoxical that such negativity also becomes an animating force behind the elevation of ethical self-consciousness. Like Spirit (*Geist*), which is the sublation (*Aufhebung*) of its previous historical moments, psychic maturation is the sublimation (*Sublimierung*) of primitive mental processes. Hegel and Freud would likely concede that through reason lies the hope that communities and cultures torn apart by discordant value practices can

[8] Here I am in agreement with Hegel's architectonic dialectical trajectory that reason is a developmental achievement borne of conflict and negation, the poignant striving for self-consciousness. Ethical reflection becomes a necessary part of the sublation (*Aufhebung*) of reason despite the fact that it resonates within the feeling soul (*Seele*) and comprises our most basal desires and strivings, which he nicely enumerates in the Anthropology section of his *Philosophie des Geistes*, which is Part 3 of the *Encyclopaedia of the Philosophical Sciences*. In other words, we do not have a clean bifurcation between desire, reason, and ethical self-consciousness, for they are unified within the synthetic strands of the dialectic as a complex holism.

be united through collective ethical commitments. If humanity is to vanquish the pathology of base desire for the optimistic voluntarism enlightened by reason, it becomes important to understand how reason itself is the knight of desire designed to transform our pathologies.

We do not have to embrace Hegel's entire philosophical system, which is neither necessary nor pragmatic, in order to appreciate how his logic of the dialectic has utility for psychoanalytic thought.[9] Through his *Logic*, Hegel may be instructive in examining the evolutionary development of history achieved through negation and conquest in which further predictive possibilities for the future of humanity may be inferred. Hegel's *Phenomenology* personifies the drama of world Spirit as the coming to presence of pure self-consciousness through the process of self-estrangement, identification, and self-recognition through the mediation of the other. World hero eventually achieves Truth, satiates the lack, and arrives at full self-actualization only after traversing the arduous and protracted terrain of alienation through the vicissitudes of desire. Spirit – civilization – is therefore a constant activity, pure unrest. "It is just this unrest that is the self."[10] Hegel refers here to the unrest of *Aufhebung*, as dialectical process continuously annulled, preserved, and transcended. Hegel's logic of the dialectic involves a threefold process by which the lower relation becomes subsumed within the higher relation, at once being canceled, surpassed, but retained.[11] This pure activity of the dialectic is constantly evolving and redefining itself through such simultaneous movements, hence becoming the architecture – the ground – of *Geist*, our shared common humanity. And the driving force behind world history, behind the very process of the dialectic, is death and destruction.

Readers unfamiliar with Hegel will likely find his degree of abstraction overly abstruse and tedious, and his grand synthesis of everything has a grandiose tenor in ambition and level of generalization.[12] This is partly based on the metaphysics of his day where the most celebrated modern philosophers and German Idealists were preoccupied with the relationship and unity of mind, nature, science, religion, ethics, and aesthetics. In the words of Derrida, "Hegelianism represents the fulfillment of metaphysics, its end and accomplishment."[13] It may be helpful to view

[9] I have attempted to show how Hegel's logic and philosophical psychology are instrumental for advancing psychoanalytic thought. See *The Unconscious Abyss: Hegel's Anticipation of Psychoanalysis*, Albany 2002, where I discuss Hegel's dialectical logic in the context of his unconscious ontology.

[10] G.W.F. Hegel, *Phenomenology of Spirit*, transl. A.V. Miller, Oxford 1977, §22, p. 12. Hereafter, all references to the *Phenomenology* will refer to *PS* followed by the section and page number.

[11] Hegel's dialectic has historically been misinterpreted and grossly misrepresented by psychoanalysts and philosophers of science, most notably Karl Popper, to be a threefold relation of thesis-antithesis-synthesis. This dialectic was advanced by Fichte in his *Wissenschaftslehre*, which referred to the process of thought and judgment; thus it is an imprecise and oversimplification of Hegel's dialectic. *Cf*. J.G. Fichte, *The Science of Knowledge*, transl. P. Health, J. Lachs, Cambridge 1982.

[12] I would argue this claim could equally apply to the celebrated theoretical physicist, Stephen Hawking, whose entire life project has been preoccupied with understanding every aspect of the universe propounded in his grand "theory of everything."

[13] J. Derrida, *Margins of Philosophy*, transl. A Bass, Chicago 1982, p. 73.

Hegel's project as an attempt to describe the fundamental processes of human activity as the progressive development of cognition and culture. As the human race evolved, so did our capacity for domestic socialization, civil obedience, ethical reflection, and rational thought. However, Hegel is primarily concerned with expatiating the universal while subordinating the particular; therefore he is first and foremost interested in offering a philosophical system that applies to all people within all historical contingencies.

Hegel's notion of spirit, and that of all of history, encompasses a process in which a subject is opposed to an object and comes to find itself in the object. This entails the mediation of its becoming other to itself, with the reflection out of otherness back to itself. The process of the development of the self and that of civilization is therefore a process of differentiation and integration. For Hegel, Being is characterized by an undifferentiated matrix which undergoes differentiation in the dialectical process of Becoming that in turn integrates into its being that which was differentiated through its projection, reclaiming it and making it part of its internal structure. The outcome of the integration is once again differentiated then reintegrated; unification is always reunification. Therefore, spirit comes to be what it already is, the process of its own becoming.

Spirit as the striving for pure self-consciousness ascends toward an absolute understanding of itself and comes to a unity constituted by the bifurcation and rigid opposition that it generates from within itself. It is precisely through such opposition that consciousness brings itself into reunification. Thus spirit, in its evolution, undergoes a violence at its own hands. By entering into opposition with itself, it raises this opposition to a higher unity and thus sublates it in a new structure. As each shape or content is confronted with radical opposition, each shape is made to collapse when its non-absolute form is exposed. Indeed, it is always driving the movement on from one shape to the next. Thus the character of the dialectic is that of negativity and conflict; it is tempestuous, feral, and dynamic. Spirit as such is the source of its own negativity as inversion and destruction pave the way for its progression forward.

There is a necessity to the dialectic that informs the internal structures of spirit, that is, there is a certain determinism to negation. The operation of such determinate negativity comes about through the collapse of each shape. As negation of a certain content takes place, it derives a certain content from the negation. Therefore, it links shapes into a necessary progression as each form turns into a new one. However, as each form is surpassed, the experience of its alteration is that of death, its end. But for Hegel, death always leads to rebirth. The dialectic is therefore the oscillation between life and death, never separate from one another. For Hegel, spirit is always "tarrying with the negative" – confronting Death, for

> to hold fast what is dead requires the greatest strength; ... the life of spirit is not the life that shrinks from death and keeps itself untouched by devastation, but rather the life that endures it and maintains itself in it. It wins its truth only when, in utter dismemberment, it finds itself.[14]

[14] G.W.F. Hegel, *PS*, §32, p. 19.

As determinate negativity, spirit vanquishes itself as it destroys itself. It kills itself as it gives itself life. This is the "tremendous power of the negative,"[15] staring death straight in the face, converting it into the positive. It is precisely through such negativity that there is progression, destroying itself in the service of raising itself – the positive significance of the negative.

If the dialectic becomes a logical model in its application toward a global amelioration of psychopathology, then we must be able to logically demonstrate whether it has the potential to bear any fruit. We may appeal to historical facticities that trace the epigenesis of humankind and perhaps even come to the conclusion that, despite all the carnage and social decay, we have evolved into more a civil and enlightened species, even though human aggressivity and immorality captures the locus of our attentions everyday as a TV screams. But a historical account alone carries less predictive value, for we have no means of being able to predict with much accuracy the future contingencies that will affect the teleological progression of human psychodynamics, hence contingencies that always inform the mediatory interventions mind assumes in each immediate shape it encounters. We may do better to stay on ontological ground, a ground that informs our collective anthropology; and we must be able to demonstrate the internal consistency and systematic coherency of the dialectic that Hegel's Logic affords.

The self, as well as world spirit, or what we may call the universal soul – psychic processes that belong to us all, is an epigenetic construct, thus a teleological movement that is a procreative self-articulated complex holism.[16] As a self-generative telic will, cognition is free in its encounters with the contingencies of its reality, taking into account the exigencies of its environment and the novelties of immediate experience. Therefore, spirit is not pre-designed or predetermined toward a presupposed end, but rather its end is a transformed achievement – "the logical and ontological Alpha of the cosmos, but only after it has emerged as its logical and ontological Omega."[17] It emerges through the process of mediation and negotiation with the existential realities it confronts.

Our faith in the transcending power of mind over the combative regimens of world discord is acceptable only to the extent at which we believe in a progressive trend toward increased solidarity through collective self-conscious rationality. The level of psychic development Hegel points toward is hardly achieved by intellectuals let alone the masses, for reason is often eclipsed by the primal lure of desire. If the facts of history and human nature do indeed lean toward a steady progressive self-conscious liberation of rational freedom, then to what degree is this the result of our aptitude to bridle and sublimate our primitive proclivities for the ideals of conscience and the rational demands of a civil society? The promise of increased unity in the face of disharmony augers well for a collectively shared and constructive value system; however, the ostensive prevalence of global division and chaos

[15] *Ibidem*.

[16] Sean Kelly provides a comprehensive account of Hegel's theory of complex holism in *Individuation and the Absolute: Hegel, Jung, and the Path Toward Wholeness*, New York 1993.

[17] J.N. Findlay, *Hegel's Use of Teleology*, [in:] W.E. Steinkraus (ed.), *New Studies in Hegel's Philosophy*, New York 1971, p. 93.

saturated by prejudicial conflicts may leave us with a less optimistic interpretation of the fate of humanity. The problem of destructiveness becomes the central task of our investigation, for if Hegel is correct, it becomes the stallion of unification as it gallops toward the horizon of reason. But if our aggressive trends continue to go unchecked, Freud's admonition of the possible extinction of the human race carries foreboding merit. Our analysis of the positive significance of the negative will lead us to conclude whether philosophical psychoanalysis may genuinely offer a contribution to peace.

Psychoanalytic Anthropology

The primary significance of destruction is never so forceful as in Freud's postulation of the death drive (*Todestrieb*), the foundation that governs psychic development to which "*the aim of all life is death*."[18] Negativity is always the base agitation of any organism – the destruction that constructs life – the purpose of which is to return to the original lost unity of its symbiotic state. The notion of original unity is instructive for our understanding of a principle of world harmony devoid of the more pathological instantiations of human aggression because, for both Freud and Hegel, consciousness emerges from an unconscious undifferentiated unity with its primordial nature. Just as Freud speculates on how the organic arises from the inorganic,[19] as the general object of anthropology, Hegel traces the dialectical emergence of the feeling soul from the abyss of its indeterminations; at first unseparated from its immediate universal simplicity, it then divides and rouses itself from its mere inward implicitness to explicit determinate being-for-self.[20]

For Hegel, spirit begins, like ego development for Freud,[21] as an original undif-

[18] S. Freud, *Beyond the Pleasure Principle*, 1920, *SE*, 18, p. 38.

[19] *Ibidem*, pp. 36–39.

[20] G.W.F. Hegel, *Des Philosophie des Geistes*, part three of the *Encyclopaedia of the Philosophical Sciences*, [in:] M.J. Petry (ed.), *Hegel's Philosophy of Subjective Spirit*, Vol. 2: *Anthropology*, 1977, §§388–403. Hereafter, all references to Hegel's *Philosophy of Spirit* from the *Encyclopedia* will refer to *EG* followed by the section number. This process is also the logical model Hegel follows from his *Logic* in his anthropological description of the soul, where a universal determines itself into particulars, showing how each mediation forms a new immediate, which is the general thrust of the dialectic.

[21] For both Hegel and Freud, the inchoate ego is originally encased in a unity and is therefore modally undifferentiated from external forces – the inner and outer are fused in a symbiotic organization. Freud informs us "originally the ego includes everything, later it separates off an external world from itself. Our present ego-feeling is, therefore, only a shrunken residue of a much more inclusive – indeed, an all embracing – feeling which corresponded to a more intimate bond between the ego and the world about it" (1930, *SE*, 21, p. 68). For Hegel, the natural soul moves from an undifferentiated unity to a differentiated determinate being; so too for Freud, ego boundaries gradually becomes more contrasted, constructed, and consolidated throughout its burgeoning activity. Freud notes that originally an infant is unable to distinguish between its own ego and the external world as the source of stimulation and sensation. But eventually the organism comes to discern its own internal sources of excitation, such as its bodily organs or somatic processes, from external sources of sensation, (e.g., mother's touch, breast, etc.), that become set apart and integrated within ego organization. It is not until this stage in ego formation that an object is set over against the ego as an existent entity that is outside of itself. Once the ego moves from primary to secondary narcissism, attachment to external cathected (love) objects form the initial dynamics of object relations and character development.

ferentiated unity that emerges from its immediate self-enclosed universality to its mediated determinate singularity. This is initiated through a dialectical process of internal division, self-externalization, and introjection as the reincorporation of its projected qualities back into its interior. Through the complexities of mediation and sublation, spirit achieves higher levels of unification until it arrives at a full integration of itself as a complex whole, uniting earlier finite shapes within its mature universality. The need for social order, unification, and harmony are motivational factors that inform the ideal of global tranquility which human violence, hate, and terrorism threaten to deteriorate, an ideal imbued with the residue of early symbiotic conditions.

The ego ensnared in the stage of primary narcissism,[22] like spirit asleep in the undifferentiated abyss of its self-absorption,[23] constitutes the psychological and ontological precursors for differentiation and development. To what degree do these conditions play in our wish for higher degrees of unity, concord, and moral self-realization? Are we to understand world spirit as "the universal brotherhood of man"[24] that seeks absolute unity, or is this merely a wish to return to the "oceanic feeling"[25] of symbiosis like a fetus in the peaceful sea of its mother's womb? To what degree is this an illusion that preoccupies so many minds, like the parallel wish for union with God, the exalted father who shall make our home safe and free from our helplessness and pain?[26] But whether these are fantasies or not, they represent moral ideals: "Thou shalt not take the name of the Lord thy God in vain!"[27] – don't you dare assault my desire!

One would be hard pressed to find someone who would not value the ideal of peace, with communal harmony, accord, and cooperation marshaled in the service of social progression. But the very nature of the need for progressive unification is also dialectically opposed to destructive and regressive inclinations that derive from earlier primitive shapes of our psychic constitution we seek to act out or recover during conflict precipitated by opposition. If the desire for unification is a derivative of our original psychical ontology, then both progressive and regressive desires may be said to emanate from the same mental (symbiotic) configurations which may further possibly serve the same aim. Both seek unity or peace of a different kind and in a different form: one through the attainment of higher integrated complexities, the other a wish to return to the warm blanket of its initial undifferentiated beginning – unity is nevertheless their goal. Here 'unity' should be understood within the symbolic context of psychological integration, where strong affective bonds for concord, union, and amalgamation are achieved and internally experienced as a transcendental ideal. Following Hegel, this would require some integrative function that would attempt to bind or resolve opposition,

[22] S. Freud, *On Narcissism: An Introduction*, 1914, *SE*, 14, p. 100.
[23] G.W.F. Hegel, *EG*, §408, *Zusätze*, 2.
[24] H.S. Harris, *Hegel's Development: Night Thoughts*, Oxford 1983, p. 411.
[25] See Freud on the "oceanic feeling" in relation to religious sentiment and early ego development; *Civilization and its Discontents*, 1930, *SE*, 21, pp. 64-68.
[26] S. Freud, *Future of an Illusion*, 1927, *SE*, 21; *Civilization and its Discontents*, 1930, *SE*, 21, p. 74.
[27] See Freud's discussion, *Civilization and its Discontents*, 1930, *SE*, 21, p. 74.

or for Freud, serve the pleasure principle through sublimation. But if the drive toward destruction is responsible for both progress and regress, growth and decay, then how are we to determine which one will advance and which one will succumb to the tyranny of the other? This brings into question how the nature of negativity and destruction influence the self-preservative drives in their quest for unification and mastery.

Freud tells us of two competing forces in human nature: the will toward life and the will toward death manifested as Eros or libido, the sexual force responsible for erotic life, and its antithetical companion conceived under the drive toward destruction.[28] This dual class of innate drives comprise those which seek to preserve and unite and those which seek to kill and destroy, both giving rise to what may be characterized as our caring and aggressive propensities. "Neither of these drives are any less essential than the other; the phenomena of life arise from the concurrent or mutually opposing action of both."[29] Furthermore, they scarcely operate in isolation, both borrowing from the resources of the other as an accompanied or alloyed counterpart, drawing a certain quota from the other side, which in turn modifies its aim or is even used to achieve its aim.

This union between life and death is the ontological fabric of the human mind to which all other dialectical polarities arise including the universality of Love and Hate. Self-preservation is clearly an erotic impulse but it must have aggression at its disposal in order to accomplish its task; just as in love, the aggressive drive is utilized in order to gain mastery and possession over an object in which the attachment to it brings about. Although the self-preservative drives stand in stark opposition to destructive ones, the two are dialectical complementarities that reflect their confluence. Here we have a similar structural dynamic of the Hegelian dialectic with negativity begetting progression in the service of achieving higher aims. Just as Being is in opposition to Nothing, so is life and death, two sides of a symmetrical relation, their necessary unity.

Collective identity is based on the strength and intensity of emotional ties among its members and the mutual identification with shared valuation practices, thus giving rise to diversity, opposition, and prejudicial division between individuals, cohorts, cultures, societies, and nations.[30] The greater discrepancies such as race, religion, ethnicity, nationality, and political affiliation that bring about more pronounced forms of prejudice and contempt is not surprising. The increased enthusiasm in nationalism, religious identity, and separatism among our diverse peo-

[28] S. Freud, *The Ego and the Id*, 1923, *SE*, 19, Ch. 4.

[29] S. Freud, *Freud's letter to Einstein, "Why War?"*, 1932, *SE*, 22, p. 209.

[30] It may be argued that identity is largely the result of the identification process itself which is influenced by myriad causal and overdetermined factors that are encountered throughout our life experiences and internalized within personality formation. Along with drives and their transformations, the nature of identification accounts for much of the intrapsychic motivations, intentions, desires, and conflicts that comprise psychical and social reality. Identity, whether personal or collective, is ultimately in the service of narcissism or self-interest, thereby affecting the ideals we espouse and the valuation practices we choose to identify with over others. Despite the overdetermination of identification, the values and mores individuals and societies adopt are fundamentally the result of the complexities of narcissistic object choice, the psychosocial functions they serve, and the evolutionary demands of the self-preservative drives.

ples point toward the need to define ourselves in opposition to difference, rallying greater collective fellowship among its identified members, and thereby strengthening the cultural narcissism that hold societies together – all in the service of the self-preservative drives that align with similarity and cultural identification. In *Group Psychology and the Analysis of the Ego*, Freud underscores the universality of prejudice:

> Every time two families become connected by marriage, each of them thinks itself superior to or of better birth than the other. Of two neighbouring towns each is the other's most jealous rival; every little canton looks down upon the others with contempt. Closely related races keep one another at arm's length; the South German cannot endure the North German, the Englishman casts every kind of aspersion upon the Scot, the Spaniard despises the Portuguese. We are no longer astonished that greater differences should lead to an almost insuperable repugnance, such as the gallic people for the German, the Aryan for the Semite, and the white races for the coloured.[31]

There is such a narcissism of even minor differences between individuals and cultures that the very sound of rap music blaring from an open car window could lead one so judgmentally inclined to conclude that the true meaning of "culture" is to be found growing in the bottom of a test tube. We are further informed by scientists that we now have the empirical means by which to measure the degree and intensity of our disgust for minor differences by observing the pupil size of an individual. It is a universal biological fact that regardless of light differences, our pupils dilate when we like something and become pinpricks when we perceive something to be a repellent.[32] It is comforting to know that when uncertain about whether or not someone is friend or foe, all you have to do is look at the center of their iris to determine if it is the size of a pinhead.

The nature of identification has its origins in early ego development whereby the child takes its parents as ideal objects who, along with their value systems, become internalized within personality formation and effect the germination of moral conscience as well as the capacities for love and hate.[33] Whether personal or collective, identity is defined in opposition to difference and identification with similarity. This structural dynamic alone may be said to account for the need for division, uniqueness, and prejudicial self-preferences as opposed to others who stand in marked difference. However, the confluence of destructive and self-preservative forces compound the nature of identifications and social relations where desire justifies murder and reason is manipulated to assuage primal instinctual urges in the service of narcissistic pursuits.

There are such countless examples of the polarization of values and ideals that stand in opposition to others that you could spend the rest of your life trying to

[31] S. Freud, *Group Psychology and the Analysis of the Ego*, 1921, *SE*, 18, p. 101.
[32] D. Morris, *The Human Species*, aired July 12, 1998 on The Learning Channel (TLC).
[33] S. Freud, *Group Psychology and the Analysis of the Ego*, 1921, *SE*, 18, Ch.7.

catalogue them all. In many cases of group prejudice, valuation practices assume a form of collective identification based on a simple rigid economy. Intolerance to difference that precipitates extreme forms of violence may be said to represent a regression to our most primitive constitutions when bad (self)objects are regurgitated from the mouth as poisonous projections of evil and hatred that must be annihilated, what Klein would refer to as the 'paranoid-schizoid position.'[34] In many cases, extreme prejudice is the product of pathological narcissism. Patriarchal value placed on male children over female children has historically led to infanticide that is still practiced today in parts of China, India, and the Middle East. Since the Taliban took power of the government of Afghanistan, women have had to wear burqua; and have been beaten and stoned to death in public for not having the proper attire, even if this simply means not having their mesh cover the front of their eyes. After James Byrd Jr. was dragged from his feet down a rural road by a chain secured to the back of a pickup truck until his right arm and head were literally torn from his torso just because he was black, the devastated town of Jaspers Texas was greeted three weeks later by a Ku Klux Klan rally. Genocide continues to rip through our world claiming innocent lives, from the Hutu's massacre of the Tutsis in Rwanda, to the Serbs mass extermination of Bosnian Muslims, to the systematic slaughter in Kosovo, not to mention the recent plethora of killings initiated by the Arab Spring. While these are extreme cases, one need not look further than one's own country to confirm the ethnic and patriotic narcissism that envelopes us all.

Both Hegel and Freud stress the importance that civilization is a process. But these aforementioned events hardly resemble the mores of a civilized culture as irrational fanaticism justifies barbarity broaching the brink of insanity. The primitive economy of rigid identification that justifies these extreme forms of savagery has at its disposal all the unbridled resources of the death drive turned outwards. The drive toward death is transformed into the destructive drive when it becomes projected onto external objects. In this way, self-preservation is maintained by destroying extraneous threats as objects of hate are rendered impotent. "Hate, as a relation to objects, is older than love. It derives from the narcissistic ego's primordial repudiation of the external world with its outpouring of stimuli. As an expression of the reaction of unpleasure evoked by objects, it always remains in an intimate relation with the self-preservative drives."[35] Yet self-preservation versus pleasure induced in killing are two different inflections of narcissism. Sadism, the derivative of hate, is nowhere so evident as with the deranged techniques conceived and used to torture, maim, and murder millions of victims in the Holocaust, and in the killing fields under the Khmer Rouge government, as well as in the death camps manufactured by Bosnian Serbs in the name of ethnic cleansing.

The Bosnian concentration camps were one of the most horrific human slaughterhouses, because the means of extermination were laborious and perverted, the aim of which was to produce the most excruciating amount of pain, mental an-

[34] See M. Klein, 'Notes on some Schizoid Mechanisms', *International Journal of Psychoanalysis* 27 (1946), pp. 99–110.

[35] S. Freud, *Instincts and Their Vicissitudes*, 1915, *SE*, 14, p. 139.

guish, and suffering possible.[36] Although it is illegitimate to make comparisons, the killing at Auschwitz was largely mechanized and bureaucratic, while the genocide at Omarska was emotional and personal, mainly depending upon the simple and intimate act of beating. These techniques were inefficient, time-consuming, and physically exhausting, yet they were habitually and systematically employed to intentionally demoralize and demolish, bringing warped pleasure to the guards and paramilitary units who, through their innovative means at devising methods of torture, could greatly bolster their prestige. The use of rape warfare on women – especially adolescents and children – is another such example of the chilling psychological and sociological rationale for the deliberate and systematic means of deteriorating the opposition from within their own support systems by depleting their morale, ego defenses, and will.[37] Here we can see how reason is distorted under the dictatorship of psychopathic narcissism. It is in moments like these that one can hear the voice of Luther – *die Hure Vernunft* – "reason is the whore of humanity."[38] We can rationalize away anything, even our morality.

Is the death drive so intent on persecuting humankind that it will eventually bring us to ruin? The bleak forecast of the continual historical rein of terror by sick minds in positions of power and privilege may lead us to rightfully conclude that "men are not gentle creatures who want to be loved,"[39] rather they want to exploit, con, use, conquer, humiliate, torture, and kill. In *Civilization and its Discontents*, Freud writes:

> The fateful question for the human species seems to me to be whether and to what extent their cultural development will succeed in mastering the disturbance of their communal life by the human drive of aggression and self-destruction ... Men have gained control over the forces of nature to such an extent that with their help they would have no difficulty in exterminating one another to the last man. They know this, and hence comes a large part of their current unrest, their unhappiness and their mood of anxiety.[40]

It may be argued that religion and ethnicity – including race – are the main reasons for divided group identifications. Together, ethnicity and religion form

[36] M. Danner, 'America and the Bosnia Genocide', *The New York Review of Books* 44 [19] (1997), pp. 55–56.

[37] B. Allen, *Rape Warfare: The Hidden Genocide in Bosnia-Herzegovina and Croatia*, Minneapolis 1996.

[38] This quotation is attributed to Martin Luther by E.M. Cioran in *The Temptation to Exist*. Also see "The Last Sermon in Wittenberg, 1546," *Luther's Works*, 54 vols., (Philadelphia): "And what I say about the sin of lust, which everybody understands, applies also to reason; for the reason mocks and affronts God in spiritual things and has in it more hideous harlotry than any harlot. Here we have an idolater running after an idol, as the prophets say, under every green tree [*cf*. Jer. 2:20; I Kings 14:23], as a whorechaser runs after a harlot. That's why the Scriptures call idolatry whoredom, while reason calls it wisdom and holiness.... Such wisdom of reason the prophets call whoredom." (p. 374–375).

[39] S. Freud, *Civilization and its Discontents*, *SE*, 21, p. 111.

[40] *Ibidem*, p. 145.

the social value structures that become the macrocosm of any culture, which furthermore acquire personal and collective meaning that validates nations and keeps them together. Ethnic and religious identification is so strong that even between closely related ethnic and religious groups, rigid group identifications keep societies from embracing shared qualities simply because of minute differences that threaten cultural narcissism. When dispute over land continues to flare throughout much of Eastern Europe and the Gaza Strip and West Bank, ethnicity, religion, and nationalism become more demarcated with group identifications more virulently opposed.

The stronger the intensity of emotional bonds between people, the stronger identifications become. Group identity fosters unity and progression, but it may also lead to discord and regression – the dynamic that fuels both peace and war. In my opinion, group identifications are responsible for the process and advance of civilization as collective value systems govern the ideals of a community. As a general rule, any movement that encourages greater emotional attachment to others strongly militates against the loom of destruction, for love is the engendered ideal and the heart of conscience. When people are governed by empathy and conscience, reason is marshaled in service of justice and the pursuit of the ethical. This too requires an inversion of aggression that becomes the internal judge of conscience, where guilt and shame equally inform our moral choices, as does reason.

Take for example, two different cultural responses to involvement in World War II. Germany experiences a great deal of universal shame for their infamous role in history that fractured world order, yet they still acknowledge and remember their history, teach it in the classroom, and maintain public museums, camps, and monuments, while the Japanese still live in collective denial of their involvement in the war. The official government policy does not recognize its historical atrocities or war crimes, which are omitted from textbooks and prohibited from being taught in public schools. Here we have two responses to collective shame: acknowledgment with the educational concern that history should not repeat itself, and denial in the name of 'saving face.' We shall not deviate far in saying that one is a healthy response of remorse to guilt and shame, while the other is an infantile attempt to maintain a cultural narcissism where the superiority of the Japanese race is inculcated in school children every day and institutionally solidified by national identity.

I use these previous examples arbitrarily as merely illustrative of particular instantiations of collective identification, yet they occur everywhere. Regardless of what examples we focus on, which always runs the danger of introducing ancillary distractions, whether they be distantly historical or more contemporary in world attention, the phenomena of collective identification equally applies to any culture or society throughout time. Although particularity is culturally relative and contingent upon the interaction between group relations and the social environs that inform collective identity, it also transpires within a greater universal process governing object relations. Here it is important to stay focused on the universal rather than the particular, for the more crucial locus of our inquiry concerns human nature itself, namely, that which is common to us all.

The End of the World

In this age of terrorism post 9/11, world anxiety becomes our number one preoccupation. When deranged minds are willing to sacrifice their own lives during the suicidal terrorist acts of committing murder, no one or nation is immune from threat. In the wake of such persistent anxiety, 'unhappy and paranoid' becomes the epithet we shall apply to characterize collective humanity. And just as Freud's seminal work, *Das Unbehagen in der Kultur*, described the unease, trepidation, and unhappiness within the culture of his day during the early rise of Hitler, we may justifiably conclude that the scale and ferocity of cross-cultural/interfaith/interethnic aggression has intensified and gotten much worse since his time. The fantasy that men are inherently gentle creatures who are born good, free of dispositional sin, and untainted by primitive intent can no longer be sustained by critical reason. It is an empirical fact that by all historiographical accounts of cultural anthropology, human civilization has been forged on human conflict, attachment deficits in parent-child rearing practices, emotional, physical, and sexual abuse, traumatization, and war. Given the historical progression of civilization, what reasonable trajectory do we posit for the future of humanity?

The astrophysicist and cosmologist Brandon Carter has provided a mathematical formulation that predicts the probability of human extinction.[41] Given that there are over 7 billion people alive on this planet today, and that we are amongst the most people who have ever lived in the history of the human race, from a predictive statistical standpoint, it is speculated that there is an approximate 5% chance that we will expire within a couple hundred years and a 95% chance that complete human extinction will occur in approximately 7 thousand years, with a possible degree of freedom extending this figure to just over 9 thousand years. This is known as the "Doomsday Argument." In other words, if all the humans alive today are in a random place in the human history timeline, then we are nearer to extinction than not.

While there are different versions of this scenario that vary in scope and formulation, including critiques, refutations, and rebuttals, philosopher John Lesley has championed this argument in his chilling speculations on the end of the world.[42] One cannot entertain the actual risks of complete human annihilation based on such brute evidence without sinking into worrisome pessimism. Lesley draws alarming attention to the underestimated dangers that threaten human extinction, including the notion that we could become extinct fairly soon. Despite the recognized risks of natural disaster, including volcanic eruptions, the earth colliding with asteroids or comets, astronomical explosions like supernovae, galactic center outbursts, and solar flares, or a complex breakdown of the earth's biosphere, we are well aware that most of the immediate threats to the survival of the human race come from man. We damage our own ozone layer, dump toxins into our air, lands, and seas via mass industrial pollution, increase greenhouse effects that

[41] B. Carter, 'The anthropic principle and its implication for biological evolution', *Philosophical Transactions of the Royal Society of London* 310 (1983), pp. 346–363.

[42] J. Lesley, *The End of the World: The Science and Ethics of Human Extinction*, London 1996.

ruin our ecosystems, and introduce fatal viral diseases and new varieties of plague that infiltrate our continents. Soon the world could become uninhabitable. From nuclear war to germ warfare, radiation poisoning, biological and chemical warfare, terrorism, criminality, technological manipulations such as genetic engineering disasters, food infections (i.e., salmonella bacteria), computer-initiated network malfunctions, internet viruses, or techno-war that jeopardize human survival, and scientific hubris – like biohacking, nanotechnology, or careless physics experimentations "at immensely high energies, [that] will upset a space-filling 'scalar field' and destroy the world"[43] – these are but a few very serious reasons not to dismiss the ubiquitous threat of world annihilation.

In our contemporary era of political and religious violence that legitimizes the morality of war with the support of military science, our conflict of cultures begets and bears witness to increased human tragedy and traumatization. There are pernicious threats associated with the subversive activities of fundamentalist religious groups that embrace ethical relativism and prescriptivism based on collective ideology just as there are repercussions from foolish decisions made by narcissistically grandiose politicians who hype up a country's citizenry based upon an appeal to emotion in the pursuit of national self-interest. If we don't kill each other by destroying our environment through chemical, biological, and nuclear war, leading to loss of biodiversity, disease, disastrous climate change, greenhouse calamity, desertification, and pollution of our planet, then overpopulation will surely erode our environment and tax our natural resources to satisfy basic human needs, which will likely lead to mass panic, mayhem, and global warfare. When people have no grain to eat, the moral principle of human rights becomes a vacuous concept.

The Ontology of Prejudice

The polarity of human desire, the nature of personal and collective identifications, and the combative forces of social and cultural oppositions all operate within anthropological and ontological structures that give rise to civilization and the historical manifestations of pathology. Civilization, even more so than nature itself, is responsible for most of our malaise; but it is also responsible for our remarkable advances in technology, science, medicine, human rights, aesthetics, and moral conscientiousness that enhance the quality of human life, most of which having occurred in our lifetime. But along with these advances have also come the technology to extinguish the entire human race. This is especially disturbing when fanatical and paranoid minds have means and access to weapons of mass destruction. This places us in the precarious position of attempting to anticipate the possible fate of humanity, for the predictive validity of the progression of civilization hinges on whether aggression will be restricted, displaced, inverted, and sublimated for higher rational, ethical, and aesthetic pursuits.

With regard to the question of the possibility of global amelioration of pathological enactments, the issue becomes one of degree. Since prehistory, culture has undergone an evolutionary process of becoming, which is responsible for what

[43] *Ibidem*, p. 1.

we have come to call civilization, our evolved contemporary valuation practices. However, like Freud observed, "uncultivated races (sic) and backward strata of the population are already multiplying more rapidly than highly cultivated ones."[44] While there are many socio-economic, political, and psychological reasons for this, they nevertheless obstruct the optimal transformation of our *pathos*.

Prejudice forms a basic constituency in our psychic constitutions, for we pass judgments on others based on our preferential appraisal of what we value and are accustomed to find familiar and/or pleasing. The double edge of the dialectic (as negativity resulting in higher unity) exposes us to a dilemma, for the dialectic is the ontological dynamic underlying prejudice itself. Part of the problem facing us is that prejudice is ontologically constituted in the most rudimentary aspects of human consciousness. Like the nature of the dialectic, prejudice has both negative and positive valences. While violence and destruction are the instruments of prejudice, so too are caring and love. Prejudice is not merely a negative construct; prejudice defines our valuation practices, which are the Mecca of individual and communal life. Rather than conceive of prejudice as simply a pathological anomaly, prejudice is also responsible for our most revered ideals. As I have said elsewhere,[45] prejudice in its essence is the preferential self-expression of valuation. All prejudicial disclosures express value preferences. Preferences are prejudicial because they signify discriminatory value judgments that are self-referential. Preference presupposes prejudice for preference typifies the priority of determinate valuation. To prefer is to value and to value is to judge: judgments by nature are valuative. All judgments are imbued with value, which presuppose self-valuation and self-interest, because valuation is a particular form of subjective self-expression. Thus, valuation is prejudicial, for it involves a relation between difference and similarity that is necessarily self-referential. Every human being by nature is prejudiced; it is simply a matter of degree, and toward what particular object one's prejudice is directed.

Prejudice is a neutral psychological predisposition that informs the ontology of human subjectivity. Prejudice is an elementary aspect of conscious and unconscious life that gives rise to the self, the nature of personal identification, individual and collective identity, culture, and shared value practices. Prejudice as valuation is therefore responsible for our shared ideals as well as the deviations of abnormality and perversion. In its ideal condition, prejudicial valuation informs our social mores and ethical practices. In its larger scope, ethics is the harvest of subjective universality. As such, selfhood and culture give rise to morality that is individualistic and interpersonally bound within a psychosocial matrix of negotiation and intersubjective validation. Value determinations, I suggest, are the result of interpersonal mediations and identifications with collective ideals and are therefore intersubjectively constructed and validated through the dialectical process of our social and cultural prejudices. Applying Hegel, the succession toward greater unity, cooperation, and peace among nations is progressively forged by the movement of the dialectic as prejudice constantly gives rise to new and higher order

[44] S. Freud, *Freud's letter to Einstein, "Why War?"*, 1932, *SE*, 22, p. 214.
[45] J. Mills, J.A. Polanowsky, *The Ontology of Prejudice*, Amsterdam–Atlanta 1997, pp. 11–13.

forms of novelty and creative complexity. These existential complexities ontologically stand over and above individual practices, for they are mediated in the face of social and cultural interpersonal forces that negotiate and intersubjectively affirm collectively shared value systems and practices over others. As with the epigenesis of the self, the process of this negotiation rests on the nature of identification.

Ideals do not exist in a moral vacuum: they are created by the larger sociocultural milieu that becomes individually and idiosyncratically internalized throughout development; yet they are always open to change and transmutation. These early internalized ideals become the formative basis of a cohesive self and social structure, which remain in flux and unrest due to the dialectical unfolding of the nature of subjectivity and social relations. The parallel process of valuation in individual and collective development is constituted *a priori* within the larger ontological structures that make worldhood possible. Such pre-established ontological conditions (as thrownness) provide the ideal objects of identification that are necessary for selfhood and for the emergence of values – yet they are always up for renegotiation. This emerging process of valuation gives rise to greater *aporiai* in selfhood, communal forces, socio-political drift, and international relations. However, the dialectical nature of prejudice that gives rise to civilization leads to an internal ambivalence, a dilemma it fights within itself. In this sense, values can never be fixed truths or universal essences. Instead, they necessarily materialize out of prejudice, negation, and conflict. Acquiring new life in the wake of destruction, the death of particular values is preserved in the ashes of history, nostalgia, and desire. As humanity elevates itself to higher degrees of complexity, so do its ideals.

From this account we may say that valuation inherently yearns for greater levels of unification and complexity. This would seem to suggest that the structurally constituted dynamic progression of the dialectic ensures that civilization will remain ontologically predisposed to seek and maintain order, accord, and social progression while allowing for a vast variance of novelty, freedom, and complexity to emerge. But with complexity and freedom come the inherent risk of individual psychopathy and social regression that threatens the progressive unification and self-preservative acclivity toward holism. To what degree will progression win out over regression in the face of our contemporary ethnic and religious conflicts? In order to provide a more systematic and rigorously justified account of the constructive forces of civilization within the destructive shapes of worldhood, we need to closely examine Hegel's logic of the dialectic and determine if the positive significance of the negative will in the end vitiate the primitive propensities that compel human relations toward destructive acts.

The Logic of the Dialectic

One of the more interesting aspects of Hegel's dialectic is the way in which a mediated dynamic forms a new immediate. This process not only informs the basic structure of his *Logic* which may further be attributed to the general principle of *Aufhebung*, but this process also provides the logical basis to account for the role of negativity within a progressive unitary drive. The process by which

mediation collapses into a new immediate provides us with the logical model for the improvement of civilization. And it is precisely this logical model that provides the internal consistency to its specific application to the amelioration of pathology. As an architectonic process, spirit invigorates itself and breaths its own life as a self-determining generative activity that builds upon its successive shapes and layers that inform its appearances; therefore, collective mind constructs its own monolith. It is this internal consistency that provides us with a coherent account of the circular motion of the progressive drive toward higher manifestations of psychical, social, and cultural development.

Hegel's use of mediation within the movements of thought is properly advanced in the *Science of Logic* as well as the *Encyclopaedia Logic*, which prefaces Hegel's anthropological and psychological treatment of Spirit. In the *Logic*,[46] Being moves into Nothing which then develops into Becoming, first as the "passing over" into nothing, second as the "vanishing" into being, and third as the "ceasing-to-be" or passing away of being and nothing into the "coming-to-be" of becoming. Becoming constitutes the mediated unity of "the unseparatedness of being and nothing."[47] Hegel shows how each mediation leads to a series of new immediates which pass over and cease to be, as that which has passed over in its coming to be, until these mediations collapse into the determinate being of *Dasein* – its new immediate. Being is a simple concept while Becoming is a highly dynamic and complex process. Similarly, *Dasein* or determinate being is a simple immediacy to begin with, which gets increasingly more complicated as it transitions into Essence and conceptual understanding. It is in this early shift from becoming to determinate being that you have a genuine sublation, albeit as a new immediate, spirit has a new beginning.

In Hegel's treatment of consciousness as pure thought represented by the *Logic*, as well as his treatment of world history in the *Phenomenology*, and the Anthropology and Psychology sections in the *Encyclopaedia*, spirit continues on this circular albeit progressive path conquering each new opposition it encounters, hence elevating itself in the process. Each mediation leads to a new beginning, and spirit constantly finds itself confronting opposition and overcoming conflict as it is perennially engaged in the process of its own becoming. In the *Logic*, the whole process is what is important as reason is eventually able to understand its operations as pure self-consciousness; however, in its moments, each mediation begets a new starting point that continually re-institutes new obstacles and dialectical problems that need to be mediated, hence eliminated.

But thought always devolves or collapses back into the immediate. This dynamic is a fundamental structural constituent that offers systematic coherency to Hegel's overall philosophy of spirit as well as its specific relevance to the problem at hand. Culture mediates opposition and conflict it generates from within its own evolutionary process and attempts to resolve earlier problems unto which new immediacy emerges. Mediation is therefore an activity performed from within society and cultural forces that in turn make new experience possible. When disparate

[46] G.W.F. Hegel, *Science of Logic*, trans. A.V. Miller, Atlantic Highlands 1969. All references to Hegel's *Science of Logic* will refer to *SL* followed by the section and page number.

[47] G.W.F. Hegel, *SL*, §C, 2, p. 105.

cultures and societies are taken together as a conglomerate with endless processes within overdetermined processes, the whole movement of civilization itself becomes an ever increasing logical synthesis.

Hegel envisions this general structural dynamic throughout all contexts of spirit, giving the movement of spirit its logical substance. Each immediacy has a new kind of claim that tests spirit's past shapes, that which in turn must be put into practice in the novel experiences it confronts. Spirit is faced with the tussle of having to take each new immediate and integrate it within its preexisting internal structure, thus incorporating each novelty within its subsisting mediatory faculties. This structural dynamic takes into account the ubiquitous nature of contingency, for spirit is simply not just extending a part of itself as mediation that is already there; it has to incessantly vanquish each new experience it encounters in all of its freshly discovered and potentially unacquainted future environments. The ongoing process of confrontation is the burden of spirit's odyssey, with each encounter signaling a spewing forth from the unconscious well of what it has already incorporated from its past, thus defining the emerging context for each new stage it confronts as unexpected reality.

The Infinite Progress of the Infinite Regress

Through the interaction of mediated immediacy, teleology becomes defined in each moment, with each immediacy being only a moment in the process of civilization. As spirit passes into new stages, it educates itself as it transforms itself, taking on new forms, expanding and incorporating larger aspects of its experience into its inner being. Preparing itself for its next confrontation, it guarantees there will always be a new stage. Because civilization is the self-sublation – what might not be inappropriately called sublimation – of its earlier primitive activity, the logic of the dialectic provides us with a prototype for understanding the underlying functions and power of the negative that propels civilization to overcome its increased oppositions, which it generates from within itself. Because civilization generates division and opposition within itself, each new mediated immediacy allows for contingencies and complexities to operate within existing dynamic structures. This is the freedom of the power of the negative, for it may seek to operate within a destructive and regressive fashion rather than align with the upward current of human growth, social consciousness, tolerance, acceptance, and ethical progress. This further insures that there will always be pathological forms of human activity: the thought that we could ever stop thinking in terms of difference such as ethnicity, religion, or race is simply an illusion.

Human beings will always seek separate unique identities (whether as individuals or as groups) in opposition to others based on the values they choose to identify with. This tendency further guarantees that nationalism, ethnic and religious identity, and separatist movements energized by rigid group identifications will never perish, for identity is what keeps people together; we may only hope that their pathological instantiations will abate and become marginalized to minor aberrations that fail to identify with greater collective global visions. But in all likelihood, we will only see spheres in the amelioration of *pathos* determined by

contingent world events. If we may offer a prediction of the future of civilization based on Hegel's logical model, then perhaps we will see many infinite progresses of many infinite regresses insofar as civilization climbs up the rungs of the ladder, it will also experience slippage, regression, and withdrawal back to earlier manifestations of its being. In this century, this explanation may be said to account, in part, for Hitler and the Holocaust, Stalin's gulags and reign of butchery, Pol Pot's killing fields, Saddam Hussain's gassing of his own people, the genocide in Bosnia and Rwanda, Milosevic's relentless crusade of ethnic cleansing, the grisly 'choppings' in Sierra Leone, and more recently, the terrorist attacks on the U.S. Pentagon and World Trade Center. But as evinced by Hegel's *Logic*, as well as our empirical social advancements, given the synthetic and upwardly mobile acclivity of the dialectic, as a rule, increased ascendance and social unification overreaches the regressive instantiation of annihilating forces.

Yet with each destruction comes a new construction of mediated immediacies that give rise to new values and social ideals. Freud's contribution is invaluable to Hegel's position because, as they both maintain, negativity is the core constituent of life: death and destruction will not only be a universality in all possible future worlds, it is a necessary ontological dynamic that assures the upward progression of change, prosperity, and maturation – the very essence of the striving for our ideal possibility-for-Being.

With most of the world's continents engaged in some form of military conflict, it may be argued (at least theoretically) that the United Nations becomes an ethical paragon for global peace and unity; and with the fight for freedom, democracy, and human rights, social consciousness has made an advance forward. But it has done so at the cost of condemning and displacing other cultural valuation practices that imperil international security, where *might* becomes the *right* of a community in the service of the collective whole. But collective identification has its limits even among nations with focused mutual goals, which further leads to resistance and stifling efforts at negotiation and diplomacy. Because each nation has loyalty to its own self-interests, an ethical diaspora is inevitable: cultural narcissism is highly recalcitrant to outside interference pressuring political reform. This may be observed by the fact that despite indubitable knowledge of the slaughter and concentration camps in Bosnia, former President George Bush Sr. and his administration was not about to send U.S. troops to intercede fearing the ghost of Vietnam, a repetition the Clinton administration faced dogged by a country absorbed by its own concerns. It may be further said that the international community's failure to appropriately intervene in the Rwanda massacre as well as the question of ground troops surrounding the crisis in the former Yugoslavia reflects a collective preoccupation not to uncritically jeopardize the lives of its own citizens.

In these situations, collective identifications that sustain national identities ultimately serve self-preservative functions, for we are bound to identify more with our own kind than a stranger in a foreign land. The brute fact is that we value our own over others: the general principle of human life becomes an abstraction when compared to the concrete social realities each country faces. This is particularly relevant when internal division and upheaval fractures the cohesion of a country's

infrastructures, such as the separatism movement in French Quebec Canada and the omnipresence of discrimination and racism that torments the United States. When industrialized countries, such as the ones in North America, are unable to shelter and provide food and clothing for their own homeless populations who die every night on the streets, they find themselves in the conundrum of determining the most optimal means of disseminating their social resources. Value is ultimately prioritized under the rubric of a particular society's self-interests, but this often encompasses wasteful concessions to popular prejudice. It is truly sad when the American public is more concerned about where the President's penis has been rather than helping the needy through humanitarian aid.[48] This is a fine example of Heidegger's *das Man*, where the herd is lost in the corrupt fallenness of idle talk and curiosity of "the *they*."[49]

The process of civilization vacillates between dialectical moments of progress versus regress as the process itself secures and mobilizes an infinite progression with infinite points of regression. Following the logical coherency of the upward ascendance of the dialectic, we may further estimate that progress will surpass the regressive and destructive forces that tyrannize world accord. What is truly infinite about the evolution of humanity is the process itself. What we see is an infinite (universal) pattern, each side being contrary moments as each merges into the other. This pattern is genuinely infinite for it is a self-maintaining process; each alteration collapses into a new moment, which is its being-for-self in its mediacy. By standing back and seeing the recurrent pattern within a new context, world spirit is enabled to effect the transition to a new immediacy that is truly sublated. Civilization, like Spirit, is always faced with the relative novelty of each new shape. Yet it approaches each new opposition not as a static antinomy doomed to stalemate, but rather as a self-contained pattern; the infinite generates new finites as a fundamental repetition of itself – a self-maintaining process that generates its own process as a dynamically self-articulated complex holism.

Hegel's odyssey of spirit may be applicable to our understanding of the trek of culture and its march over the ever increasing proliferation of human aggression. As our world confederations gain greater amity, consensus, and cohesion, the intersubjective negotiation of valuation gives rise to new novelties, complexities, and increased unity. But the more convoluted social realities become, destructive forces continue to grow in abundance. We may surmise that the insidiousness of human pathology will recede in certain pockets of communal affiliation but flow in others as the valences of prejudice undergo the vicissitudes of transformation. No longer is the standard of culture measured by whether or not one uses a bar of soap, but rather by the values we espouse in relation to others, especially the promise to keep our aggressions in check. Now our degree of civility is to be equated by the mutual agreement not to point our missiles at one another – quite an accomplishment for decades of fear and cold war! Yet this existential reality

[48] Here I am referring to the Clinton-Lewinsky affair.

[49] See Heidegger's discussion of fallenness and inauthenticity in the forms of gossip, curiosity, and ambiguity in Chapters 4-5, *Being and Time*, trans. J. Macquarrie, E. Robinson, San Francisco 1962.

underscores the fact that aggression will always play a part in our value practices and the ontological relations that comprise worldhood.

Whether in politics or business, advances in culture are due to the process of negotiation and mutual recognition, which leads to the mutual desire to understand, communicate value preferences, and support each other cooperatively despite vast differences that define our identities. The need for mutual recognition, validation, and affirmation of cultural values and worth leads to understanding; and in turn, understanding leads to empathy and care. Despite the continuing tumultuous Mid-East peace negotiations and the tenuous Irish settlement where rigid religious identifications insure irreconcilable division, the process signals the human willingness to seek viable solutions in the name of peace, which is itself a productive dialectical movement. Whether they advance in peaceful resolution through mutual negotiation remains a possibility only the future can command. But given the current state of military conflict in the Middle East, where social chaos, lawlessness, mass trauma, civilian revolt, and religious fundamentalism including messianic fanaticism fueled by systemic hatred for the Other, where The End of Days is preceded by a prophetic apocalypse, it becomes a logical prediction that World War III is just around the corner.

When social and psychic conflict remains irresolute, the human species has the compulsion to repeat its traumas in the effort to resolve them. Not unlike Nietzsche's eternal recurrence, the compulsion toward control and mastery may be generally attributed to the *Aufhebung* of cultures as civilization becomes more integrative, refined, and balanced. But as Freud points out, the repetition of destruction is a retrograde character of our species that must be harnessed and channeled into appropriate directions if we are to survive as a human race. It is in the austere face of violence and havoc that continue to pollute our globe where we may observe the pessimistic resonance of Freud's dismal conclusion that offers us "no consolation."

As long as people are deprived of the most basic needs that comprise human necessity, there will always be atypical suffering, seething envy, hate, and murder. And in the uncultivated masses that bleed world tranquility, destruction and violence will be the primary instruments of human deed with each act of aggression begetting new aggression in order to combat it. With the perseverance of peace, perhaps this cycle will culminate in a more docile set of human relations. Through mutual dialogue and the open exchange of value preferences, new ideals, conventions, and policies will emerge, even though this may in all likelihood require the aggressive encroachment on societies and cultures that fail to develop shared global identifications.

Living in the End Times

Hegel once said that human history is the "slaughterbench"[50] of happiness – a progressive yet poignant achievement. But happy or not, happiness is neverthe-

[50] G.W.F. Hegel, *Reason in History*, the Introduction to the *Lectures on the Philosophy of History*, trans. R.S. Hartman, Indianapolis 1953, §27.

less what we covet, what Aristotle called "the highest good attainable by action," that is, "living well" and "doing well."[51] Some of us live well and some of us do well; but for most of the world population, happiness is a foreign reality. Ephemeral moments of pleasure are not happiness: they are not the *eudaimonia* Aristotle envisioned. Even the satisfaction of life's simplest pleasures is often minimized, postponed, or held in abeyance for other desires that have not yet been actualized. Desire is such a complicated creature that it is responsible for generating our most detestable beastly attributes as well as our most cherished and exalted ideals. As being-in-relation-to-lack, desire seeks to assuage its anxiety, to go beyond its finite appearances and fill the hole, the lacunae in its being – simply the wish for wholeness we call peace.[52] The nature of value inquiry is a lived existential ordeal that must endure the gauntlet of anxiety and dread that pave the successive path toward the fulfillment of human *ethos*. It is this positive significance to the power of the negative that becomes the engine behind our moral prosperity even when the dark shadow of our aggressivity and destructive inclinations loom over the sky like a black plague.

"*Fundamental insight*. – There is no pre-established harmony between the furtherance of truth and the well-being of mankind" says Nietzsche.[53] Harmony is made by humankind through the call of conscience and the puissance of reason – a rational passion. And as Freud tells us, the "intellect ... is among the powers which we may most expect to exercise a unifying influence on men – on men who are held together with such difficulty and whom it is therefore scarcely possible to rule... . Our best hope for the future is that intellect – the scientific spirit, reason – may in process of time establish a dictatorship in the mental life of man."[54] Is it such a utopian expectation to think that we can subordinate our pathological natures to the monarch of reason? Perhaps this is the true meaning of faith. For even if there are no emotional ties that exist between people, cultures, or nations, the bonds of reason conjoin us in mutual appreciation for the *ought* that dictates even our most irrational moments.

Having offered this optimistic gesture, we are still left with uncertainty, tenuousness, and platitudes. We are currently experiencing a crisis of liberal capitalism in North America and the UK, which has led to economic exploitation of the masses engineered by those who want to make obscene amounts of money quickly and unabashedly at others' expense. And the rest of the world has its own financial crises, most recently as the United States and members or the European Union are in chaos over fiscal mismanagement with the added irritant of various central banks and China calling for economic reform. When there is the U.S. deregulation of financial institutions, with a conglomerate of powerful lawyers, politicians, economists, accountants, and lobbyists manipulating laws around the exchange of commerce, taxation, and corporate profits, then citizens will naturally be enticed,

[51] Aristotle, *Nicomachean Ethics*, trans. M. Ostwald, Englewood Cliffs 1962, Bk. 1, §4, 15.
[52] See J. Mills, *Origins: On the Genesis of Psychic Reality*, Montreal 2010.
[53] F. Nietzsche, *Human, All Too Human*, in *A Nietzsche Reader*, trans. R.J. Hollingdale, London 1977, p. 198.
[54] S. Freud, *New Introductory Lectures on Psycho-Analysis*, (1933 [1932]), *SE*, 22, p. 171.

taken advantage of, and swindled – partly because some people are gullible and/or unreflective – acting on unbridled wish-fulfillment and whim – and partly because they are simply overburdened by so many demands and responsibilities within our stressful contemporary societies that the pleasure principle is sure to ring its bell. 'We want relief! And be quick about it.'

As Slavoj Žižek points out, like many reformed or neoMarxists before him, such greed is generated by the seductive system of capitalism itself, which beckons only desire and personal self-indulgence. In Freudian terms, we only think about our own gratifications. The vulgar popular expression is: 'Fuck everyone else! The only thing that matters is me.' This is a common sentiment. I further suggest this is a failure of value inquiry, a painful revelation about our collective moral psychology. We are destroying our planet, sullying and stealing natural commodities from impoverished countries that are sold off by corrupt government officials, hence causing new poverty and hunger where people turn to rioting, looting, and killing because they have no rice or wheat, thereby generating social calamities that potentially have world-wide effects. When people are deprived of the basic necessities for human survival, it is illogical to reproach the outrage fueling such desperate acts of aggression and social discord. 'Would you want to be treated like an object that does not matter to the rest of the world?' That is rational conscience speaking. The object here is actually a unity of subjects – a collection of people who are exploited because the Law says they can be. Remember that the Law is what the people in political power determine to be the case. It can change at any time, contingent upon the instrumental functions that bring that about. When collective social systems facilitate ecological risk and allow world erosion to happen, we must critically target the ultimate sources of responsibility. In industrial countries, the populace clearly experiences such trickle-down effects when spikes in gas prices immediately affect the cost of transportation and availability of food and produce. And when people in non-industrial or developing countries have no credit cards to rely on to offset their lack of capital or debt, it becomes a matter of life or death.

As our planet faces increasing desertification due to climate change, environmental pollution, and human spoilage, water wars are generating global concern and economic exploitation.[55] The shortage of drinking water compounded by corporate pollution is resulting in the poisoning of our water supplies. Industrial pollution is credited for producing biologically contaminated water due to the mass dumping of chemicals, pesticides, rocket fuel, and pharmaceuticals discarded by large animal factories and sewage treatment plants. Furthermore, the capitalist privatization of water leads to large-scale neglect of water systems by introducing contaminants that are directly hazardous to citizens, particularly those in disenfranchised nations. Exacerbated by a lack of proper sanitation, millions of people die each year due to water-born diseases, primarily small children up to the age of 5, and as many as 1 out of 10 in India. Industry is always thinking about the 'bottom line' – the cheapest way is the desired norm if you are a businessman. As

[55] For a disturbing account of this phenomenon, see the recent film documentaries, *Blue Gold: World Water Wars*, by Sam Bozzo, and *Flow: For Love of Water*, by Irena Salina.

long as capitalists control water supplies, they can exploit the rest of the world. It is understandable that, under these circumstances, how large groups of peoples would succumb to squalid instinct where 'Everyman for himself!' is the bottom line. Living is squalor, without dignity or a dime, leads to perpetual insecurity, sorrow, bitterness, and the need for revenge. If this continues without global intervention, we can predictably foresee the eruption of mass social psychopathy, where disenfranchised countries become a breeding ground for death, depravity, and terrorism.

Žižek alerts us that the two principle dangers confronting our world today is unbridled capitalism and fundamentalist religious extremism,[56] which he believes is leading to an "apocalyptic zero-point." But the real culprit he focuses on is the global capitalist system itself.

> [The] 'four riders of the apocalypse' are comprised of by the ecological crisis, the consequences of the biogenetic revolution, imbalances within the system itself (problems with intellectual property; forthcoming struggles over raw materials, food and water), and the explosive growth of social divisions and exclusions.[57]

From "invisible" migrant workers deprived of all privileges and used as slave labor in order for us to have dollar stores and cheap blue jeans, to "crazies" who announce their "irrational" intent on using a nuclear device or potent biological and chemical weapons – to a large degree the question of money is always looming in the background. Money motivates everything. We are erecting walls to keep our bordering countries out, and constructing walls within our own nations like 'gated communities' designed to protect us from 'the criminal other.' 'Love thy neighbor' is replaced with mistrust and fear, whereby the Other becomes a threat, source of envy, or persecutory object. Ignoring all warning signs, we are living in collective denial and omnipotent disavowal of our impending peril under the illusory fantasies of grandiose hubris sustained by hegemonic ideologies. 'When will the next natural disaster occur? When will the next bomb go off, or plane fly into a building?' As Žižek puts it: "we know very well that this will happen at some point, but nevertheless cannot bring ourselves to really believe that it will."[58]

Despite the fact that reason tells us to stop the exploitation, corruption, indulgences in excess, and the destruction of our natural resources, we still want. Here wanting becomes a perennial quandary. And human desire wants immediate satisfaction. Its popular motto is: 'Fill the Lack! Chop, Chop!' As a society, we often live in the moment. We want, and want more. This inner mantra becomes an incessant insatiable whine. Restraint, compromise, and self-restriction become an unwelcome trespass. The reality principle is conveniently forgotten when urge, impulse, and caprice are beguiled by immediate objects of tantalizing pleasure.

[56] S. Žižek, *Living in the End Times*, New York 2010, p. 131.
[57] *Ibidem*, p. x.
[58] *Ibidem*, pp. x–xi.

And conscience becomes compartmentalized because the complexities of social order and world discord pose overwhelming impasses to pragmatic resolutions. Here the collective psyche, insofar as it personifies the universal psychological dispositions inherent to humanity, remains in an ambivalent state of inner worry and unmediated conflict.

If we entertain the possibility of the doomsday argument, and that we may very well be living in the end times, then we are approaching a cataclysm that may no longer be preventable. Unlike Freud who had an ambiguous air of skeptical guardedness, yet one with a pessimistic undercurrent, Hegel was more optimistic. I am not so sure we can readily amalgamate the two poles. Each represents a dialectical position and creates an almost insurmountable tension, what Žižek calls a "parallax gap," namely a point of irreconcilable contradiction or antinomy where there is no discernable synthesis.[59] Žižek points toward, in his words, a "pathetic case for communism"[60] as a possible reference point for rectifying our world crises, but this suggestion violates human nature.[61] As long as a macrosystem supports capitalist self-interest, human need and greed will gravitate toward self-pursuit over that of shared equality or the uniform distribution of material wealth through financial egalitarianism. Such a condition, namely global communism, would have to be superimposed on us by a supreme force, hence state intervention, even if it was for our own good. It is more plausible that we will increasingly seek modifications and amendments in neocapitalism reflective of controlled or regulated democratic systems that institute structural improvements through refined checks and balances safeguarding shared collective interests valuing socialistic commitments. Short of a global centralized currency that is regulated, policed, and affects everyone, perhaps some transnational system of socialistic democracy will emerge in the future as a logical synthesis informing the sublation of humanity through natural compromise.

A more pessimistic outcome is that the dialectic would reach an implosive climax or irresolvable breaking point where we have a deadlock of oppositions that lead to snowballed eruptions within multiple social climates that contaminate the world scene with war, famine, and ecological disaster. Even if one pole eventually

[59] Žižek makes the proper locus of his philosophy the parallax gap, where there is a fundamental displacement of difference that poses an "irreducible obstacle to dialectics" based on a pure shift of perspective that can lead to no higher synthesis. See S. Žižek, *The Parallax View*, Cambridge 2006, p. 4.

[60] S. Žižek, *Living in the End Times*, p. 5.

[61] Freud criticizes communism for its naïve philosophy of human nature based on a fantasy principle that ignores the instinctual basis of human aggression. In *Civilization and its Discontent*, he argues that even if private property was not allowed by the state and material wealth was distributed generously among peoples, it would do nothing to eradicate our aggressive proclivities. He elaborates the fantasy that: "If private property were abolished, all wealth held in common, and everyone allowed to share in the enjoyment of it, ill-will and hostility would disappear among men. Since everyone's needs would be satisfied, no one would have any reason to regard another as his enemy; all would willingly undertake the work that was necessary ... [T]he psychological premises on which the system is based are an untenable illusion. [...] Aggression was not created by property. It reigned almost without limit in primitive times, when property was still very scanty, and it already shows itself in the nursery" (*SE*, 21, p. 118).

vanquishes the other, it would be at everyone's expense. If we accept Žižek's premise that unbridled liberal capitalism is a world nemesis bringing the end of days, then it would likely take a series of catastrophic global events before we are forced to curb its enthusiasm. This would necessitate a mass mobilization of political, diplomatic, and economic reform. But as long as there are political hegemonies that drive world relations based on capitalist principles that on one hand clamor 'We are all equal!' but on the other hand admonish 'You have to earn it. There are no free handouts!' then I am afraid we will have to accept the premise of natural law theory – namely, what is natural to do is what is right to do (e.g., 'It is not natural to give away your resources, for this is contrary to self-preservation and self-interest,' or 'Others will have to work for it just like me,' or 'We can't take care of the whole world!' and so on).[62] We can equally imagine the inverse (another dialectic) if we were on the receiving end of exploitation and abuse – 'It is *natural* to kill those who cause us suffering!' And here we return to the question of *pathos*. Some of us *by necessity* have to suffer more than others.

In order to envision salient global solutions, we need more than mere awareness or collective self-consciousness, we must be willing to give up what we are comfortable with in our immediate lives for the sake of the symbolic Other and *act* for a higher principle of valuation, even if this realization is simultaneously motivated by enlightened self-interest and/or self-preservation. I can't see that happening anytime soon. Perhaps the message itself is important enough to reiterate. But if I were a betting man, I would say we are on the brink of extinction.

[62] I would argue that it is 'logical' to have government exercise some control, regulation, or oversight over capitalism based on utilitarian and socialistic commitments that improve the financial stability, social security, and quality of life of the populace, including those controlling the wealth, but it is not necessarily 'instinctual.' What is instinctual or natural is to be concerned first and foremost with one's own affairs, including securing resources for one's immediate familial and communal priorities out of pragmatic necessity. We primarily make decisions based upon emotional unconscious factors – both innocent and prejudicial – that resonates within the deeply felt interior of our beings. These unconscious agentic processes, I argue, have an affective exuberance that is somatically absorbed within our embodied sentience that simultaneously, emotively and semiotically interjects an overarching valuative tone. This naturally includes an amalgamation of the most innately intuitive, feeling, moral, and spiritual sentiments that coalesce within the personality, but also the most primitive, conflictual, agonizing, and phantasmatic. And I would argue much more so than cognition (see J. Mills, *Origins*). Even if we accept the evolutionary argument that memes – the cultural equivalent of genes – drive complex social systems through a replicatory process of mimesis, they would not likely extend past one's immediate social milieu unless other contingencies or environs would ingress upon the motivation systems comprising such social organizations. In other words, we prefer to stay close to home. And sometimes, perhaps more than others, what is natural is not always good. Therefore, natural law or desire governing human nature and interpersonal relations must be subjected to developmental and educational forces that introduce critical analysis, ethical self-reflection, domestic gentrification, and collective valuation practices for the good of all.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

JEFFERY L. NICHOLAS
Providence College
Rhode Island, USA

Toward a Radical Integral Humanism: MacIntyre's Continuing Marxism[1]

Abstract

I argue that we must read Alasdair MacIntyre's mature work through a Marxist lens. I begin by discussing his argument that we must choose which God to worship on principles of justice, which, it turns out, are ones given to us by God. I contend that this argument entails that we must see MacIntyre's early Marxist commitments as given to him by God, and, therefore, that he has never abandoned them in his turn to Thomistic-Aristotelianism. I examine his reading of Marx, with its emphasis on the concept of alienation as a Christian concept, and explain how this reading differs from the dominant scientific-determinist reading of Marx. This examination then leads to a discussion of why MacIntyre abandoned both Marxism and Christianity in 1968. Finally, I turn to his more recent writing on Marx. I contend that if we view them through his argument about the principles of justice and which God to worship, we see MacIntyre's mature philosophy as more Marxist than most people, perhaps even MacIntyre himself, would allow.

> MacIntyre should, therefore, still be read along with Thompson and Marx, not with either conservatives or conventionally academic philosophers. The philosophers have only *interpreted* the world in rival ways; the point is to *change* it.[2]

We have four reasons for reconsidering MacIntyre's relationship to Marxism. First, Paul Blackledge and Neil Davidson note that, despite the importance of MacIntyre's work on Marxism, few books dedicated to his ideas address that work; further, those works that address MacIntyre's engagement with Marxism have failed to fully explore the relationship between MacIntyre's theoretical essays and

[1] This essay is a revision of a lecture delivered at Universidad Sergio Arboleda. I thank Liliana Irizar, Rodrigo Pumbo, Fr. Mauricio, the professors and students of Sergio Arboledo for comments on that lecture.

[2] K. Knight, *Revolutionary Aristotelianism*, [in:] I. Hampsher-Monk, J. Stanyer (eds.) *Contemporary Political Studies*, vol. 2, 1996, p. 896.

his political essays on Marxism.[3] Kelvin Knight contends that MacIntyre "never abandoned Marx's idea of revolutionary practice".[4] So, despite MacIntyre's claim in *After Virtue* "that Marxism is exhausted as a political tradition,"[5] Marxists ideas, such as revolutionary practice and the link between theory and practice, prove important for understanding MacIntyre's mature theory. Further, MacIntyre insists that his critique of liberalism has always been Marx's critique of liberalism.[6]

Yet, the most important reason for discussing MacIntyre's early work on Marx, I shall argue, is that one cannot separate MacIntyre's Christianity – Thomist though it may be – from MacIntyre's political, Marxist commitments. Peter McMylor[7] has thoroughly explored MacIntyre's understanding of Marxism as the inheritor of Christianity in the West. McMylor writes that "the theological nature of MacIntyre's stance is an essential conditioning factor in understanding his initial relationship to Marxism." He cautions, however, that "it would, of course, be foolish to deny development and discontinuity in MacIntyre's thought".[8] Though I agree with McMylor's overall point, I contend that for too long, those engaged with MacIntyre's work have over-drawn the discontinuity between MacIntyre's early Marxist stage and his mature work, perhaps because MacIntyre himself overdrew that discontinuity and emphasized his break with Marxism too forcefully. In fact, unlike others who have criticized liberal capitalism, including the former Marxist Jürgen Habermas and the philosopher-pope John Paul II, MacIntyre has never conceded to capitalism as an acceptable form of economic organization.

I argue that we must read MacIntyre's mature theory – his "Revolutionary Aristotelianism" or Thomistic-Aristotelianism – through Marxist eyes. Even if we appreciate his Marxist past and recognize that he has never rejected Marx fully, we often do not see his current trajectory as Marxist in any sense. In fact, perhaps even MacIntyre himself does not recognize the extent of his Marxist leanings. Yet, if we put MacIntyre's Marx in context, we can come to see a different MacIntyre. Further, we can come to see the need for greater dialogue between Christians and Marxists or, if you will, Thomists and Marxists. Most importantly, we must come to understand our practice and theory, not simply as influenced by Marx, but as inherently Marxist.

I divide my argument into five parts. First, I examine MacIntyre's 1986 essay on how to choose which god to worship. In this article, MacIntyre contends that our initial judgments of justice are ones by which we must choose which god to worship. In turn, however, we later discover that those same judgments were ones inspired by God. MacIntyre's essay sets up my reading of his work in the rest of this essay. Second, I explore the theological nature of MacIntyre's Marxist beliefs.

[3] P. Blackledge, N. Davidson, *Introduction*, [in:] P. Blackledge, N. Davidson (eds.), *Alasdair MacIntyre's Engagement with Marxism*, Chicago 2009, p. 15.

[4] K. Knight, *Aristotelian Philosophy: Ethics and Politics from Aristotle to MacIntyre*, Cambridge 2007, p. 122.

[5] A. MacIntyre, *After Virtue. A Study in Moral Theory*, Notre Dame 1984, p. 262.

[6] A. MacIntyre, 'An Interview with Giovanna Borradori', [in:] K. Knight (ed.), *The MacIntyre Reader*, Notre Dame 1998, p. 258.

[7] P. McMylor, *Alasdair MacIntyre: Critic of Modernity*, London 1993.

[8] *Ibidem*, p. 4.

The concept of alienation proves central to understanding Marx as a Christian heresy. Third, I examine how MacIntyre's reading differed from other Marxisms and why those led to Stalinism. MacIntyre eventually abandons Marxism because it resulted in disunity and inequality. Fourth, I uncover four reasons for why MacIntyre also abandon Christianity. Like the dominant forms of Marxism, the dominant forms of Christianity led to disunity and inequality. In short, MacIntyre abandons Marxism and Christianity for Marxist and Christian reasons. Fifth, I look at MacIntyre's more recent engagement with Marx. I find in this engagement reasons for attending more closely to Marx, which MacIntyre especially articulates in the 1990s. I conclude with some reflections on philosophical practice. Always at the center of my analysis is the concept of alienation and the rejection of disunity and inequality.

I. Reconsidering MacIntyre's Marxism

My argument begins with trying to understand the personal as well as theoretical meaning of a passage that McMylor cites from MacIntyre's 1986 essay "Which God Ought We To Obey and Why?" We have several gods from which to choose. To make that choice, MacIntyre contends that we can use only two criteria: the identity of the god and the nature of the god. The identity of god, however, is revealed only within the sacred texts of this or that religious tradition. So our choosing must begin with the nature of god; that is, we begin with an understanding of God as just. "[U]nless that god is just and is justly owed obedience by us, such obedience cannot be justly required of us".[9]

McMylor points out, correctly, that in making this argument, MacIntyre demonstrates that reason and faith are not two distinct realms or separate aspects of our lives, but unified. "From the fact that we can at one stage in our progress towards God evaluate the divine claims, using standards of justice acquired and elaborated independently of the knowledge of God, it does not follow that in so doing we are judging the Word of God by something external to it".[10] We are able, according to MacIntyre, to reasonably progress in our moral life in relationship to God. Such progress is reasonable because our initial assent to the divine commands followed from our judgment that these divine commands were just. Later, we come to see "that the standards by which we judged God is itself a work of God, and that the judgments that we made earlier were made in obedience to the divine commands... God, it turns out, cannot be truly judged of by something external to his Word, but that is because natural justice recognized by natural reason is in itself divinely uttered and authorized".[11] Reason and faith are unified because both are gifts from God.

I propose that, without over-emphasizing or psychologizing them, we understand these words and this argument on a personal level. As MacIntyre recognizes,

[9] A. MacIntyre, 'Which God Ought We To Obey and Why?', *Faith and Philosophy* 3 [4] (1986), p. 359.
[10] *Ibidem*, p. 370.
[11] *Ibidem*.

his philosophical journey has always been a personal one,[12] and readers of MacIntyre must keep this personal aspect in the forefront of their thinking about MacIntyre, not to dismiss his philosophy, but rather to understand it and to recognize the essential link between theory and practice which lies at the heart of this philosophical work. MacIntyre's essay, "Which God Ought We To Obey and Why?," is written shortly after his conversion to Thomism. I contend that we should read his larger philosophical development in the light of this essay. That reading allows us to recognize, perhaps in a way that at the time he did not, that MacIntyre never abandoned his initial judgments about the divine commands. Rather, he has come to see his earlier beliefs about justice – essentially Marxist beliefs – as, instead, divinely authorized.

He is proposing a unity between his early judgments determined by "natural reason" with his later judgments that the divine commands and the early ethical judgments are both expressions of God's Word. In short, his early Marxist beliefs – the ones he held on to and for which he rejected the IS and Marxism itself – are expressions of God's Word, are, in fact, divine commands.

From the perspective of a new convert to Catholicism, this judgment about the unity of faith and reason points to the convert's initial judgments about both the divine commands and his perception of justice. That is, as I read this essay, it entails that MacIntyre's more mature philosophical position constitutes a progress in moral judgment that is reasonable because his earlier Marxist judgments are now seen as inspired by God and as leading him to the position he now occupies. As shall be evident in my discussion later, MacIntyre's early Marxist writings comprised a very personal attempt to understand how to live both as a Christian and a Marxist. McMylor writes, "It seems clear that what impels MacIntyre towards Marxism, as it is to do a later generation of so-called Liberation Theologians, is in the Christian commitment to practice and to encounter God in the world, amongst, the poor".[13]

II. Alienation From Hegel to Marx

At the age of 24 in 1953, Alasdair MacIntyre published *Marxism: A Critique*. This particular book is next to impossible to acquire, though Blackledge and Davidson have published excerpts in their very important *MacIntyre's Engagement with Marxism*. MacIntyre, however, published a revised and much rewritten version titled *Marxism and Christianity* in 1968. Despite the difference in title, MacIntyre is just as much concerned with Christianity in 1953 as he was in 1968. By 1968, however, he decided he must abandon both Marxism and Christianity: "[in 1953] I aspired to be both a Christian and Marxist, or at least as much of each as was compatible with allegiance to the other and with a doubting turn of mind; now I am skeptical of both, although also believing that one cannot entirely discard either without discarding truths not otherwise available."[14] Notice that even

[12] A. MacIntyre, 'An Interview with Giovanna Borradori', [in:] *The MacIntyre Reader*, pp. 255–266; 'An Interview for *Cogito*', [in:] *The MacIntyre Reader*, pp. 267–275.

[13] P. McMylor, *Alasdair MacIntyre...*, p. 8.

[14] A. MacIntyre, *Marxism and Christianity*, Notre Dame 1968, p. vii.

in 1968, when MacIntyre claims he is skeptical of both Marxism and Christianity, he still holds that each has truths one should not discard. This line of thought supports my contention that MacIntyre's Thomism is Marxist.

In both *Marxism: A Critique* and *Marxism and Christianity*, MacIntyre argues that, with the "division of human life into the sacred and the secular," Marxism is the inheritor of Christianity in the West. "When the sacred and the secular are divided, then religion becomes one more department of human life".[15] The attempt by rationalists in modernity to replace Christian theology with a secular doctrine failed in all but one case: "Only one secular doctrine retains the scope of traditional religion in offering an interpretation of human existence by means of which men may situate themselves in the world and direct their actions to ends that transcend those offered by their immediate situation: Marxism".[16] Marxism offers an undivided understanding of human life. In making this claim, MacIntyre offers the reader a particular understanding both of the function of religion and of Marxism – to provide an interpretation of human existence. "Every individual finds himself with a given social identity, a role or set of roles which defines his phase within a set of social relationships, and these in turn constitute the immediate horizon of his life."[17] An interpretation of human existence allows one to understand and orient herself within her social existence and, thus, provides opportunities for her to seek meaning. Once Christianity has been displaced, only Marxism can provide the individual a social identity, a social identity that defines the horizon of one's life.

MacIntyre sees Marxism and Christianity not as strictly antagonistic to each other. Rather, Marxism is a "transformation of Hegel's secularized version of Christian theology, [and thus] has many of the characteristics of a Christian heresy rather than a non-Christian belief".[18] This transformation is necessary for a secular age.

Following Emile Durkheim, MacIntyre contends that in primitive religions, the concept of the divinity represents the "structure of social life." This representation, then, makes "religious consciousness [...] profoundly conservative." Yet, continues MacIntyre, religion also can be an instrument of change. The great historical religions "have been rich enough *both* to express and to sanction the existing social structure *and* to provide a vision of an alternative."[19] The critical function of religion is possible only "because and insofar as [religion] enables individuals to identify and to understand themselves independently of their position in the existing social structure." MacIntyre contends that religion and society each tell the individual what he is. This disjunction between the voice of society and the voice of religion provides "grounds both for criticizing the *status quo* and for believing that it is possible for him to act with others to change it."[20]

[15] Quoted in P. McMylor, *Alasdair MacIntyre...*, p. 3.
[16] A. MacIntyre, *Marxism and Christianity*, p. 2.
[17] *Ibidem*, p. 2–3.
[18] A. MacIntyre, '1953, 1968, 1995: Three Perspectives On Marxism', [in:] P. Blackledge, N. Davidson (eds.), *Alasdair MacIntyre's Engagement with Marxism*, Chicago 2009, p. 412.
[19] A. MacIntyre, *Marxism and Christianity*, p. 3, original emphasis.
[20] *Ibidem*, p. 4.

Marxism arises as an interpretation of human existence, on this MacIntyrean account, when traditional religion, co-opted by a capitalistic market, can no longer satisfy the longings of the poor, oppressed, and disenfranchised – that is, during the industrial revolution when children are forced to work fourteen hour days and men and women have no time for religious communion on Sundays because they are too tired from work. Marxism is a "social doctrine of man and society that would have the scope and functions of religion" and yet be rational, or "open to amendment by critical reflection at every point, and that would enable men to self-consciously and purposefully achieve such transformation of social life as they wished to see it."[21]

This claim proves all the more important if my thesis is correct that MacIntyre's Thomist-Catholicism cannot be divorced from his Marxism. This Catholicism, like the great religions of the past (of which it is one) that empowered people with a critical insight, must open individuals up to critical reflection "at every turn" and, further, give them the ability to consciously transform social life "as they wished to see it."

In making his argument that Marxism is the inheritor of Christianity, MacIntyre wishes to avoid the weak claim that Marxism simply inherited the function of religion without inheriting any of its content. In particular, MacIntyre contends that the concept of alienation remains central to Marx's thought throughout its development and was abandoned or lost by poor interpreters of Marx, beginning with Engels. To defend this stronger version of his thesis, MacIntyre traces the concept of alienation from Hegel, through Feuerbach, to Marx.

Hegel borrows the concept of alienation from religion. For him, alienation is the condition of human life in a fallen state. Human agents are divided in themselves and from each other. This division is a division, primarily in consciousness. For instance, human agents see morality arising, not from within, but without and opposed to the agent. One fails to obey the moral law and, thus, develops a bad conscience. Likewise, the human agent sees society as something external to his or her participation in it. Thus, the individual agent tries to resist the bonds of society as much as possible, developing bad conscience.

Marx takes this concept from Hegel and combines it with Feuerbach's materialism. Feuerbach's materialism focuses on how human beings reproduce material culture. "Man as a being sprung from nature is a creature of nature, not a man. Man is the product of man, of culture, of history."[22] Marx takes Feuerbach's materialism and focuses on the means of the production of subsistence. For Marx, the concept of alienation describes the situation of the human agent working to satisfy his or her needs. Human agents encounter nature and must work on nature to produce their means of subsistence. This means of subsistence satisfies the needs of human agents, but, in so doing, produces other needs of a material and social kind. The division of labor cleaves society "making of each individual a hunter, a fisherman, a shepherd, and so on" who must now fulfill the demands, not of his

[21] *Ibidem*, p. 5.
[22] Quoted *ibidem*, p. 25.

nature, but of society.[23] The interest of the society, then, takes political form in the state. Though the agent acts in ways that produce the state, the state itself is not a voluntary association, and the human agent sees it as alien.

In Marx's and our day, society is organized according to capitalist markets. "The essential antagonism of society is that between the worker and capitalist".[24] In this society, human agents are alienated from their products, from their work, from their species-being, from community. A worker makes objects that are, not hers, but a commodity she must buy. Work becomes, not a meaningful occupation, but drudgery. The agent seeks to satisfy, not her real needs, but the needs created by the capitalist system. Agents relate to each other, not as comrades or members of a community, but as competitors whose interactions are based on the model of economic-exchange. The religious conception of alienation, which Hegel used to describe abstract human consciousness, Marx uses to describe concrete social relations. "The achievement of Hegel in Marx's eyes was to 'see history as a process in which man is estranged from himself, exteriorizes himself and his work, and then finally comes to his own once more.' The error of Hegel is to see this as a history, not of men, but of abstractions... The achievement of Marx here is to have given historical form to a concrete view of what man in society ought to be, of what he is, and of how his estrangement from his own true being comes about."[25]

The MacIntyre of 1968 reads Marx as formulating a Christian heresy. Marx takes an originally religious concept – alienation – from Hegel and transforms it into a materialist concept. In doing so, Marx is able to provide both an analysis of the human condition and a vision of un-alienated life. The articulation of a materialist conception of alienation as a centerpiece to an interpretation of human existence makes Marxism a secular version of Christianity. It fulfills a role that no other modern philosophy has been able to do. Disagreement over the role of alienation in Marx, however, separates MacIntyre from other Marxists of his time.

III. MacIntyre against the Marxists

MacIntyre believed that his reading of Marx differed significantly from the more dominant reading. The dominant reading provided by Engels and Lenin understood Marx to be giving a scientific-determinist reading of history, one in which the concept of alienation has little place. This reading allowed for the violent Russian Revolution of 1917 and eventually led to Stalin's rise to power. Stalinism is, on MacIntyre's reading, the anti-thesis of Marxism. In 1968, frustrated with the scientific-determinist interpretation of Marx and the atrocities of Stalin and, later, Khrushchev, in the name of history, MacIntyre abandons Marxism and Christianity. In this section, I examine his differences with the scientific-determinist reading of Marx. In the next, I will examine MacIntyre's reasons for leaving Christianity. In the final section, I will contend that his return to Christianity must be understood in light of his earlier Marxist commitments.

[23] *Ibidem*, p. 62.
[24] *Ibidem*, p. 50.
[25] *Ibidem*, p. 57.

If we understand, as Marx did, that human beings are alienated from product, work, species being, and society under capitalism, then we must ask, how do we achieve the communist state. The depression of 1928 did not lead to capitalism's failure, nor did the economic crisis of 2007. Today, capitalists – major stockowners – make more profits than they ever have before. Nor has the increase of industry led to a revolutionary consciousness in the proletariat. Today, just the opposite appears true. In the United States, worker productivity is higher than it ever has been; yet wages remain stagnant and have remained stagnant for close to thirty years.[26] Why has the revolution not come?

MacIntyre believes that Marx offers two reasons for believing in the inevitability of the communist society, one that is optimistic about the proletariat and one that, pessimistically, relies on an elite vanguard to lead the way. MacIntyre will favor the optimistic interpretation, while Engels and Lenin will rely on a scientific-determinist interpretation. In his mature work, Marx defends a theory according to which capitalism is inherently self-destructive and also defends a philosophy of history "according to which all forms of social order are in the end likely to be self-destructive and at the same time creative of new social orders".[27] The theory about capitalism's self-destructiveness is scientific, whereas the second is not. If we treat Marx's claims about capitalism as part of a larger scientific philosophy of history, however, MacIntyre argues that it fails.

According to Marx's theory of capitalism, capitalism must either expand or perish. In a competitive, unplanned economy, the capitalist desires to purchase labor as cheaply as possible – which limits the ability of the proletariat to purchase goods thus diminishing profits – but also desires to sell commodities at a profit – which means that demands for commodities must exceed purchasing power. (The reality of this situation can be seen in Wal-Mart's current dilemma in which they must provide collection baskets for their own employees who do not earn enough to feed themselves and their families.[28]) Thus, in the long run, both profits and standards of living fall. From this analysis, Marx draws two predictions: first, that capitalism chronically cannot distribute goods – that is, make a profit – and that the "large-scale growth of industry will produce an organized and self-conscious working class which realizes that it has no interest in the continuance of this form of social and economic system."[29]

One condition of Marx's theory, according to MacIntyre, is that neither proletariat nor capitalist can exercise agency. Marx's theory relies on the contention that individuals are assigned roles in capitalism, which roles replace "their individual wills," and that these roles are fixed and immutable.[30] Importantly, MacIntyre

[26] L. Mishel, 'The Wedges between Productivity and Mean Compensation Growth', Economic Policy Institute [2012 August 26.]; http://www.epi.org/publication/ib330-productivity-vs-compensation/.

[27] A. MacIntyre, *Marxism and Christianity*, p. 81.

[28] L. Halloran, 'Wal-Mart Food Drive Unwittingly Fuels Talk of Minimum Wage Hike', NPR, [2013 November 22.]; http://www.npr.org/blogs/itsallpolitics/2013/11/22/246558453/wal-mart-food-drive-unwittingly-fuels-talk-of-minimum-wage-hike.

[29] A. MacIntyre, *Marxism and Christianity*, p. 83.

[30] *Ibidem*.

asks, why cannot the individual capitalist become aware of his role and why can he not, therefore, alter his (or her) behavior.

Let me provide an analogy to make this point. Consider Newton's theory of gravity. According to this theory, material objects are attracted to each other, and this attraction means that objects fall to the ground. If you jump off of a building, you do so knowing that you will hit the ground and die. If, however, I am there to catch you, you would not hit the ground. We do not consider that scenario – the one in which I catch you when falling – as a reason to reject Newton's theory. Rather, we see it as an example of outside interference. In acting as an agent, I prevent you from falling to the ground.

Likewise, if the capitalist acts as an agent, then his action can be seen as outside interference. That is, a capitalist's or proletariat's agency, from the perspective of Marx's theory, appears, not as a contradiction to the theory, but as outside interference, for which we should not discount the overall theory. Capitalism has survived, not because Marx's theory is inherently wrong, but because capitalists have exercised agency to prevent the collapse of capitalism.

One problem with MacIntyre's interpretation, is that at times Marx writes as though, not only his theory of capitalism, but his philosophy of history is scientific and that the theory of capitalism is one parcel of the philosophy of history. Karl Popper believes that this confusion in Marx's writing resulted from Marx's confusing a law and a trend. In contrast, MacIntyre reports that Marx, in a letter to a Russian journal, makes just this distinction between law and trend. Instead of confusion in concepts, MacIntyre believes that the problem in Marx results from Engel's interpretation of Marx that emphasizes scientific-determinism and ignores the role of the concept of alienation in the mature Marx.

Engels, according to MacIntyre, conceives of Marxism "as a systematic philosophy of nature as well as of society." On this conception, certain highest-order laws govern all natural and social processes, and the "evolutionary order of nature is matched by that of social progress."[31] Engels, in fact, believed that Marx was similar to Darwin. Where Darwin discovered the science of evolution and its basic laws, Marx discovers the science and laws of history. Engels, in fact, assimilates social science to natural science. In doing so, however, he opens up Marx to the charge of bad science, of confusing a law with a trend.

In contrast, MacIntyre contends that we should read Marx such that the concept of alienation is, not singular to Marx's early writings, but central to his mature work. Alienation essentially names the inability of the proletariat to recognize that the frustrations she feels in the economic system arise from her own agency – her "patterns of behavior."[32] In the communist society, these frustrations are destroyed because human behavior no longer creates contradictions between its needs and its activity. The concept of alienation, as a Hegelian concept borrowed from religion, for MacIntyre, "rests upon a hope. . . on a confidence in what human

[31] A. MacIntyre, *Marxism and Christianity*, p. 88; see also A. MacIntyre, 'Notes from the Moral Wilderness', [in:] *Alasdair MacIntyre's Engagement with Marxism*, pp. 53–53.

[32] A. MacIntyre, *Marxism and Christianity*, p. 89.

beings will be able to make of their lives when certain barriers and frustrations are removed."[33] This Marxist hope, however, is neither religious nor scientific.

On the one hand, Marx's prediction cannot be the same as a scientific prediction. The scientist predicts exactly, and if his prediction either does not come through or the reality differs from his prediction in some way, he must go back to theory and assess what went wrong. Marx, however, denies that we can know what the future communist society will be like. Living as alienated beings, we cannot imagine what the concrete world of the un-alienated – the reconciled – will be like. (Similarly, as Christians, we only know that in heaven we will live without sin, but what such a life will be like, we cannot predict.) Agents consciously create a socialist-communist society and design institutions that serve human purposes (and not institutional ones). This point proves pivotal – because self-conscious agents construct the society, they must be the ones who drive emancipation. No one can bring about emancipation for them (for us).

Thus, concludes MacIntyre, Marx's "prediction" of a communist society rests, not on a scientific law, but on a humanistic hope.[34] Marx's theory that capitalism will eventually self-destruct is different from his theory that human history aims at an emancipated state. For MacIntyre, this point entails that Marxism "is a secularized version of a Christian virtue." Both Marxism and Christianity, however, prove more capable of describing the alienated or fallen state of humanity than of "describing the future nature of unalienated" humanity.[35] Yet, the concept of unalienated humanity is not empty. The end of transformation entails the transformation of "work into a creative activity to be judged by aesthetic standards."[36]

Needless to say, MacIntyre's interpretation of Marx differed not only from Engel's but from a number of more prominent and active interpreters, specifically Lenin[37], Stalin, and Lukacs. The reason for this difference is two-fold: first, Engels' interpretation guided many early Marxists because, second, Marx's economic and philosophic manuscripts, which spell-out the concept of alienation, were not published until 1930. Lenin would only have access to Marx's post-1848 writings. Similarly, Lukacs, writing *History and Class Consciousness* in 1918–1922, would have been limited in his understanding of the importance of the concept of

[33] *Ibidem*.

[34] P. McMylor (*Alasdair MacIntyre*, p. 14) notes that MacIntyre's account in 1968 differs from that of 1953 on this point.

[35] A. MacIntyre, *Marxism and Christianity*, p. 92.

[36] *Ibidem*, p. 93.

[37] Paul Blackledge ('Alasdair MacIntyre's Contribution to Marxism: A Road Not Taken', *Analyse & Kritik* 30 (2008), pp. 215–227) makes a strong argument that MacIntyre simply mis-interprets Lenin in this regard. Specifically, Blackledge argues in contrast to MacIntyre that Lenin did not believe that a vanguard had to lead the people in revolution. Blackledge also contends that Lenin did not necessarily accept a scientific-determinism. While I think Blackledge's argument is fairly strong on this point, it's importance for this work is minimal because I am concerned with what MacIntyre perceived to be a weakness. Given Blackledge's attempt to make Lenin not hold the positions MacIntyre claims he held, I take it that Blackledge would consider these positions weak, if not wrong, as well. See É. Perreau-Saussine's ('The Moral Critique of Stalinism', [in:] P. Blackledge, K. Knight (eds.), *Virtue and Politics*, Notre Dame 2011, 134–152) discussion of this issue.

alienation in Marx. Lukacs, then, defines Marxism as a "consciousness which is constitutive of contemporary social reality."[38] He placed that consciousness, not in the proletariat, but in the communist party. Because of that move, the Communist International denounced Lukacs' writings as voluntarist. Instead, they adopted a scientific-determinist understanding of history, in which history is seen as marching inevitably forward toward the communist state. This thesis becomes paramount for Stalin. "History, according to Stalin, moves forward whether we will it or not; we can assist it or try to retard it, but we cannot change its direction or its goal."[39] MacIntyre concludes that, given the choice between Lukacs' voluntarism and Stalin's scientific-determinism, one is left either deifying the Party or deifying history. It becomes Stalinism. Stalinists [...] made the working class serve the needs of the party and the bureaucracy rather than vice versa [...] [and] believed that the end of achieving communism justified unlimited terror and unlimited deceit as a means.[40]

In "Notes from the Moral Wilderness," an essential essay for understanding MacIntyre's early Marxism and the path that eventually lead to *After Virtue* and a philosophy of the rationality of traditions, MacIntyre contrasts the Stalinist from the moral critic, particularly the ex-Stalinist moral critic. The Stalinist judges his morality according to what he believes is the determined outcome of history – the communist state. The moral critic, on the other hand, judges morality according to a standpoint outside of history. This autonomous standpoint "is the essence of the liberal tradition of morality."[41] The problem lies deeper, however. The moral critic takes from Stalin his understanding of theory. Popper, for instance, attacks historicism, by which he means Stalin's notion of a determinist history governed by laws through which the future is predictable. Popper, nor Stalin nor the moral critic, sees outside the confines of the definition. He identifies theory with Stalinism, and, having rejected Stalinism, can only choose liberalism. MacIntyre seeks something outside the straightjacket of Stalinist dialogue; he seeks, not a scientific-determinist reading of history, but a humanist hope.

MacIntyre proposes a "theory which treat[s] what emerges in history as providing us with a basis for our standards, without making the historical process morally sovereign or its progress automatic." This position entails that certain moral questions need to be re-examined: "What is the relation between what I am, what I can be, what I want to be, and what I ought to be."[42] The key question of *After Virtue* that forces MacIntyre to turn to Aristotle is clearly articulated here. Yet, if MacIntyre's rejection of Stalinism, liberalism, and the ex-Stalinist critic force him to turn to Aristotle, it does so because of his commitment to a different reading of Marx, a reading which, I insist, informs his mature theory. The next section explores the reasons why MacIntyre abandoned Christianity at the

[38] A. MacIntyre, *Marxism and Christianity*, p. 98.
[39] *Ibidem*, 100.
[40] Ch. Lutz, *Reading Alasdair MacIntyre's "After Virtue"*, New York 2012, p. 20.
[41] A. MacIntyre, 'Notes from the Moral Wilderness', [in:] *Alasdair MacIntyre's Engagement with Marxism*, p. 47.
[42] *Ibidem*, p. 57.

same time that he abandoned Marxism. The reason for discussing that rejection of Christianity here is that it helps us to understand how MacIntyre came back to Christianity. This return to Christianity parallels, I believe, a more silent return to Marxism, one that I will explore in the final section.

IV. MacIntyre Against the Christians

> The difficulty is that all the formulations of the Christian religion are politically double-edged. 'All men are equal before God and God wills them to be at one' can either be interpreted to mean that inequality and disunity are a scandal that Christians ought to strive to abolish, or they can be interpreted to mean that it is only before God that men are equal, and only God that can make them at one, so that a merely human equality and unity are neither desirable nor possible. I do not doubt that the original Gospel commands imply the former interpretation; but any Christian who wants to can always rely on the second. As most do.[43]

If the dominant interpretations of Marx put Marxism into question for MacIntyre, then the dominant interpretations of Christianity put it in question. The two opposing and competing functions of religion are, one, to sanction the established modes of social relations and, two, to place those social relations in question by reference to a more perfect state. MacIntyre believes that Jesus meant for Christianity to put into question the social relations of the day, especially those that support disunity and inequality; yet, he believes that many Christians have interpreted Jesus to mean that only God can abolish such disunity and inequality. For the young MacIntyre, Christianity too easily gave in to dogmatism and subverted the Gospel message to politics. Based on this initial juxtaposition, MacIntyre comes to reject Christianity in 1968 on four grounds.

First, historically speaking, Roman Catholicism and Protestant Christianity were both corrupted and justified their corruption through reference to God's word. Too easily, then, Christians have used the Gospel to either ignore injustice in this world or to justify that injustice. Historically, the role of indulgences is only the most egregious example of *pleonexia* in the Church. Further, MacIntyre is all too familiar with Weber's analysis of Protestantism as the root of capitalism. According to Weber, Protestantism places an emphasis on the accumulation of wealth as a sign of God's favor. This accumulation, then, supports a capitalistic approach to the market, accepting with it injustice. In using God's Word to justify *pleonexia*, Christians commit blasphemy.[44] Moreover, the corruption of God's word leads to greater inequality and to a division between those who have and those who have not.

Second, Christianity played reductionism with salvation. "For communism inherits from Christianity the notion of a redemption, a reconciliation of all mankind.

[43] *Ibidem*, p. 64.
[44] *Cf.* P. McMylor, *Alasdair MacIntyre...*, p. 14.

Then, just as Christianity turned salvation for man into salvation for Christians, so communism turned reconciliation for man into reconciliation for the proletarians."[45] In Marxism, this reduction of reconciliation plays out in the Leninist notion of a vanguard to lead the masses to the fully communist society. In Christianity, however, we see different pronouncements of who is and who is not worthy of salvation – Jews, non-Christians, Muslims, Christians of a different sect, Marxists, modernists, etc. This reductionism takes its worst form as orthodoxy: "the corruption of the gospel is the kind of preaching that restricts the new creation to those who are doctrinally orthodox".[46] Obviously, denying salvation to some people creates disunity and enforces present inequality.

Third, Christianity has remained a stranger in a strange world. Of course, MacIntyre writes these words in 1953, but we should not dismiss them too easily. In 1953, the Church had not yet accepted Darwinian evolution and had not yet made an apology to Galileo. It rejected science, despite the work of Gregor Mendel, which is the foundation of modern genetics. As such, Christians educated their children in the classics and in theology. Their education was separate and distinct from non-Christians, which is a cause of concern itself. Because that education was in the classics, moreover, it too easily associated Catholics with "liberal humanism" and the leisure classes of the 18th and 19th century who had time to devote to the study of the classics and the Bible. Thus, we see in practice a division that supports an inequality in the world.

Fourth, MacIntyre, mistakenly he says in 1995, equated Christian theology with the theology of Karl Barth.[47] Barth's theology, however, could not provide an adequate account of moral life. While this point concerns theory, MacIntyre claims that his conclusion was supported by what he saw in the world: "platitudinous emptiness of liberal Christian moralizing... in which the positions of secular liberalism reappeared in various religious guises."[48] MacIntyre has always rejected this liberalism. If such "liberal Christian moralizing" is empty, it will have nothing to say about living morally, which means it cannot address inequality and disunity. In 1968, MacIntyre seems to be associating such liberalism with the attempt by people in the Church to make Christianity as relevant to the secular world. Of central concern is the demythologizing of the Gospel. Yet, for MacIntyre, this move merely acquiesces to the status quo. In contrast, "Hegel, Feuerbach, and Marx humanized certain central Christian beliefs in such a way as to present a secularized version of the Christian judgment upon" the secular world – i.e. a judgment about its injustices[49].

While MacIntyre rejected Christianity as a way of life, he believed a core of Christianity still served subversive purposes: The religion that is untouched by the

[45] A. MacIntyre, 'Extracts from *Marxism: An Interpretation*', [in:] *Alasdair MacIntyre's Engagement with Marxism*, p. 9.

[46] Quoted in Lutz (*Tradition in the Ethics of Alasdair MacIntyre: Relativism, Thomism, and Philosophy* New York 2004, p. 16). See Lutz's discussion of MacIntyre's critique of Christianity (*ibidem*, 15–21), which provides a biographical context to MacIntyre's religious struggles.

[47] A. MacIntyre, '1953, 1968, 1995...', p. 419; *Cf.* 'An Interview with Giovanna Borradori', *passim*; 'An Interview for *Cogito*', *passim*.

[48] A. MacIntyre, '1953, 1968, 1995...', p. 419.

[49] A. MacIntyre, *Marxism and Christianity*, p. 143.

Marxist critique is that which proclaims not the justification of every social order, but the inadequacy of every social order. The grounds of this inadequacy spring from the radical nature of human sin and from the fact that no human order can ever be adequate to the perfection, which God ordains and which is displayed in Jesus Christ. "Such a religion is one that will also be at odds with Marxism in that it will see the corruptibility of communist society as clearly as that of any other society".[50]

In 1953, MacIntyre believed he could live a Christian-Marxist life. In 1968, he abandoned, not just the Christian-Marxist life, but also the Christian life and the Marxist life. If he saw that no human order could be adequate to the Gospel vision in 1953, in 1968 he believed that neither could any Christian order. MacIntyre, through Aristotle, recovered his Christianity – a Thomstic Catholicism. If my argument is right, however, he also recovered a form of Marxism, one which kept alienation at the center of a critique of the world and which was initially inspired by God.

V. MacIntyre's Marxist Thomism

I have been arguing that MacIntyre has a particular reading of Marx that sees Marxism as a Christian heresy because the religious concept of alienation is central to Marx's analysis of the world. The young MacIntyre attempted to live life as a Christian and a Marxist, based on his judgments about inequality and disunity and how both Marxism and Christianity resisted and attempted to overcome such inequality and disunity. He left Marxism because the dominant interpretation instantiated in the acts of Stalin increased disunity and inequality. He also left Christianity because the dominant form of Christianity spread disunity and inequality. In short, his initial judgments about justice, which were inspired by the Gospel and by Marx's *1844 Economic and Philosophic Manuscripts*, drove him to abandon Christianity and Marxism without abandoning those judgments of justice. If we read his 1986 essay "Which God Ought We To Obey and Why?" as a personal statement, however, we must conclude that he believes (or at least ought to believe) that those initial judgments were and continue to be inspired by God. Thus, I contend that MacIntyre's mature work must be read through Marxist eyes.

I wish to contrast the rejection of Marxism as articulated in 1982 in AV with the new engagement with Marxism from 1994 to 1997. In particular, by 1997 in "Politics, Philosophy, and the Common Good" MacIntyre endorses Kelvin Knight's understanding of his politics as "Revolutionary Aristotelianism." Key to Knight's analysis is the idea that MacIntyre has never fully abandoned Marxism. We see in the writings of 1994 and 1997 a way of thinking about MacIntyre's work as a continuation of a Marxist project abandoned by Marx after 1948.

Many label MacIntyre the ex-Marxist, not only because of his public break with Marxism in 1968, but also because of his work in 1982's *After Virtue*. Here, MacIntyre claims that as a political tradition, Marxism is dead. Towards the end

[50] Quoted in Lutz, *Tradition in the Ethics of Alasdair MacIntyre*, p. 16.

of AV, MacIntyre attempts to respond to possible objections to this thesis, the main ones coming from some form of Marxism. He contends in his response to the supposed objections that "[s]ecreted within Marxism from the outset is a radical individualism.".[51] In *Capital* volume 1, Marx imagines the individuals of the communist society to be like Robinson Crusoes who enter civil society through free contracts. According to MacIntyre, no Marxist has been able to explain why such individuals would enter into the communist society and, historically speaking, Marxists have always in the end resorted to some form of Kantian or utilitarian individualism, "the kind of moral attitude which they condemn in others as ideological."[52]

Second, in practice, Marxism has failed on its own grounds. MacIntyre, rightly, notes that the failures of the Soviet Union do not speak against Marxism because they are not honest attempts at Marxism. The attempts of Marxists at their best in Yugoslavia and Italy, however, show that even the best attempts at practice become Weberian. At the core of AV, however, is exactly a critique of this Weberian approach to politics. Thus, Lutz writes "MacIntyre's critique of the social political, pseudoscientific abuse of the social sciences in AV does not mention Marxism or the politics of the Left explicitly, but MacIntyre's Marxist and post-Marxist friends had no question about its intended object."[53]

However, I want to examine what else MacIntyre says here. MacIntyre asserts once more, as he did in 1968, that Marxism is a philosophy of optimism. Yet, the Marxist of the 1980s would "be forced into a pessimism quite alien to the Marxist tradition, and in becoming pessimist he would in an important way have ceased being a Marxist."[54] MacIntyre admits that he shares such pessimistic views because, not only Marxism, but "every other political tradition within our culture" is exhausted "as a political tradition." AV is, from beginning to end, a pessimistic book that contends that the barbarians have been ruling us for some time. As such, we should not expect it to be very Marxist, but for that reason we cannot expect it to be very Thomist (or Christian) either.

Turning to 1995, MacIntyre reflects on his early Marxism. He contends that the original *Marxism* sought to reaffirm central elements of Christianity that many Christians ignored. These elements "are most aptly and relevantly identified by asking what attitude Christians ought to take to capitalism and then noting how that attitude relates to Marxist analysis of capitalism."[55] God calls us to relationships of love; relationships that, through charity, expand upon and still rest on justice. Justice itself requires that we resist and, where possible, abolish "institutions that systematically generate injustice."[56] Referring back to my opening thesis, we can see that, indeed, early MacIntyre judged justice according to God's justice. Those judgments, moreover, are essentially Marxist judgments about alienation

[51] A. MacIntyre, *After Virtue*, p. 261.
[52] *Ibidem*.
[53] Ch. Lutz, *Reading Alasdair...*, p. 35.
[54] A. MacIntyre, *After Virtue*, p. 262.
[55] A. MacIntyre, '1953, 1968, 1995...', p. 412.
[56] *Ibidem*.

and the incompatibility of capitalism with human flourishing. "There is on the one hand the large range of particular injustices perpetrated against individuals and groups on this or that particular occasion, where those other individuals who committed the injustices could have done otherwise consistently with conformity to the standards of profit and loss, of commercial and industrial success and failure, enforced by and in a capitalist economic and social order. The immediate cause of such injustice lies in the character of those individuals who commit them. But there is on the other hand a type of injustice which is not the work of a particular person on a particular occasion, but is instead perpetuated institutionally."[57]

While the first type of evil is individual, it is still systematic, conditioned by the parameters of capitalism. Individuals under capitalism develop as particular types with particular vices that, even without the pressures of the market or the corporation, still result in the oppression and domination of others. The second, however, results from the institutions that agents establish in a capitalist mode of production. This second kind may be more dangerous because often agents alienate themselves from the institutions that their actions produce when they interact with each other. Both of these evils produce a variety of injustices: the original injustice of individuals entering the market on unequal terms which gives power to those who have more over those who have less so that contracts are not free; the absence of any justice of desert, which is found in a just wage and a just market price; the (mis)educational system that trains individuals, not to express their needs and agency, but to serve as cogs in a pre-made machine aimed at capital accumulation; finally, the injustice of the accumulation of money, which makes riches, not an affliction as in the Bible, but an end.[58]

This Christian analysis of the injustices of capitalism "relies in key part, even if only in part, upon concepts and theses drawn from Marxist theory."[59] Thus, "Christianity in turn needed and needs to learn certain truths from Marxism." Again, we see evidence that supports my reading of MacIntyre's 1986 essay. This evidence, moreover, means that we have to understand MacIntyre's mature work, especially its emphasis on practices and traditions, as inherently Marxist.

Continuing in this essay, MacIntyre writes that he rejected more than he should have of Marxism. Moreover, he writes that, free of problematic philosophical assumptions, his return to biblical Christianity and the Catholic Church helped him to come to a new understanding of Marxism – "not only what had been right in official Catholic condemnations of Marxism, but also how much had been mistaken and rooted in obfuscating and reactionary social attitudes. Part of what Catholic theologians – and more generally Christian theologians – had failed to focus upon sufficiently was the insistence by both Marx and Marxists on the close relationships of theory to practice, on how all theory, including all theology, is the theory of some mode or modes of practice."[60] Theology is expressed in historical times. When in good order, the Church makes "intelligible in a variety of contexts

[57] *Ibidem*, p. 413.
[58] *Ibidem*, p. 413–416.
[59] *Ibidem*, p. 416.
[60] *Ibidem*, p. 424.

and practices... the Word of God revealed to it and the world."[61] When theology claims, not subordination, but independence of the Word of God, it reduces to simply another set of competing opinions. MacIntyre concludes by emphasizing the need to learn from Marx's writings of the 1840s.

This conclusion points to the essay "The *Theses on Feuerbach*: A Road Not Taken." Published in 1994, it brings us once more to Marx's writings of the 1840s. MacIntyre's thesis is that, in the third thesis on Feuerbach, Marx laid out a plan for philosophical analysis that he soon abandoned. The third thesis insists that human beings change their historical circumstances and their consciousness through revolutionary practice. This changing of historical circumstances and consciousness rejects the perspective of civil society for something more. Yet, Marx did not have the philosophical tools to say what more that was.

Yet, the concept of revolutionary practice underscores Knight's description of MacIntyre's philosophy as revolutionary Aristotelianism. For Knight, MacIntyrean practices are revolutionary because they challenge the power structures of institutions. "In going beyond the exposure of rational inconsistency in legitimations of modernity, MacIntyre draws on Marx for a critique of its characteristic institutions. He indicts 'the institutional injustice of capitalism' for the alienation and exploitation of labour."[62] While institutions pursue external goods like money, in practices, agents pursue internal goods and virtues. Both are needed for a good life, as Aristotle recognizes in the *Nichomachean Ethics*, but the heart of the flourishing life consists in the pursuit of internal goods. Such MacIntyrean practices as chess and fly-fishing are Marxist, on the account provided here, in two ways: first, they bring together theory and practice, and, second, they imply a revolutionary critique of capitalism.[63]

Importantly, in the 1997 essay "Politics, Philosophy, and the Common Good," MacIntyre endorses Knight's reading of his work, including Knight's insistence that MacIntyre has never abandoned all of Marx. "For an accurate and perceptive discussion of my political views see Kelvin Knight."[64] I want to draw attention to what MacIntyre writes about Marx in this essay. While MacIntyre endorses Aristotle, he holds that Aristotle needs to be corrected on a number of issues. "[A]nd a philosopher who can provide much of what we need at this point is Marx [...] The questions that we now need to put to Marx's texts are [...] questions – about the relationship, for example, of the ineradicable defects of the so-called free market economy to the nature of social activity – answers to which are badly needed by any form of Aristotelianism that aspires to contemporary relevance."[65]

MacIntyre currently calls himself a Thomistic-Aristotelian. Even though Thomist, then, this Aristotelianism needs Marx. Primarily, it needs Marx to identify what is wrong with a free-market economy. As just seen, however, it also needs

[61] *Ibidem*.
[62] K. Knight, 'Revolutionary Aristotelianism', p. 892.
[63] *Cf.* K. Knight, 'Revolutionary Aristotelianism', *passim*; *Aristotelian Philosophy...*, *passim*;
[64] A. MacIntyre, 'Politics, Philosophy, and the Common Good', [in:] *The MacIntyre Reader...*, p. 235.
[65] *Ibidem*, 251.

Marx to understand social reality, not as civil society, but as the foundation for our identity and consciousness. This rejection of civil society returns us, once more, to the concept of alienation as articulated in the early Marx. Marx brings the I and the we together in class struggle. Only in the concept of human nature alone can "morality and desire ... come together once more."[66]

Further, it needs Marx to understand how "revolutionary Aristotelianism" is a theory of some mode of practice, and, further, how Christianity is also a theology of some mode of practice. At the center of MacIntyre's philosophical program from the very beginning is a rejection of disunity and inequality. He rejected Marxism in the form of Stalinism and Christianity in a liberal form because they rested on and perpetuated disunity and inequality. We must see, however, that MacIntyre's conception of revolutionary practice, as well as his conceptions of the narrative unity of life and tradition which I have not discussed, are responses to disunity and inequality. In that sense, then, Thomistic-Aristotelianism also needs Marxism to highlight disunity and inequality in practice so that the theory of practices does not itself become another failed Marxism or another failed Christianity.

***

I began with a discussion of MacIntyre's argument that Marxism is the secular inheritor of Christianity, that, in fact, it is a Christian heresy. MacIntyre puts the Hegelian-religious concept of alienation at the center of his reading of Marx. Marx transforms this conception of alienation from an abstract concept to a concrete one that allows us to examine the objective activity of human life. In contrast to other Marxists, who may not have been familiar with Marx's *1844 Economic and Philosophic Manuscripts*, MacIntyre displaces a scientific-determinist reading of Marx in favor of a humanist reading. Marx's prediction for the collapse of communism is, in the end, based on humanistic hope.

MacIntyre came to reject both Marxism and Christianity. Marxism and Christianity seemed always to lead to disunity and inequality. They both seemed to eventually end in a liberal morality that could not critique the worst actions of their adherents. In turning from Marxism and Christianity, MacIntyre found Aristotle. Through Aristotle, he found Thomas; through Thomas, he found Christianity. Yet, his seminal work, *After Virtue*, seemed to emphasize his split with Marxism.

We have seen that this split was never total. In 1968 he insisted that to abandon Marxism would be to abandon important insights. His later work, while still critical of Marxism, shows that his concept of practices is inherently Marxist. That concept proves Marxist and revolutionary because it provides a focal point for agents to criticize the actions of the market. I have furthered suggested that he remains committed to some of Marx's program – especially the unity of theory and practice and the Marxist critique of a capitalist economy.

My primary aim in this essay has been to show that MacIntyre's mature work must be read through a Marxist lens. This claim means more than acknowledging the Marxist critique of the market in his conception of practices. For one, MacIntyre's Marxism also entails a critique of civil society. More importantly, I claim

[66] A. MacIntyre, 'Notes from the Moral Wilderness', p. 63.

that we need Marx still; we need to understand where Marx went wrong and what we can learn, both from where he went wrong and from what he got right. Further, and this point is most critical, we must recognize that to be Christian in today's world is also to be Marxist.

The 1986 essay "Which God Ought We To Obey and Why?" is crucial for this argument. Regardless of what MacIntyre himself wished to say or imply in this essay, we cannot help but conclude that our initial judgments of justice, Marxist that they are, are in fact inspired by God. They are also God's judgments, and, in this limited though important sense, God also is a Marxist. This conclusion, of course, has significant implications for both Marxists and Christians. Primarily, it entails that Christians and Marxists should, not condemn each other, but dialogue with each other in solidarity. Sadly, this solidarity does not exist in the United States or Western Europe. One practical task we have, then, is to provide opportunities for such dialogues.

It also means, however, that we can begin to share a particular vision of society, one which MacIntyre proposed in his early work. It must be, not a pessimistic vision of AV, but an optimistic one of 1844. "The true Christian community will be one of poverty and prayer. In one sense it will not be specifically Christian, for it will be concerned above all with the truly human [...] But in another sense this new community will be both human and Christian. For its prayer will be the classical prayer of Christendom."[67]

[67] A. MacIntyre, 'Extracts from *Marxism...*', p. 22–23.

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

ANDREW N. CARPENTER
Northcentral University

Historical Materialism, Ideological Illusion, and the Aristotelian Heart of Marx's Condemnation of Capitalism*

Abstract

In the paper author advocates rejecting a prominent criticism of Marx, which holds that his condemnation of capitalism fails because it is based on incoherent, inconsistent moral reasoning. To rebut this criticism he investigates Marxs conception of ideological illusion, arguing that some moral judgments could be true even if people always possess moral beliefs because of ideological illusion. To support this thesis he provides epistemological argument about the nature of epistemic justification, proving that on any reasonable interpretation of knowledge, justification of moral beliefs is possible under even the extreme conditions of ideological illusion. He then introduces the final theme of the essay, namely that strikingly Aristotelian themes lie at the heart of Marxs ethical condemnation of capitalism. By discussing the similarity of Marxs account of alienation to Aristotles account of self-actualization he shows that both thinkers explicitly connect their accounts of development and actualization of human potential and freedom within a society with a sharp critique of economic forms that block or retard that development and actualization.

This paper advocates rejecting two prominent criticisms of Marx. The first criticism, most famously advanced by G.A. Cohen, amounts to a dilemma urging that historical materialism is a fundamentally flawed theory because, even under its most plausible and sympathetic interpretation, it is explanatorily impotent. The second criticism holds that one of the most striking applications of Marx's historical materialism, his condemnation of capitalism, fails because it is based on incoherent, inconsistent moral reasoning; this line of thought is advanced by a number of critics, including, most famously, Richard Miller and Allen Wood.

Assessing both criticisms requires understanding the status and significance of social and cultural phenomena under historical materialism; in particular, and as I argue below, getting to the heart of the first criticism requires investigating

* Polish translation of the paper was published in *Studia Philosophica Wratislaviensia* 5 [4] (2010), pp. 41–64

Marx's account of the relation between productive forces, relations of production, the economic base, and superstructural phenomena and dissolving the air of paradox that motivates the second criticism requires investigating Marx's conception of ideological illusion. In the course this second investigation, I introduce the final theme of this essay, namely that strikingly Aristotelian themes lie at the heart of Marx's ethical condemnation of capitalism. I use this theme to suggest how, *pace* the arguments Marx himself developed , there exists a plausible reconstruction of Marx's condemnation of capitalism that does not rely on the doctrine of historical materialism and conclude that this constitutes an independent reason to conclude that much of value in Marxism survives the powerful critiques I discuss throughout this essay.

I

Historical materialism is a theory which, if true, charts, explains, and predicts the general course of human history. It charts history, first, by providing a method of individuating historical epochs and, second, by producing a systematic account of both the transitions between epochs and the main currents within each epoch. The theory also provides the means of explaining why human history has taken the course thus described. Finally, Marx's theory is meant to show why this course is necessary and, indeed, what pattern will inevitably occur in the future. Famously, Marx maintains that the key to each of the elements in his theory of history–description, explanation, and prediction–is the economic structure of society, understood in terms of the interactions between the productive forces and relations of production.

Viewed from another direction, historical materialism appears more radical. For Marx also holds that, in some strong sense, the economic structure – and in particular the level of the productive forces – "determines" and "explains" the non-economic activities of the so-called cultural 'superstructure'; this is the point of the notorious base-superstructure metaphor. This claim often makes the theory seem implausibly reductive, not to mention absurd. But despite many provocatively misleading passages, Marx's considered opinion is not that human social, religious, intellectual, etc. experience are mere epi-phenomena; far from dismissing superstructural phenomena as unreal, Marx and Engels both acknowledge explicitly that the base and superstructure interact causally, i.e. that the superstructure affects the base as well as *vice versa*. Even so, Marx's claims about the base and superstructure cannot but seem peculiar and radical. He holds, for instance, that the legal and political institutions of any time, or at least their general character, is determined by the nature of the economic structure existing at that time. Likewise, he holds that certain ideas become widespread at a certain time because at that time, by promoting the interests of the ruling or revolutionary class, those ideas serve to stabilize or develop the economic structure. He makes similar claims about, e.g., class composition, class consciousness, artistic trends, social and religious customs – in short, about all of ostensibly non-economic culture. As Marx might put it, on the materialist conception of history consciousness in general at a given time (during a certain epoch) is explained by the nature of production

at that time (of that epoch). One crucial question, of course, is precisely what relation Marx was asserts by making such explanatory claims.

On the reading of Marx developed in the next section, the heart of historical materialism is the thesis that the growth of the productive forces explains the broad contours of human history. Although this "technological determinist" reading is disputed by some contemporary Marxists, I maintain that Marx himself held such a view. More specifically, I believe that Marx held that each epoch of human history is characterized by a distinctive "mode" of production, roughly a distinctive level of productive ability and set of relations of production, and that these modes of productions rise and fall as and because they promote and impede the development of society's productive capability. Although historical materialism is easily summarized at this level of generality it is, as suggested, extremely difficult to understand how one should interpret and attempt to justify the theory's explanatory claims.

But before addressing some of these problems, a rough sketch of how Marx motivates his doctrine of historical materialism is in order. In a letter to C. Schmidt, Engels writes that "according to the materialist conception of history, the ultimately determining element in history is the production and reproduction of real life" (5/8/1890, p. 498). One way of understanding Marx's defense of historical materialism is in terms of arguments which specify necessary conditions for the existence of certain forms of human life, namely the conditions necessary for (1) the existence of any human being at any time, (2) the continued existence of the human species, and (3) the continued existence of human civilization. Marx holds that these various conditions are interconnected in complex ways that motivate the radical claims of historical materialism. Viewed in this way, Marx's general strategy for motivating historical materialism has three steps:

(1) Arguing that the conditions of their existence is such that to exist, human beings must "have a relation with Nature."

(2) Arguing that the conditions of continued human existence (i.e., survival of the human species), and thus *a fortiori* the conditions of humans flourishing in a civilization, are such that each person's necessary relation with nature must be mediated by social interaction with other people.

(3) Arguing that the social interaction necessary for civilization is historically unstable in the sense that a condition of the continued existence of civilization is that the forms of human interaction must change in time and, indeed, must change in certain ways which explain (and necessitate) the broad contours of human history.

The first two steps are relatively straightforward. As Marx asserts in the *German Ideology*, historical materialism begins with premises which are capable of straightforward (and simple) empirical confirmation: "the premises from which we begin are not arbitrary ones, not dogmas, but real premises... [about] real individuals... and their material conditions of life... [and they] can be established in a purely empirical way" (Bottomore, p. 69).

A necessary condition for the existence of any living creature is that certain biological needs be met, e.g., through nutrition. Of course, there are non-biological

conditions for human life as well: astronomical (we must exist in a solar system, on a planet where organic chemical reactions are possible), physical (we must exist in a spatially-temporally extended region where, e.g., atoms are stable and complex elements can be created), and more mundanely geological, climactic, and so on. The crucial difference is that whereas the manner in which those non-biological conditions are fulfilled does not change importantly through history (the general structure of our world explains why they are fulfilled; in Marx's words, they are conditions which people "find already in existence"), humans consciously act to fulfill at least some biological needs and the way they act to fulfill them has changed importantly throughout their history.

Marx uses this distinction in the *German Ideology*, where he remarks that although "all histiography must begin" by assuming that the non-historical conditions for life are met, "historiographic analysis" proceeds by examining our practices of fulfilling certain significant biological needs, specifically by studying "their modification in the course of history by men's activity" (Bottomore, p. 69). He also uses this distinction to criticize theories of history which (he feels) fail to acknowledge the importance of distinguishing the different types of conditions for biological existence. In such theories, Marx complains, "the real production of life appears as *a*historical... thus the relation of man to Nature is excluded from history and in this way the antithesis between Nature and history is established" (Bottomore, p. 71). Significantly, he suggests that part of the reason why historical materialism may seem so radical and absurd is that its critics are captivated with this (in his opinion) bogus picture of history. Thus he argues that because they fail to understand the nature of humanity's relation to Nature, "the exponents of this conception of history have consequently only been able to see in history the political actions of princes and States, religious and all sorts of theoretical struggles" (*ibid*). For this reason, it should be strongly emphasized that Marx's necessary "relation with Nature" is not simply humanity's relation with the universe in general: it is our relation with those needs which we purposively meet in ways which vary through history. In Marx's terminology, our relation with nature revolves around our human ability to produce our means of subsistence.

The second step in Marx's strategy is to defend the wholly plausible assertion that human beings can only meet their needs of subsistence by cooperating, i.e., through social interaction. Although it is possible for individuals to exist in isolation, there are good reasons to think that the human species can be perpetuated in the long run only if individuals strive to meet their needs of subsistence collectively. One set of reasons involves the relative frailty of the human animal (its vulnerability to predators, the environment, long gestation period, etc) which can be offset by social behavior; another involves the conditions of scarcity in which humans are situated (the relative efficiency of organized hunter-gathering, agriculture, etc.). Moreover, it is undeniable that, in fact, humans have interacted collectively to meet their subsistence needs, and also that their doing so is a necessary condition of the existence of human culture and history, and thus also for the truth of any theory of history.

Marx's claim that culture and history could not have developed if people had not come together to fulfill certain biological needs thus seems eminently plausible. What is extremely controversial – and distinctive of the radical nature of historical materialism – is the third step in Marx's strategy, viz. arguing that the course of human history can be explained by appeal to the "story" of humanity's collective production. Historical materialism makes a much stronger claim than that attaining a certain level of productive ability is a necessary condition of possessing a certain level of culture. Marx holds that, in a strong sense I have yet to define, production e x p l a i n s culture and, indeed that the development of our productive abilities e x p l a i n s the development of our civilization. Thus he writes in the *German Ideology* that his "conception of history...rests on the exposition of the real process of production... [which] explains all the different theoretical productions and forms of consciousness, religion, philosophy, ethics, etc., and traces their origin and growth" (Bottomore, p. 70)

In the same passage, Marx suggests what form his argument for this radical claim might take. He writes that the historical materialist conception of history:

> ...rests on the exposition of the real process of production, starting out from the simple material production of life, *and on the comprehension of the form of intercourse connected with and created by this mode of production*, i.e. civil society in its various stages as the basis of all history....From this starting point, it explains all the different theoretical productions and forms of consciousness, religion, philosophy, ethics, etc., and traces their origin and growth, by which means *the matter can of course be displayed as a whole (and consequently, also the reciprocal action of these various sides on one another)*. Unlike the idealist view of history... [historical materialism] shows that history does not end by being resolved into 'self-consciousness', as 'spirit of the spirit', but that at each stage of history there is found a sum of productive forces, a historically created *relation of individuals to Nature and to one another*, which is handed down to each generation from its predecessors, *a mass of productive forces, capital, and circumstances, which is indeed modified by the new generation but which also prescribes for it its conditions of life, and gives it a definite development, a special character. It shows that circumstances make mean just as much as men make circumstances*. (Bottomore, pp. 70–71; my emphasis)

In the highlighted passages, Marx suggests that historical materialism is motivated by an account of the complex interaction obtaining between persons, their productive abilities, their relations in civil society, and their cultural products. The last category includes theological, philosophical, and ethical phenomena and, more generally, peoples' "forms of consciousness." Marx suggests that justification for the thesis that, ultimately, the development of the productive forces explains the course of human history (as well as the general character of civil society, and people's forms of consciousness, etc.) is cashed out in terms of (1) interaction statements specifying the subtle and complex ways in which these things inter-

act and (2) explanation statements specifying that the interaction is such that some of these things can be used to explain the others. Thus my interpretation of Marx's texts considers "the reciprocal action of these various sides on one another" crucially important and takes seriously both parts of the Marx's assertion that "circumstances make men just as much as men make circumstances." As we shall see, there are deep difficulties with interpreting Marx's explanation statements and also with understanding how they and the corresponding interaction statements could be confirmed. In the next two sections I shall discuss two of Marx's arguments about historical materialism from each of two central texts, viz. the *German Ideology* and the 1859 *Preface*.

II

The first argument I will discuss is largely due to statements Marx makes in the *German Ideology*. It is an argument from premises about the material conditions of subsistence to what is commonly called the Development Thesis. In Cohen's words, the Development Thesis is the claim that "the productive forces tend to grow in power throughout history" (1988:3). Cohen's formulation seems misleading in one important respect. As I shall argue below, Marx's theory of history requires – and his arguments for historical materialism attempt to motivate – the stronger claim that *necessarily*, the productive forces tend to grow in power throughout history.

The argument from the *German Ideology* is more general than the argument in the 1859 *Preface.* In particular, most of historical materialism's substantive claims follow from the latter argument. Though detailed, the *Preface* argument is quite obscure; consequently, looking first at the more general argument from the *German Ideology* might help us achieve a deeper understanding of Marx's position.

This is my reconstruction of Marx's general argument for the development thesis:

1. Like all living things, human beings have a necessary relation to nature; a necessary condition for continued biological existence of humans is the continued fulfillment of certain material (biological, physiological, chemical, etc.) needs.

2. Human beings' relation to nature is different from that of other species because they produce their means of subsistence: uniquely, human beings are able consciously to create the means of satisfying some of their biological needs.

3. To satisfy the conditions of continued biological existence human beings must cooperate and thus produce culture.

4. A necessary condition for the "social, political, and spiritual processes" of human life is the ability to produce more than what is required for collective subsistence; humans' productive ability provides them with "a definite way of expressing their life, a definite mode of life."

5. The existence of a particular mode of life presupposes humans attaining a certain level of productive ability – more advanced culture requires more, and more sophisticated, production than less advanced culture.

6. The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, change the productive forces.

6'. The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, develop the productive forces.

7. But a change in a productive forces produces new relations of productions.

8. Therefore, no mode of life can be sustained indefinitely; by virtue of the way that relations of production sustain a mode of life, necessarily the productive forces advance and different modes of life develop.

The first three premises are claims about the material conditions of human existence and correspond to the first two steps of Marx's general strategy described above. They are plausible, easily-justified empirical statements; indeed, as I suggested above it seems that any plausible historical theory must be committed to similar premises. Similarly, I think that the fourth and fifth premises are empirical claims which, though less obvious, also can be adequately motivated. However, the sixth, seventh, and eight steps are quite troubling. Peculiarly, Marx seems to consider (6) and (6') as equivalent. Below I argue that although there are reasonable grounds on which (6) is true, Marx produces no plausible motivation for (6'). This is unfortunate, because I will argue that it is (6'), not (6), which is essential for the derivation of the Development Thesis.

Having already discussed the first three premises in Section I above, I'll immediately turn to the fourth premise. This step has two parts, one fairly obvious and commonplace and another less obvious and faintly Hegelian in character. The more obvious part is the claim that to maintain the social interaction necessary for mutual survival, society's total production must be above the subsistence level. In the *German Ideology* after characterizing the way in which communities produce their subsistence needs (in Marx's language, "reproduce their physical existence") as a mode of production, Marx writes "this mode of production should not be regarded simply as the reproduction of the physical existence of individuals" (Bottomore, p. 69). It is certainly true that in the history of our civilization the level of production has been above the subsistence level. It also seems plausible to assert that communities which only produced at the subsistence level would not be stable in the long term. On the one hand, it seems that if people didn't (at least partially) recognize the benefits of mutual cooperation, communities would tend to be unstable because individuals would have no preference to stay in a community. On the other hand, it seems true that stability-enhancing sociological or political structures can be maintained, at least in the long run, only by the use of additional production. Given the claim in the third step that social interaction is necessary for the continued existence of our species, it would follow from this that producing above the subsistence level is required for continued existence of humanity. Thus there seems to be an adequate empirical justification for the first part of the fourth premise.

In any event, the claim that there is surplus production is, like the claims in the previous premises, conditionally necessary in the sense that any theory of history or civilization must accept it as true: if there were no surplus production, there would be nothing for such theories to discuss. It is important that this conditional necessity, e.g., that all theories of history must assume that there is some produc-

tion beyond subsistence, not be confused with stronger modal claims, e.g., that it is necessarily true that there be production beyond subsistence. It seems possible that humans culture could not have progressed beyond the stage of subsistence existence (and, given steps 1-3, that it might therefore have become extinct). Likewise, the claims in steps one and two are conditionally but not strictly necessary. To specify the conditions necessary for human history is not to prove the necessity of human history. Because they are premises in an argument for the Development Thesis, it is important that the claims in the first three steps not be interpreted as having more than conditional necessity. Obviously, interpreting these premises as asserting that it is necessarily true that these necessary conditions are fulfilled would make the argument for the Development Thesis invalid; likewise, if an interpretation with only conditional necessity cannot be sustained, then it is proper to conclude that Marx's argument begs the question.

In the *German Ideology* Marx connects the unproblematic claim that culture requires production above level of bare subsistence with a rather strange, Hegelian claim linking personal identity with modes of production. He writes:

> The way in which men produce their means of subsistence depends in the first place on the nature of the existing means which they have to reproduce. This mode of production should not be regarded simply as the reproduction of the physical existence of these individuals. It is already a definite form of activity of these individuals, a definite way of expressing their life, a definite *mode of life*. As individuals express their life, so they are. What they are, therefore, coincides with their production, with *what* they produce and with *how* they produce it. What individuals are, therefore, depends on the material conditions of their production. (Bottomore, pp. 69–70)

Marx seems to be making at least two distinct (but related) claims here, viz.: (1) a person's mode of life in some sense depends on the means of production of that individual's community; and (2) there is some relation of correspondence between different modes of production and different modes of life. These claims – and especially the explanatory statements which correspond to them – clearly are extremely relevant to the full doctrine of historical materialism.[1] However, I believe that a rather truistic reading of them will suffice for our discussion of Marx's more limited argument for the Development Thesis. For our purposes these weak claims will suffice: (1') living in a society which produces above the level of subsistence is a necessary condition of possessing a cultural identity (i.e., a distinctively cultural mode of life) and (2') some modes of life are not possible in a society with a level of production x which are possible in a society with a level of production y, where x is much smaller than y. Each of these claims is rather obvious, if not trivial.

The fifth and sixth steps follow quite easily from step four. Step five is a paraphrase of weak condition (2'). It merely asserts that there is some relation of

[1] They are also interestingly reminiscent of Hegel, namely because Marx seems to claim that a person's sense of identity arises from his productive interaction with his environment and society,

correspondence between different modes of production and different modes of life, viz. a relation such that to more advanced forms of life there corresponds more advanced forces of production. Given the reasonable assumption that in a stable society the relations of production serve to stabilize the forces of production, it follows that to a stable mode of life there will correspond relations of production which stabilize the productive forces, which is the first part of step six.

A summary of the premises just discussed might be helpful before continuing:

4. A necessary condition for the "social, political, and spiritual processes" of human life is the ability to produce more than what is required for collective subsistence; humans' productive ability provides them with "a definite way of expressing their life, a definite mode of life."

5. The existence of a particular mode of life presupposes humans attaining a certain level of productive ability – more advanced culture requires more sophisticated production than less advanced culture.

6. The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, change the productive forces.

I believe that claim that the relations of production must act to change the productive forces follows trivially – too trivially – from the earlier steps. For if the relations of production have a stabilizing function on the productive forces, they must affect them. But if the productive forces are affected by the relations of production, then they are changed by them. Earlier I argued that it would be absurd for a theory of historical materialism to deny that the productive relations and forces interact, i.e., that the relations of production and productive forces are both effected by the other and affect the other. It trivially follows from this that they also change each other: something cannot be effected without changing *in some way*.

The important question is whether this assertion of trivial change is sufficient for Marx's purposes and, if not, how he might justify less trivial statements about changes in the forces of production. Consider the rest of his argument:

6. The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, change the productive forces.

7. But a change in a productive forces produces new relations of productions.

8. Therefore, no mode of life can be sustained indefinitely; by virtue of the way that relations of production sustain a mode of life, necessarily the productive forces advance and different modes of life develop.

We have seen that it follows from the plausible assertion that the relations of production (perhaps constantly) act to stabilize the productive forces that they (perhaps constantly) change them. On the surface, the transition from step six to steps seven and eight seems reasonable. Truly, since to different productive forces there corresponds different relations of production, the interaction between them entails that through history there is a (perhaps constant) succession of different productive forces and relations.

But this is true only as far as it goes, which is not far enough. Namely, only this follows from steps six and seven:

> (T) It is true (merely) that the relations of production continually change the productive forces *in some way*, and that this entails that there is in history a succession of productive forces and relations that are different *in some way*.

If it is thought that the relations of production and productive forces undergo constant interaction (as they surely do), then it follows from Marx's argument that they are always changing. But this type of change is too frequent for historical materialism, which tries to account for discontinuous (epochal) change. This suggests that the change involved in (T) is not the type of change which historical materialism attempts to explain intended to explain, and that the constant interaction posited between the productive relation and forces is the wrong kind of interaction.

Yet couldn't Marx could produce an argument whereby epochal change is explained by appeal to the constant interaction between the productive forces and relations? I believe that this is impossible, at least on the basis of (T). Presumably any such argument would specify certain necessary conditions for epochal change to occur, for instance that epochal change occurs when constant change has changed the relations of production to extent x or when it has gained the set of features y while lost set of features z, etc. That is, any such account will have to explain why the constant interaction tends to produce *certain determinate changes*. But it is in principle impossible for (T) to ground judgments like this. (T) is a truth about changes conceived qua changes, in a completely general and indeterminate way. Thus (T)'s truth is consistent with the constant change generating a state of equilibrium where the changes become progressively smaller and more insignificant and do not provoke new modes of production or life. Similarly, (T) is consistent with the changes causing a kind of negative feedback loop such that the productive relations and forces change in a regular, cyclical cycle; indeed, *prima facie*, this kind of change appears more plausible than change where the relations' stabilizing function inevitably leads to them "contradicting" the very productive forces they stabilize.

Truly, (T) is consistent with *any* proposed account of change. But this is tantamount to saying that it can support no conclusions of the kind required by Marx, because it cannot provide evidence for no account of change. In particular, it can provide no evidence either for the claim that the changes are of a certain type, nor that the changes will tend to take place in a certain pattern. Thus in Marx's argument there is no way that steps six and seven can support the conclusion that the productive forces and relations of production change in a way that makes modes of life unstable. The problem is that nothing in Marx's argument guarantees that changes which advance the productive forces need occur; unless he – somehow – can support a conclusion much stronger than (T), Marx's general argument for the Development Thesis is invalid.

It seems appropriate to turn to Marx's texts, to see whether they exhibit any signs of a better attempt to motivate the Development Thesis. Unfortunately, Marx rarely explicitly formulates the Development Thesis, much less explicitly

defends it. Consider this passage, from *The Poverty of Philosophy*, where Marx describes how the forces of production affect modes of life:

> Social relations are intimately connected with the forces of production. In acquiring new forces of production, men change their mode of production, their way of earning their living; they change all their social relations. The hand mill will give you a society with the feudal lord, the steam mill a society with the industrial capitalist.
> The same men who establish social relations in conformity with their material power of production, also produce principles, laws, and categories in conformity with their social relations. Thus, these ideas and categories are no more eternal than the relations which they express. They are *historical and transient products*. (Bottomore, pp. 108–109)

Although this passage is about "new forces of production," in the text quoted Marx says nothing about how these new forces are acquired. First Marx notes that one of the effects of acquiring a new level of effect is that the economic structure of society, including the relations of production, is transformed: "the hand mill will give you a society with the feudal lord, the steam mill a society with the industrial capitalist." That is, to changes in the productive forces (of the relevant size) there corresponds changes in the mode of production. Next he asserts that the new productive force also has great effects on superstructural phenomena like "ideas and categories," which, he asserts, correspond to – are "in conformity with" – the new economic structure. Thus he is asserting that the change in productive forces elicits change in the economic structure which, in turn, is reflected in things like ideological phenomena. As I have interpreted him, Marx explains the acquisition of new productive forces by asserting that these changes further reflect upon and develop the productive forces – he closes the circle, as it were.

It is no mere coincidence that in the lines immediately following the passage just cited, Marx asserts the Development Thesis and, indeed, characterizes it in terms of a closed circle of interaction with the relations of production and culture generally. The next lines are:

> There is a continuous movement of growth of the productive forces, of destruction of social relations, of formation of ideas; nothing is immutable but the abstract movement – *mors immortalis.* (Bottomore, p. 109)

Marx clearly states that the production of the productive forces is part of a "continuous movement.' Namely, he lists two movements: (1) growth of the productive forces and (2) the creation of new social relations and new ideas. Obviously, Marx believes that these motions are necessarily related: they are two aspects of "*a* continuous movement," a single movement which is also "*the* abstract movement." This coheres fully with my reading, according to which Marx accounts for the necessary development of the productive forces by appeal to a constant interaction between the forces and relations of production. And consider how Marx

characterizes the nature of this constant interaction: "nothing is immutable but the abstract movement – *mors immortalis*." This surely tells us what Marx believes is the "immoral death" at the heart of the Development Thesis: by this he can only mean that "the destruction of social relations" is the effect of the prior growth of productive forces, and the cause of the next.

As we've seen, the bare conclusion that the modes of production and modes of life and the forces and relation of production causally interact cannot establish the Development Thesis. This strategy of argument from mutual interaction can succeed only if it is shown that these changes must have a particular determinate character, viz. only if Marx can show that why the relations of production must affect the productive forces in a way that causes them to develop. Yet there seems to be no such argument in his texts. In fact, to return to my reconstruction of his argument, Marx seems merely to *assume* that these claims are equivalent:

(6) The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, change the productive forces.

(6') The continued existence of a particular mode of life presupposes the existence of certain relations of production which, through their interaction with the forces of production, develop the productive forces.

The conflation of these claims is, to borrow Strawson's famous criticism of Kant, nothing less than a non sequitur of numbing grossness. Not only does (6) entail only indeterminate claims about changes – and therefore is not even a similar kind of statement to (6') – but, as I've shown above, it is incapable of grounding any account specifying the type of change which must take place.

The failure of Marx's strategy is disappointing, because it seemed a promising way of establishing the necessity of the Development Thesis. Recall that on my reading the Development Thesis is the thesis that:

(DT) Necessarily, the productive forces advance through history.

One response to the line of thought I've pursued is to argue that a weaker, contingent version of the Development Thesis will suffice for historical materialism, viz.:

(DT') The productive forces tend to advance through history.

I do not think that this line of thought, – which, for example, seems to be defended at Cohen 1988:22ff[2] – could be used to defend Marx's own views: his unqualified claim that the communist age is inevitable commits him to defending the strong version of the Development Thesis (and many other troubling modalities as well). In the next section I will sketch the traditional, detailed account of historical

[2]Thus after "admitting... that I do not have a good answer to the question of how productive forces select economic structures" (1988:17), he defends something like (DT') by an argument that "the *source* of the development of the forces" is a *preference* for labor-reduction which, in turn, is at least partially explained by appeal to (the obviously contingent fact of) humanity's being situated in conditions of scarcity.

materialism in the 1859 *Preface* which includes the prediction of communism and will discuss several aspects of the notorious "problems of explanation" that that many critics believe cripple this account.

III

Perhaps it is better to characterize the text in the 1859 *Preface* as a summary of, and not an argument for, historical materialism. Certainly if Marx attempted to justify historical materialism in this text, his attempt was a dismal failure; the *Preface* provides almost no insight into why Marx believes the theses promulgated there. Nevertheless, the *Preface* is often treated as if it were Marx's classical argument about historical materialism. Most of the "argument" in the *Preface* consists of discussing various features of the complex network of interactions which Marx maintains occur between (1) the base and superstructure, (2) the relations of production and the productive forces, and (3) the development of the productive forces and the relations of production and superstructure. Marx believes that some of the interactions are explanatory: thus, for example, he asserts that "the mode of production of material life determines the general character of the social, political, and spiritual processes of life" (Bottomore, p. 67). This is my reconstruction of Marx's main claims in the *Preface*:

1. The relations of production within a society constitute an economic structure which is the basis 'on which rises a legal and political superstructure, and to which correspond the definite forms of social consciousness.'

2. To a given set of relations of production in a society there corresponds 'a definite stage in the development' of that society's productive forces; together these things constitute the 'mode of production' which explains 'the general structure' of a society's superstructure.

3. One feature of the mode of production is that the relations of production which correspond to a society's productive forces promote the development of those forces.

4. But as a society's productive forces develop, there gradually (but inevitably) occurs a certain disequilibrium in the mode of production; 'the material forces of production come in conflict with the existing relations of production' in the sense that 'from forms of development of the forces of production the relations of production turn into their fetters.'

5. When they interact so as to fetter the productive, the productive forces and relations of production (inevitably) interact with the superstructure in a manner which 'begins an epoch of social revolution.'

6. In this revolutionary period, 'the immense superstructure is more or less rapidly transformed.' In particular, these changes (inevitably) follow the fettering of the productive forces: (i) the normal division of society into competing classes is intensified and (ii) people become conscious of their position qua members of a class and consciously wage class revolution, viz. the exploited class engages in revolution against the ruling class.

7. The interaction between the productive forces, relations of production, and superstructure explains why, necessarily, the success of a revolution against the ruling classes is inevitable.

8. To the revolutionary transformation of the superstructure there (necessarily) corresponds a 'change in the economic foundation'; in particular, after every successful superstructural revolution there exists new relations of production which promote the development of the productive forces.

9. This cycle continues: the new relations of production come to fetter the development of productive forces and a new historical epoch is ushered in through class revolution.

10. But, inevitably, at some point this process of historical development comes to an end and a new era of history begins: necessarily, at no time after the revolution which overturns 'the bourgeois relations of production' will the relations of production come to fetter the development of the productive forces; 'the bourgeois relations of production are the last antagonistic form of the social process of production.'

Thus as I understand his text, central to Marx's historical materialism are claims about a complex series of interactions among the productive forces, relations of production, and a wide variety of superstructural phenomena. More specifically, the *Preface* portrays historical materialism as holding that many of these things interact in such a way that they are functional for the historical development of each other. The complexities of the interactions in Marx's account of historical materialism are neatly highlighted in this passage from *Capital III*:[3]

> We have seen that the capitalist process of production is a historically determined form of the social process of production in general. This process is, on the one hand, a process by which the material requirements of human life are produced and, on the other hand, a process which takes place under specific historical and economic conditions of production and which produces and reproduces these conditions of production themselves, and with them the human agents of this process, their material conditions of existence and their mutual relations, that is, their particular economic form of society. For the aggregate of the relations in which the agents of production stand to Nature and to each other, and within which they produce, is precisely society. (Bottomore, p. 164)

In addition to summarizing several strands of interactions important to historical materialism, this passage also evokes the second central feature of the *Preface* argument, viz. that all the interactions posited happen because they are functional for "reproducing the conditions of production themselves." That is, in addition to maintaining that the productive forces, relations of production, and all the varied superstructural phenomena interact with each other, it is a central claim of historical materialism that some of these things *explain* the others. In fact, to most of the interactions statements listed above corresponds an explanation statement.

[3]In addition, this passage is also evocative of Marx's general strategy, according to which an analysis of the "process by which the material requirements of human life are produced" is of crucial importance.

In addition to those listed in steps two and seven above, the principal explanatory claims are:

1' A society's base (i.e., its relations of productions) explain its superstructure.

2' From (2), by the transitivity of explanation, the development of a society's productive forces explains that society's superstructure; "the mode of production of material life conditions the social, political, and intellectual life process in general."

4' The growth of productive forces in a society explains why this fettering occurs.

5' The existence of fettering relations in a society explains why revolution is inevitable in that society.

6' The fettering relation between the relation and forces of production explains this intensification of class warfare and development of class consciousness.

7' Revolutionary victory is inevitable because successful revolutions are functional for developing new relations of production which are better adapted to the productive forces.

8' Their being functional for the continued development of the forces of production explains the development of the new relations of production.

9' Thus, by the transitivity of explanation, the general course of human history is explained, ultimately, by its being functional for promoting the development of the productive forces.

10' In particular, this explains why the communist mode of production will inevitably develop: necessarily, the 'prehistory of human society comes to an end' because its ending is functional for promoting the development of the productive forces.

I alluded in section I above to deep problems in understanding historical materialism's explanatory statements. One fundamental problem is understanding precisely what Marx thinks is explained by what. For instance, when he asserts in the *German Ideology* that "the struggles with the State, the struggle between democracy, aristocracy, and monarchy, the struggle for the franchise, etc. are merely the illusory forms in which the real struggles of the different classes with each other are fought out" (Bottomore, p. 111) he seems to mean that political struggles are "illusory" (in some sense) because they are explained (in some way) by the existence of the class struggle. What exactly is the class struggle meant to explain? For instance, is the political conflict "illusory" because the class conflict determines it in every detail? Or, does Marx have in mind that what is explained by the class struggle (and, ultimately, by the development of the productive forces) is why certain political ideas and movements emerge at a certain time?[4] Whereas the former interpretation would make historical materialism absurd, the latter claim would support an interpretation on which historical materialism seemingly has a much greater chance of being true.

In fact, I think it is fairly clear that Marx intends his explanations to be understood as being of the second kind. It is less clear what kind of explanatory relations

[4]In this case the thing that Marx is denouncing as illusory would be something like the alleged "historical autonomy" in the "marketplace of ideas" of political opinions.

Marx could be talking about. Although the issue is too complex to be resolved by simple citation, this passage from *Capital I* implies that Marx believes that the explanations in historical materialism are of a piece with functional explanations in evolutionary theory:

> Darwin has aroused our interest in the history of natural technology, i.e., in the formation of the organs of plants and animals, as instruments of production for sustaining life. Does not the history of the productive organs of man, of organs that are the material basis of all social organization, deserve equal attention? And would not such a history be easier to compile, since, as Vico says, human history differs from natural history in this respect, that we have made the former, but not the latter? Technology discloses humanity's mode of dealing with Nature, the process of production by which he sustains his life, and by which also his social relations, and the mental conceptions that flow from them, are formed. (Bottomore, pp. 78–79)

Suppose that the explanations of historical materialism are interpreted (i) as making claims of the less radical, more plausible type discussed above and (ii) as functional explanations. To what sorts of claims would the theory be committed, for example when Marx asserts (iii) that the base explains the superstructure? If (i) and (ii) are assumed, I believe that explanation (iii) would then be unpacked roughly as follows:

> One sense in which the development of the relations of production explains the main course of human history is that (1) within a certain epoch, it explains the "general features" of the major superstructural phenomena and also the broad outline of their development throughout that epoch, (2) it also explains why the general features of these superstructural phenomena differ in certain major ways in each different historical epoch, (3) it explains the general nature of the superstructural phenomena occurring during transition between epochs, and (4) it gives a functional explanation of why the base explains the superstructure, viz. because the general features of the superstructure thus explained are functional for enhancing (in the early parts of an epoch) or impeding (in the later part of an epoch) or optimizing (in the period of transition between epochs) the rate of the development of the productive forces.

Of course, a functional reading of historical materialism would be committed to many other functional explanations as well, for example to (4') there would correspond the functional claim that the relations of production vary as they do throughout history because their development is functional for producing the optimal rate of advancement of the productive forces.

Successfully sustaining a functional reading would require overcoming daunting exegetical theoretical issues and also serious problems. about functional explana-

tions. Although these problems seem to represent a thicket best avoided,[5] the bare possibility of a functional reading suggests a potentially serious challenge to my critique of Marx's argument for the Development Thesis. For perhaps it is not mysterious why changes in the production relations promote the development of productive forces: perhaps this can be explained functionally. But no matter what deep theoretical questions plague the theory of explanatory relations, it is manifestly true that an explanatory relation cannot be asserted for which there is no justification whatsoever. Yet for the same reason that they cannot provide evidence for causal explanations, the interaction statements in Marx's argument are wholly incapable of providing any justification for functional explanations: in principle, the sixth and seventh steps can only support wholly indeterminate claims about the changes effected on the productive forces by the relations of production. Unless my analysis of non sequitur can be defeated, on no interpretation, functional or otherwise, can Marx's argument for the Development Thesis be valid.

Even if this general argument were good, we have seen that Marx's detailed "argument" for historical materialism in the *Preface* faces a host of exegetical and critical questions which becloud the prospects of defending Marx's theory of history. And without a plausible argument for the Development thesis, Marx's theory seems as good as dead. On the one hand, that Marx apparently does not even attempt explicitly to motivate most of the substantive claims made in the *Preface* gives a strong general reasons for doubting his controversial theory. On the other hand, the failure of Marx's argument for the Development Thesis provides a specific reason for rejecting his motivation for the thesis which constitutes the heart of his historical materialism.

If the prospects defending Marx's historical materialism look bleak, the possibility of developing other – but distinctively Marxist – theories of historical materialism seems to be slightly better. We have seen that the question of what types of explanation it is possible for historical materialist theories to sustain is hotly contested. While some philosophers (e.g., Cohen) insist that only functional interpretations of historical materialism are possible, others (e.g., Elster) argue that historical materialist explanations must be interpreted causally because it is in principle impossible to justify functional explanation in social theory. Lastly,

[5] To give one example of the complexity of these questions:

It can be argued (controversially) that these are necessary conditions for distinguishing between mere functionality (B's being functional for A) and instances of genuine functional explanation (A's explaining B because of B's functionality for A):

(1) One must be able to confirm the claim that whenever conditions are propitious for optimizing (the relevant feature of) A, Bbecomes functional for A; and

(2) One must be able to explain this pattern by confirming that there is a mechanism such that "behavior at one point in time gives rise to consequences that have the effect of maintaining similar behavior at a later time." But, it can be argued (controversially), the functional explanations of historical materialism cannot be confirmed because in their case neither condition can be met; for example, it seems that some of the events which are allegedly explained functionally do not repeat enough times for the first condition to be met. (Interestingly, the debate about condition (2) has strong parallels with certain questions about epistemic justification. At places Cohen exploits the parallels by construing the question as "what would it be for one to justifiably believe that such a mechanism exists."

I shall attempt to demonstrate that the failure of this argument shows why much of the discussion of historical materialism "in general" – that is, about the possibility of constructing some theory of historical materialism – doesn't consider relevant interpretive possibilities.

As I argue above, the claim that the superstructure effects the base is fundamental to Marx's account of historical materialism. Truly, it seems almost absurd to deny this claim, and thus presumably any reasonable interpretation of historical materialism must be committed to it. Yet we have also seen that it can be hard to reconcile the claims that the superstructure affects the base and the base explains the superstructure. One might suspect that a deep philosophical problem lies behind this difficulty. For the joint assertion of these claims can seem problematic: is it not absurd – or, perhaps, paradoxical – to assert that an effect affects its own cause?

If there exists a philosophical problem here, then any possible defense of historical materialism seems to face a dilemma. If the theory is not interpreted as holding that the base and superstructure causally interact, then it is committed to an absurd epi-phenomenalism. Yet it seems that if the theory holds that the superstructure affects the base, then the problem just described means that it cannot hold that the base "determines" or "explains" the superstructure. On similar grounds, G.A. Cohen seems to think that any causal interpretation of Marx's explanatory relations is impossible and concludes, apparently by an argument from elimination, that the explanatory relation must be functional. After discussing instances of superstructural effects on the base, Cohen writes:

> ... That seems to refute the doctrine of base and superstructure, since here superstructural conditions – what legal rights – determine basic ones – what their economic powers are. Yet although it [this instance of the base being affected by the superstructure] seems to refute the doctrine of the base and superstructure, it cannot be denied. And it would not only seem to refute it, but actually would refute it, were it not possible, *and therefore mandatory* (for historical materialists), to present the doctrine of base and superstructure as an instance of functional explanation. (1988:9; Cohen's emphasis)

If the dilemma I sketched above were sound, then functional interpretations of historical materialism would not only be possible, but would be "mandatory for historical materialists."[6] Although I do not deny that a functional construal of historical materialism is possible, I do not believe that it is necessary. For both

[6] And perhaps functional interpretations would be possible only if (as Cohen thinks but Elster seems to deny) an adequate account of functional explanation as a *distinctive* form of explanation could be defended. If the dilemma were good and Elster were right in characterizing functional explanation as a kind of imperfect causal explanation – see, e.g., Elster 1986:206 – then it seems that historical materialism would be incoherent under any interpretation (unless some third type of explanation was developed and shown to be relevantly different).

However, it seems that Elster's conception of functional explanation is wrong for roughly the reasons that Cohen gives: some accepted biological explanations would, if interpreted causally, generate claims of reverse causality. (Somewhat ironically, Cohen's argument about the logical

the dilemma I've sketched and Cohen's (similar) argument are incompatible with two fundamental features of historical materialism, viz. its historical perspective and its fine-grained analysis of the base and superstructure and the productive forces and relations of production.

Consider the joint assertion of these two causal claims:

(1) A's causal action on B explains B; and

(2) B causally interacts with A.

Consider, next, this unproblematic reading of (1) and (2): at some time t, A causes B and at some time t', after time t, B causally interacts with A. This, clearly, is an instance where it is consistent to assert that one thing can causally explain another thing with which it interacts; *pace* Cohen and the putative dilemma, statements like (1) and (2) can be jointly asserted without contradiction.

Since any serviceable reading of the main notions of historical materialism (base and superstructure, productive forces and relations, etc.) obviously attributes temporal persistence to them, causal interpretations of historical materialism of this type are possible in principle. A logically consistent causal interpretation of historical materialism could hold, for example, that the relations of production at the beginning of an epoch (1) interact with that epoch's forces of production and (2) were caused by the forces of production obtaining at the end of the prior epoch.

Indeed, one could consistently hold both that the base explains the superstructure and undergoes constant interaction. Consider these assertions:

(3) For any time t, at time $tA's$ causal action on B explains the "general nature" (or "macro characteristics" or "broad course", etc.) of B.

(4) For any time t, at time $tB$and Aundergo "micro level" causal interactions.

In addition to ignoring historical materialism's historical perspective, the putative dilemma wrongly treats the base and superstructure as if they were simple – i.e., atomic – objects. This is clearly mistaken: the superstructure is complex, both because it contains many parts (including political features, sociological features, artistic features, religious features, intellectual features, etc.) and because it can be described at both a micro-level – in terms of the details of the kinds of features just listed – and at a macro level, for example as a superstructure which supports the interests of the ruling class and inculcates a bourgeois ideology. Assertions (3) and (4) seem perfectly compatible when read in terms of this structural dimension: at any time the base could causally interact with the superstructure so as to explain its "general nature" (i.e., an explanation of its macro form) while, simultaneously, the base and superstructure causally interact on a micro level.

All this is not to argue in support of "causal" versions of historical materialism. My point is that neither the apparent dilemma facing causal readings, nor Cohen's similar *a priori* "logical" argument for functional readings justify excluding causal historical materialisms from consideration. It seems that the types of interrelations posited by historical materialism (viz., the complex interconnections between

necessity of a functional interpretation is wrong for reasons which Elster would accept: as I argue, in some circumstances B can causally interact with A and A can causally explain B without generating reverse causality.)

productive forces, relations of production, the economic base, and superstructural phenomena adumbrated above) could be established, in principle at least, by either causal or functional accounts.

I submit that this conclusion should promote a broader debate about historical materialism. It obviously provides richer interpretive possibilities than many readers of Marx have considered: we now understand that non-functional materialisms are possible, and also that the presumptive greater plausibility often accorded to functional interpretations is misguided. Moreover, rejecting considerations of reverse causality advances discussions about the general plausibility and possibility of defending (some interpretation of) historical materialism, namely by separating this discussion from the question of whether or not functional explanation in social sciences is possible. Since non-explanatory versions are a distinct possibility, what often appears to be implicitly assumed is false: it is not true that the discussion of the possibility or impossibility of historical materialism rises and falls with debates about the conditions for successful functional explanations.

Put another way, historical materialism has more interpretive possibilities – and perhaps more life – than many have acknowledged.

IV

My argument thus far seeks to demonstrate that Marx's signature doctrine of historical materialism is more sophisticated and more interesting than prominent critics have assumed. In addition to seeking to highlight some aspects of Marx's subtlety and sophistication, this exegetical work also has a wider purpose: I wish to rebut some prominent objections to Marxism in order to create space for contemplating the value of Marx's ideas in our century. As I seek to demonstrate below, one source of significant value is Marx's ethical critique of capitalism, which I believe applies, in a novel and powerful way, a strikingly Aristotelian conception of human flourishing to a trenchant analysis of economic conditions that remain in effect today but of which Aristotle could not conceive.

To make this case, however, I must address a second prominent criticism of Marx, namely that Marx's views on morality are paradoxical and that his condemnation of capitalism is inconsistent. On the one hand, Marx seems to believe that any set of moral principles is ideological in the sense that their acceptance at a given time can be explained by reference to the mode of production then existing (i.e., to the development of the relations and forces of production) and to their being functional, at that time, for promoting the interests of a particular – usually the ruling – class. Marx's belief that moral judgments are grounded on such ideological illusion' is commonly, and mistakenly, I believe, taken to show that Marx cannot consistently hold that there are eternal (i.e., trans-historical) moral truths, because he is committed to holding that moral concepts are wholly and irredeemably historically-relative. It is this claim – that Marx cannot consistently maintain that there are historically-independent moral truths – which lies behind the charge of paradox. For the paradox is allegedly generated because, on the other hand, manifestly a sustained and incisive moral critique of capitalism is present in Marx's writings.

As one would expect, attempts to resolve this alleged paradox have centered on questions like 'could Marx consistently criticize capitalism by appeal to *proletarian* justice,' 'could Marx's condemnation of the evils of capitalism be based on a conception of *non-moral* evil,' 'could Marx's claims that morality and justice are ideological be intended only ironically or figuratively,' or even 'could Marx sometimes have been mistaken about Marx's own considered opinion about justice.' If Marx's views were paradoxical in the way that is assumed, then these would be interesting and important questions to ask. However, I believe that such questions are non-starters: because Marx's writings (about morality) are in no way paradoxical, it is much easier to provide a sympathetic reading of his texts than is commonly held.

The reason why there is no deep problem of coherence or consistency among Marx's "anti-morality" and "pro-morality" texts is that an ideological account of the genesis or spread of moral *opinions* is perfectly consistent with a "moral realism" about moral *truths*. That is, there is no reason why Marx could not consistently hold (1) that every set of moral principles, and every moral judgment based on those principles, arises or becomes widely-held at a certain time because this is functional for realizing some class interests and (2) that some of *these very same principles and judgments* are (eternally) true, and others are (at all times and under all modes of production) false. To think otherwise is to confuse genesis with justification, which is a fallacy. There is no question of a paradox: as I shall now attempt to show, some moral judgments could be true – and known to be true – even if people always possess moral beliefs because of ideological illusion.

To be successful, charges of moral paradox in Marx's texts must be supplemented with arguments about the nature of epistemic justification. The question of paradox turns on questions about knowledge. One relevant question is what conditions must obtain for someone to justifiably believe a moral truth; a second is whether or not it seems plausible to hold that these conditions can be met under circumstances of "ideological illusion." I shall argue that on any reasonable interpretation of knowledge, justification of moral beliefs is possible under even extreme conditions of ideological illusion.[7]

Consider this strong thesis about ideological illusion:

(1) For any person at any time, if that person has a moral belief, then he has that belief because he is under the spell of ideological illusion.

According to (1), in principle every moral belief of every person is caused by ideological illusion. That is, (1)'s extremely strong claim is that – regardless of a person's historical or cognitive state regardless of whether, for example, a person

[7]One problem with this line of approach is that Marx's comments about knowledge are few and far between, not to mention cryptic. Even if he had a coherent view, it seems highly unlikely that Marx's theory of knowledge could ever be determined. Yet if I am correct in holding that the "ideological" status of moral judgments is consistent with the possibility of moral knowledge under all but highly implausible conceptions of knowledge, then it follows that it is unreasonable for critics to focus on the alleged paradox in Marx's moral theory and practice. Unless it can be shown to arise on plausible theories of knowledge, discussion of the paradox seems unimportant both to an evaluation of Marx's texts and, more generally, to a discussion of Marxist condemnation of capitalism.

lives under communism or knows all about the theory of ideological illusion – humans cannot have a moral belief which does not arise *because* of its ideological character (e.g., because it serves to promote the interests of one class or another in the mode of production at a certain point in history). I note that (1) is consistent with (2):

(2) Certain moral judgments are true.

Suppose that people happen to entertain and accept certain moral notions when and only when and, indeed, because their doing so is functional for the optimal development of the forces and relations of production. Clearly, this does not entail that people never have true moral beliefs. For it seems at least *possible* that (1) some moral principles or judgments are in fact correct and (2) at some time for "ideological" reasons people might come to accept certain moral judgments which happen to be true. Again, it seems possible that were there a God or were (*pace* (1)) there a time in the past or future when people are not bewitched by ideological illusion, then it *would* be possible to come to have moral truths in the normal way, viz. by engaging in non-ideologically motivated moral reflection. It being impossible in practice for us to undertake such moral reflection would have no bearing on whether, in principle, there are moral truths which such reflection could discover.

It is perhaps less obviously true that (1) is consistent with (3):

(3) Certain moral judgments can be justifiably believed (i.e., known) to be true.

Suppose that, although they are caused by ideological illusion, a person happens to have only true moral beliefs and, in addition, among his other beliefs are beliefs which justify those true moral beliefs. The alleged paradox in Marx's moral thought is motivated, I believe, by the assumption that in this case a person would not justifiably believe the "ideological" moral truths. That is, even if Marx need not deny that there are moral truths, it would be paradoxical for him to make moral judgments if his account of ideological illusion entailed that moral truths are unknowable. I believe that the claim that (1) is inconsistent with the possibility of moral knowledge presupposes a conception of epistemic justification which is untenable because under it *all* knowledge is impossible.

Philosophical questions about justification center on specifying exactly what cognitive relation with a belief's justification is necessary for justifiably believing it. It is hotly contested, firstly, whether a person must know, justifiably believe, or believe a belief's justification and, secondly, whether he must believe, justifiably believe, or know that the justification justifies the belief. The range of answers given to these questions is great. Epistemic externalists deny that a person's justifiably believing something requires the person having *any* cognitive grasp of that belief's justification. Epistemic internalists, on the other hand, hold that justified belief must involve believing (or justifiably believing, or knowing) whatever justifies the belief, and perhaps also believing (or justifiably believing, or knowing) that the belief is justified by those things. There are at least six distinct forms of epistemic internalism and several kinds of epistemic externalism.[8]

[8] *Prima facie* one might think that arguments from ideological illusion would be ineffective against any form of epistemic externalism. Yet externalists with a causal account of justification

Consider our example where ideological illusion both causes a person to have a certain moral belief, which happens to be true, and also causes a person to have other beliefs, which happen to provide evidence sufficient to justify the moral belief. In this case the person has a true moral belief and has other beliefs which justify it. Presumably the reason why there would not be knowledge under conditions of ideological illusion is that the person doesn't believe the beliefs for the right reason. Clearly, the person might *possess* any belief required for justification; the argument that he cannot have moral knowledge must hinge on the claim that he does not have at least some of these beliefs for the right reason, viz. because he has them *because* they promote certain class interests and not because of the proper epistemological reasons (e.g., because they are true, because they are known to be justified, etc.).

That is, the objection from ideological illusion seems to be based on a conception of justification according to which certain evidence justifies a person's belief only if he justifiably believes or knows that the evidence justifies the belief. By holding that ideological illusion explains *why* the person has the latter belief, claim (1) defeats this kind of justification: the person merely believes, and does not justifiably believe or know (because the belief is only due to ideological illusion), that the evidence justifies his belief. Consider these definitions of two types of epistemic internalism:

> (D1) Theory T is an instance of Strong Justification Internalism = DF. T holds that if e justifies $S's$ belief b, S justifiably believes b only if (1) S justifiably believes e; and (2) S justifiably believes that e justifies b.
>
> (D2) Theory T is an instance of Strong Knowledge Internalism = DF. T holds that if e justifies $S's$ belief b, S justifiably believes b only if (1) S knows e; and (2) S knows that e justifies b.

Under either theory of justification, the strong thesis of ideological illusion (claim (1) above) entails that the second condition is not met for moral beliefs: person S could merely believe that evidence e justified belief b, but could not justifiably believe or know this because (by hypothesis) he believes this *because*

would probably agree that beliefs believed *because* they promote class interests do not have the proper causal link to whatever makes them true. (Indeed, such externalists are likely to be moral skeptics precisely because on nearly any account of morality there are no such causal links to moral truth.) But other externalists would be unmoved by such considerations: they argue that one may justifiably believe something just in case it is true (or perhaps if it could in principle be justified, for example by God or by people not in the grip of ideological illusion.

Thus one point to make about the argument from ideological illusion is that on this second kind of externalist account of justification, admitting that a certain belief is true ((2) above) is tantamount to admitting ((3) above) that it can be known. Since it seems that considerations about ideological illusion do not call (2) into question, (3) could be defended on externalist grounds; certainly defending the paradox charge would require producing an argument against the second type of externalism. But epistemological externalism is a fairly unpopular view. In the text I argue that the only form of internalism under which (1) is incompatible with (3) involves a conception of justification which is implausibly strong.

he is under the grip of ideological illusion.[9] Thus a paradox might exist if such a conception of knowledge were defended (or at least if it could be shown that Marx held such a view). The trouble is that this conception of justification is implausibly strong; knowledge is impossible under this account of justification. Theories of type (D1) and (D2) hold that one justifiably believes b only if one justifiably believes that e justifies b. But, by parity of reasoning, one justifiably believes the second-level belief that e justifies b only if one justifiably believes the third-level belief that e_2 justifies believing that e justifies believing b.

These conceptions of justification are absurd because they generate this vicious "levels" regress:

1. A necessary condition for justifiably believing b is justifiably believing (a), that e justifies b.

2. A necessary condition for believing (a) is justifiably believing (b), that e_2 justifies believing that e justifies b.

3. A necessary condition for believing (b) is justifiably believing (c), that e_3 justifies believing that e_2 justifies believing that e justifies b.

4. A necessary condition for believing (c) is justifiably believing (d), that e_4 justifies believing that e_3 justifies believing that e_2 justifies believing that e justifies b.

And so on.

By contrast, consider these forms of epistemic internalism:

> (D3) Theory T is an instance of Strong Belief Internalism = DF. T holds that if e justifies $S's$ belief b, S justifiably believes b only if (1) S believes e; and (2) S believes that e justifies b.
>
> (D4) Theory T is an instance of Weak Justification Internalism = DF. T holds that if e justifies $S's$ belief b, S justifiably believes b only if S justifiably believes e.

Neither of these conceptions of justification generates the vicious levels regress. More importantly, under neither conception is justification defeated by the ideological illusion hypothesis. As we have seen, claim (1) is consistent with a person having any belief, and therefore with both having the evidence for belief and believing that the evidence justifies the belief: claim (1) is compatible with strong belief internalism because it is *possible* that ideological illusion generate all the requisite beliefs. In the case of weak justification internalism, similarly (though less obviously), moral knowledge is possible under conditions of ideological illusion: since justifiably believing evidence e does not require justifiably believing that $e$justifies b, the fact that ideological illusion causes believing e is irrelevant to $b's$ justification.[10]

[9] That is, if S believed that e justifies b, he could not justifiably believe or know this because, *ex hypothesi*, this belief is "ideological" and S therefore does not stand in the proper epistemological relationship (according to (D1) or (D2)) to it.

[10] There is a regress of justification under weak justification internalism, but it clearly is not the vicious "levels" regress generated by strong knowledge or justification internalism. Rather,

The point is that there seems to be no good reason for supposing that the hypothesis of ideological illusion would defeat the justification of true beliefs arising from such illusion. The only plausible account of why claims like (1) would defeat justification rests on an untenable account of justification; furthermore, there are alternate accounts of justification on which (1) clearly does not defeat justification.

Thus the charge of paradox can be saved only by drastic and heroic measures: to show that paradox exists in Marx's texts one must discover and defend a conception of knowledge on which ideological illusion defeats justification. In the absence of such an argument, Marx could denounce "bourgeois morality and justice" as ideological – i.e., as arising because it is functional for promoting the class interests of the bourgeoisie – and consistently condemn capitalism *on these very moral principles*. It also follows that Marx consistently could hold that all his moral beliefs arise because of ideological illusion (viz., because they promote the class interests of the proletariat) and also claim that they are eternally true.

If these extremes are consistent, so much the better for Marx's actual, more moderate position! In fact, although he need not to avoid paradox, Marx does eschew the traditional morality of rights and justice for another, somewhat Aristotelian ethical theory based on the ethical unacceptability of unnecessary alienation. Similarly, Marx's actual claims about the ideological status of morality are fairly moderate – roughly, that apologists for capitalism employ bogus ideological rationalizations – and he does not hold that his own ethical judgments are ideological. Again, however, no paradox would be generated if his claims were as strong as claim (1).

The rest of this essay considers the ethical basis of Marx's condemnation of capitalism, as well as his dismissive claims about the ideological nature of bourgeois morality and justice. Hopefully, having avoided at the outset the tortuous, false trail of the famous Marxian paradox of morality will facilitate discussion of these issues.

V

Above I argued that no paradox would ensue if Marx's condemnation of capitalism were based upon an application of those bourgeois moral principles which he dismisses as ideological. Yet I also held that Marx's condemnation of capitalism, like his commendation of communism, is not based on principles of humanitarian justice, the universal rights of man or on any other principle of bourgeois morality. Rather, Marx's evaluation of modes of production is based on other,

it is the "normal" regress of justification which epistemic internalism is intended to solve–in the case of weak justification internalism, a solution might be to adopt a holistic conception of justification. If there is a problem here with weak justification internalism, it is a wholly epistemological problem.

The point is that such considerations are inessential to my argument against the existence of the "morality paradox" in Marx's texts. For my purposes, it is sufficient to show (1) that apparently the only plausible account of why a belief's arising from ideological illusion would defeat justification relies on a rather implausible account of justification and (2) that there are alternate, more plausible accounts of justification on which even the strongest ideological illusion hypothesis does not defeat the justification of moral judgments.

non-bourgeois moral principles. Even if it were somehow inconsistent for him to employ those very judgments he denounces as ideological, I believe that Marx would not be guilty of this inconsistency: Marx's condemnation of capitalism is based on ethical principles which are not ideological.

One apparent problem is that it is difficult to square this interpretation with the seemingly plausible assumption that non-ideological moral principles must be, in some strong sense, trans-historical or non-historically-relative. For, manifestly, Marx's evaluations of modes of production are *not* based on principles which are "universal" in this sense: undeniably, he bases his moral evaluations on his dialectical theory of history. It is clear, for example, that Marx subscribes to no ethical principles which condemn capitalist exploitation in all circumstances. Truly, he praises capitalism as ethically necessary for creating the material conditions under which communism is historically possible; insofar as some alienation is essential to the capitalist mode of production then that alienation likewise is ethically necessary.

Although such moral judgments may appear problematically relative, I believe that his making them is consistent with Marx's holding a non-ideological ethical theory based on eternal moral truths. Before turning to his texts, it will be helpful for me to sketch my understanding of Marx's ethical theory. Alienation seems to be central to Marx's ethics; these three premises, in particular, seem fundamental to his position:

1. Avoidable alienation is morally wrong.
2. Actions which alleviate conditions of alienation are morally commendable.
3. These ethical principles are lexically prior to all others.

I address Marx's extremely complex notion of alienation in section VI below. For the moment, it is important merely to note that my claim is that whatever alienation is, all of Marx's moral judgments revolve around it. I believe that, for Marx, avoidable alienation is the fundamental moral evil. Similarly, I argue below that he considers the unalienated development and manifestation of human capacities to be the fundamental moral good. It follows from this that for Marx, actions which alleviate conditions of alienation are morally praiseworthy.

The third step, the claim that premises one and two override or "trump" all other ethical considerations, is necessary for understanding some of Marx's puzzling failures to condemn alienation in capitalism. My claim is that Marx, like Rawls, holds that certain ethical strictures must be met before any other moral evaluation can take place. Put another way, I believe that Marx would reject any moral reasoning which contradicted premises one or two. Thus, for example, he would reject complaints that actions necessary for alleviating alienation are morally unacceptable on grounds of justice, or because they are inconsistent with allegedly inalienable human rights.

I've suggested that Marx's ethical theory is dialectical in an important sense because it relies heavily on his dialectical theory of history. Historical materialism is an extraordinarily optimistic doctrine. It holds, for example, that human existence under communist relations of production would be radically free of alienation; Marx believes that, unlike any other period of human history, unalienated

existence would be possible during a communist epoch. Historical materialism also makes the strong claim that the alienation rampant in all other modes of production can be avoided *only* under a communist mode of production. From this, these steps follow:

4. As the doctrine of historical materialism shows, the world can be de-alienated only by the establishment of communist relations of production.

5. From 2, actions promoting the establishment of communist relations of production are morally commendable.

6. From 1, actions impeding the establishment of communist relations of production are morally wrong.

It seems possible that actions which impede the development of communism could nevertheless reduce some unnecessary alienation, and thus for Marx would be morally commendable. Yet given Marx's belief that the world can be radically de-alienated and that this can take place only under communist relations of production, it follows that "on balance" such reforms would not be morally praiseworthy. Thus steps five and six hold that for Marx, actions promoting or impeding the establishment of communism are, respectively, in general morally correct and incorrect.

In fact, Marx insists that simple (i.e., non-revolutionary) reforms of non-communist modes of production usually do not reduce alienation at all – he claims that characteristically such reforms are functional for promoting the aims of the ruling class, and therefore for increasing alienation. Thus in the light of his theory of ideology, it follows that Marx considers "bourgeois" ideological morality ethically wrongheaded because it promotes and does not alleviate alienation. These steps follow:

7. From 5 and 3, "bourgeois" reforms to capitalism motivated by bourgeois moral principles are morally commendable if they promote the establishment of communist relations of production.

8. But such bourgeois reforms – and bourgeois morality as well – are ideological: they promote the class interests of the bourgeoisie, which impedes the establishment of communist relations of production.

9. From 3 and 6, therefore bourgeois morality and bourgeois reforms of capitalism are morally wrong.

On the ethical theory sketched above, it follows that if they are ideological in the sense described, then bourgeois reforms and morality are morally wrong. Surprisingly, though bourgeois morality is unconditionally condemnable under Marx's theory, the capitalist mode of production it supports is not. On Marx's analysis of the mode of production, significant alienation is an inherent feature of capitalism. Yet although he insists that existence under capitalism involves a miserable, alienated life for the vast majority, Marx also believes that the de-alienated communist world of the future would be impossible had many not suffered under capitalism. His doctrine of historical materialism supplies the second of these ethically-significant steps:

10. Necessarily, capitalism involves massive alienation.

11. But as historical materialism shows, capitalism is necessary for making the establishment of communist relations of production historically possible.

Given my claim that unalienated existence is the primary moral good in Marx's ethics, it would seem to follow that capitalism is, unconditionally, to be blamed. Yet a central claim of historical materialism is that the existence of capitalism is necessary for creating the material presuppositions of communism. Although Marx optimistically holds that communism inevitably will follow capitalism, his historical materialism also commits him to the extremely pessimistic conclusion that capitalism – and therefore extreme alienation – is a necessary *historical* precondition of communism.

That is, although the fundamental principles of Marx's ethical theory – steps one, two, and three – are eternally true, there is a sense in which many of the particular moral judgments made by his theory are historically-relative, viz. because his theory is extremely sensitive to historical possibility. Thus even though Marx insists that capitalism necessarily involves massive alienation, from step eleven it follows that capitalism is ethically necessary if a de-alienated communist era is to be ushered in. If, as I believe, the primary principle of Marx's ethical theory is that avoidable alienation is morally wrong, then it follows that the existence of capitalism is not always to be condemned. This is not to deny that in general Marx considers alienation a moral evil. Yet, depending on historical circumstances, sometimes the judgment that, e.g., 'the alienation inherent in capitalism is morally wrong' is over-ridden or trumped by the judgment 'the alienation inherent in capitalism is ethically necessary to bring about communism.' Specifically, given the claim from historical materialism in step eleven, it follows that capitalism is ethically necessary when and only when communism is not a historical possibility. Thus from step eleven,

11. Capitalism is necessary for making the establishment of communist relations of production historically possible,

these steps follow:

12. Therefore, from 4, 2, & 3, revolutionary action against capitalism is ethically wrong insofar as establishing communist relations of production is not a historical possibility.

13. But, from 4, 1, & 3, impeding revolutionary action against capitalism is ethically wrong whenever it is historically possible to establish communist relations of production.

On the ethical theory I've sketched, capitalism is morally repugnant for two reasons whenever communism is historically possible: firstly, because capitalism involves massive alienation and secondly, and perhaps more repulsively, because this alienation is unnecessary. Because they promote the mode of production which perpetuates this avoidable alienation, bourgeois morality and justice are utterly morally abhorrent as well. The case when communism is historically impossible – i.e., because the forces of production are not yet sufficiently advanced – is more complex. In this case, I've argued, the alienation is historically necessary for producing conditions under which alienated existence can be avoided. Suppose that it could be shown that life under capitalism was far more alienated than any previous mode of life. I claim that according to Marx's moral theory, the lexically-prior principle 'avoidable alienation is morally wrong' would entail that

even if it were possible, it would be morally wrong to abandon the capitalist mode of production for an earlier, less alienated mode. The reason is that Marx insists that all alienation will be unnecessary *once humanity has endured the capitalist historical epoch*. Since the ultimate evil in Marx's ethics is permitting unnecessary alienation to exist, it follows that he must consider it morally wrong to impede the historical process whereby communism develops through capitalism.

This allows us to better understand Marx's complex evaluation of bourgeois morality. Insofar as it would impede the historical transition to communism, the premature supersession of bourgeois morality by proletarian morality would be morally wrong. My suggestion is that although he believes it false and ideological, Marx might hold that bourgeois morality has a historically important function. Note that this is a practical, empirical claim about the effect of particular standards of morality being adopted at a particular time; to say that it might have a historically important function is not to make any theoretical claim about the theoretical validity of bourgeois morality. In particular, it is not tantamount to saying that it sometimes false on Marx's ethical theory that bourgeois ideology is false and morally repugnant: witness the derivation of step nine above. The way in which Marx seems to sometimes sanction, sometimes critique bourgeois morality has often been thought to pose a grave exegetical problem. As I shall argue in part III below, clearly distinguishing Marx's evaluation of the practical effects of a certain morality at a certain time from his theoretical statements about morality will go long way towards clearing up the appearance of mystery.[11]

More generally, I believe that failure to recognize the dialectical structure of Marx's ethical theory, and in particular the failure to understand why Marx is committed to steps twelve and thirteen, explains why his condemnation of capitalism can seem so confusing. In the next section I will attempt to flesh out and justify the interpretation of Marx just sketched.

VI

Any discussion of Marx's condemnation of capitalism which, like mine, emphasizes his hatred of alienation faces the onerous task of supplying an adequate

[11] This is easier said than done because, notoriously, Marx often wrote as a revolutionary and not as a theoretician. Given his views about the historical inevitability of communism, and of the functionality of different ideologies for promoting the historical development of communism, it is clear that Marx would believe that explicitly discussing his ethical theory would have no important (and perhaps even a retarding) revolutionary function.

Of course, the seeming lack of even an implicit theoretical backdrop to much of Marx's writings is another problem altogether. My claim in this essay is not that Marx explicitly held the ethical theory which I've described. Nor do I hold even that the evidence shows that Marx consistently could have held this theory–clearly, no theoretical reconstruction can square with all of Marx's (often vague, often polemical) texts. Rather, I believe that the ethical theory sketched in this essay is merely a viable interpretive possibility: it is consistent at least with the main currents of Marx's thought.

It does seem, however, that for the reasons adumbrated above the theory I discuss represents a particularly interesting interpretive possibility. Namely, my reading seems interesting because it promises to explain away much of the mystery and paradox which is commonly thought to surround issues of "Marxism and morality."

account of that troubling concept. Notoriously, Marx's notion of alienation is extremely complex. He speaks of alienated actions (e.g., alienated labor) as well as states of alienation (e.g., states of the worker caused by alienated labor). His discussion of the things which can become alienated seems to encompass individuals, groups, institutions, and, indeed, societies as a whole. Finally, Marx invokes many senses in which these things can be alien from each other and themselves; alienation, for Marx, can involve being alienated from the results of production, from the process of production, from the physical and social environment in which production occurs, from other people involved in production, and, crucially, alienation from one's own historically possible self-development (i.e., from one's creative potential as a producer and consumer of production).

The last form of alienation is the most important. For Marx all alienation seems to involve the self-estrangement of man in two senses, namely estrangement by his own actions (as a producer, consumer, member of a group, institution, society, etc.) which is also estrangement from the realization of historically-possible development and fulfillment of one's wants. Many of these aspects of alienation are brought into play in this condemnation of alienated labor from the *Economic and Philosophic Manuscripts*:

> In what does this alienation of labor consist? First, that the work is *external* to the worker, that it is not a part of his nature, that consequently he does not fulfill himself in his work but denies himself, has a feeling of misery, not of well-being, does not develop freely a physical and mental energy, but is physically exhausted and mentally debased. The worker therefore feels himself at home only during his leisure, whereas at work he feels homeless. His work is not voluntary but imposed, *forced labor*. It is not the satisfaction of a need, but only a *means* for satisfying other needs....The alien character of work for the worker appears in the fact that it is not his work but work for someone else, that in work he does not belong to himself but to another person....The activity of the worker is not his spontaneous activity. It is another's activity, and a loss of his own spontaneity. (Bottomore, pp. 177–178; Marx's emphasis)

I believe that, at bottom, Marx's moral condemnation of capitalism revolves around the way in which "the activity of the [alienated] worker is not his spontaneous activity... [but represents] a loss of his own spontaneity." For Marx this alienation has a number of important effects, which he sketches in this passage. More generally, Marx seems to this loss of spontaneity morally repugnant because his conception of the good life cannot be realized under such conditions of alienation. Marx's account of the good life, as well as its importance in his ethical theory, seems quite Aristotelian. Thus Marx and Aristotle both emphasize the importance of exercising the most creative human activities, especially those involving purpose and intelligence. Like Aristotle, Marx insists upon the paramount moral importance of the distinctively human ability purposively to plan one's life through choice; for both a life determined by forces outside one's control is not a

good life. Finally, it seems that although he would find Marx's historical optimism utopian, Aristotle would agree with the spirit of Marx's famous dictum that the best life of all would be in "an association in which the free development of each is the condition for the free development of all."

The interesting question of precisely to what extent Marx's and Aristotle's views converge and diverge is perhaps best left untouched. For the purposes of filling out my sketch of Marx's ethical theory it will suffice simply to list some areas of broad convergence. Helpfully, Richard Miller supplies just such a list. In "Marx and Aristotle" he cites these correspondences between Aristotle's conception of the good life and Marx's denunciation of capitalism:[12]

A1.	Since happiness is an activity, it requires a minimum of material goods and physical energy.
M1.	Capitalism degrades the proletariat by depriving them of the material means and the physical energies to exercise most of their capacities.
A2.	The good life must give priority to the exercise of the best human capacities, the ones which remove a person the farthest from an animal existence.
M2.	The proletariat are forced to spend most of their waking lives providing for the physical needs they share with animals, and most of the rest recovering through mere relaxation.
A3.	Intelligence, above all, separates men from animals. The exercise of intelligence is an especially important aspect of the good life.
M3.	Capitalism, while employing the most advanced technology and industrial organization, forces the proletariat to engage in monotonous, repetitive activity of unparalleled stupidity.
A4.	The molding of one's life through deliberation and choice is an important, characteristically human ability. So far as one's life is determined by forces beyond one control, it is not a good life.
M4.	Under capitalism, proletarians' lives are largely determined by forces beyond their own control.
A5.	The best life consists of activities engaged in for their own sakes.
M5.	Under capitalism proletarians sell their labor-power as a means to obtain mere necessities.
A6.	In a good society, people care for each other for their own sake, not for the sake of the goods that can be extracted from each other.
M6.	In capitalism labor-power is a commodity competing in the market and consequently the vast majority of people treat each other as means and are separated from each other through the egoistical relations of civil society.

[12] One could also add to each pair a corresponding statement about life under communism, where communism attains the Aristotelian virtue absent in capitalism. Miller lists these correspondences in Miller, 1989, pp. 177–180.

A7.	Pleasure, although not the sole good, is a good. It is the unimpeded exercise of a human faculty.
M7.	Part of the damage of capitalism is the infliction of pain on the proletariat and the impairment of their ability to function.
A8.	The pursuit of money for its own sake is unnatural and undesirable. It leads to the atrophy of the many human faculties which have non-acquisitive goals. Money should only be acquired in order to provide the means for the exercise of these capacities.
M8.	Capitalism encourages people to pursue money for its own sake. Under capitalism, the experience of life loses its diversity and individuality.

These correspondences are striking. In Miller's words, when one compares Marx and Aristotle "an amazing similarity emerges, uniting the great opponent of exploitation with the most celebrated defender of slavery" (1989:178). Now I will examine several passages where Marx seems to denounce capitalism on the grounds that in alienated capitalist life the vast majority of human beings have little hope of realizing any of the Aristotelian virtues listed above. I shall try to show that a reading stressing the affinities between Marx and Aristotle motivates the interpretation of Marx's ethical theory sketched in section V above.

Central to my interpretation is the claim that Marx's ethical theory, and in particular his critique of bourgeois morality as ideological, is compatible with his adopting what I've called moral realism. In an interesting passage in the *Communist Manifesto*, Marx and Engels respond to the charge that the Marxist critique of ideology is inconsistent with moral realism, i.e., the claim that there are non-relative moral truths. In this passage the interlocutor condemns Marxism for rejecting all morality:

> There are... eternal truths, such as Freedom, Justice, etc., that are common to all states of society. But Communism abolishes eternal truths, it abolishes all religion, all morality...it therefore acts in contradiction to all historical experience.

Marx and Engel's response, in effect, is to appeal to trans-historical moral judgments based on analyses of alienation:

> What does this accusation reduce itself to? The history of all past society has consisted in the development of class antagonisms, antagonisms that assumed different forms at different epochs. But whatever form they may have taken, one fact is common to all past ages, viz., the exploitation of one part of society by another. (McLellan, p. 236)

For Marx and Engels exploitation is a crucial notion for the analysis of alienation in a society; roughly, the more people who are exploited, and the more exploited they are, the more alienated is their society. It seems quite significant,

therefore, that Marx and Engels cite exploitation as the concept "common to all past ages" which shows that they do not mean to "abolish all morality." This coheres with my interpretation, on which alienation the central notion of Marx's ethical theory and his moral judgments are historically-relative insofar as they are sensitive to its "different forms at different epochs" and to the developing historical possibilities for its eradication. Moreover, my Aristotelian reading of Marx's ethical theory – e.g., its concern with the good life, and evaluation of institutions and practices by appeal to their consequences for its realization – helps explain why Marx's critique of bourgeois morality is not a critique on morality as such. As Marx and Engels suggest, their critique is of those non-Aristotelian moralities (utilitarian, deontological, etc.) which, by purporting to offer "universal" moral principles, are unacceptably insensitive to historical possibility.

Of course, my reading of Marx would be highly implausible if his celebrated critiques of private property, the hidden hand conception of the market, bourgeois freedom, etc. could not be understood in this Aristotelian light. In this passage from the *German Ideology* Marx and Engels argue that the alleged which proletarians allegedly exercise over their labor power is an illusion concealing the alienated character of life under capitalism. Significantly, they find bourgeois freedom bogus precisely because it is compatible with the debased state of the alienated proletariat, who are "subject to the power of things" alien to their own intentions and purposes:

> In theory, therefore, individuals appear to have greater freedom under the rule of the bourgeoisie than before; in reality of course they are less free, because they are more subject to the power of things... For the proletarians... the condition of their own lives, labor, and with it all the conditions of existence of modern society, have become something accidental, over which the individual proletarians have no control... The contradiction between the personality of the individual proletarian and the condition of life imposed on him, his labor, becomes evident to himself, for he is sacrificed from his youth onwards and has no opportunity of achieving within his own class the conditions which would place him in another class....The proletarians, if they are to achieve recognition as persons, will be obliged to abolish their own conditions of existence, which are at the same time those of [capitalist] society as a whole. (Bottomore, pp. 255–256)

Despite their formal freedoms as owners of their labor power, not to mention their protection by the canons of justice and assorted and sundry universal rights, the proletarians are debased because they are alienated. Thus "sacrificed from their youth onwards," the proletarians are "subject to the power of things" to such an extent that "the conditions of existence of modern society" afford them "no control" and, indeed, do not provide them with the possibility of "recognition as persons." Clearly, in this passage Marx morally condemns the capitalist mode of production because it affords a miserable, alienated existence to the great majority of its members.

Similarly, Marx and Engels denounce the bourgeois claim that the "hidden hand" of the market realizes some supposed common good by discussing how the market causes proletarians' lives to be alienated. In the *German Ideology* they write:

> Just because individuals seek only their particular interest, which for them does not coincide with their common interest (for the 'general good' is an illusory form of community life), the common interest is imposed as an interest 'alien' to them, and 'independent' of them....The social power, i.e. the multiplied productive force, which results from the cooperation of different individuals as it is determined by the division of labor, appears to these individuals, since their cooperation is not voluntary but natural, not as their own united power but as an alien force existing outside them, of whose origin and purpose they are ignorant, and which they therefore cannot control. (Bottomore, p. 228)

Presumably the sense in which the cooperation of the proletariat is 'natural' as opposed to 'voluntary' is that under the capitalist mode of production economic forces compel them to work under a certain division of labor, which ideologically appears inevitable to them. Instead of acting in their genuine class interests, the proletariat's "multiplied productive force" serves to promote an absurd "common interest [which] is imposed as an interest 'alien' to them, and 'independent' of them." But the fundamental problem is deeper than a lack of class consciousness; the situation described is one where *individuals* are deeply alienated from their creative capabilities, viz. because under the capitalist mode of production they do not select what they produce, when they produce it, how much they produce or may consume, or indeed which skills they will cultivate and develop as producers and consumers. In short, what is true for the class is true for the individuals comprising that class: the proletariats are alienated from their essential human powers, which appear to them as "an alien force existing outside them, of whose origin and purpose they are ignorant, and which they...cannot control."

In the light of his analysis of how it maintains alienated existence, it becomes clear why Marx is so unsympathetic to claims that the market efficiently satisfies at least some important needs. Consider his sarcastic portrayal in *Capital I* of capitalism's commodity market as "a very Eden of the innate rights of man:"

> The sphere that we are deserting, within those boundaries the sale and purchase of labor power goes on, is in fact a very Eden of the innate rights of man. There alone rule Freedom, Equality, Property, and Bentham. Freedom, because both buyer and seller of a commodity, say of labor power, are constrained only by their own free will. They contract as free agents, and the agreement they come to is but in the form in which they give legal expression to their common will. Equality, because each enters into relation with the other, as with a simple owner of commodities, and they exchange equivalent for equivalent. Property,

> because each disposes only of what is his own. And Bentham, because each looks only to himself. The only force that brings them together and puts them in relation with each other is the selfishness, the gain, and the private interests of each. Each looks to himself only, and no one troubles himself about the rest, and just because they do so, do they all, in accordance with the pre-established harmony of things, or under the auspices of an all-shrewd providence, work together to their mutual advantage, for the common weal and in the interest of all. (McLellan, p. 455)

Far from representing a concession to the "historically-relative justice" of capitalism, this passage summarizes what Marx considers the most contemptible features of capitalism: illusions to the contrary, there is no pre-established harmony or benevolent guidance shaping the market, and the alleged freedom, equality, and justice provided by fair competition are merely handy pretexts for the bourgeoisie to exploit and alienate the proletariat. Although in a sense Marx here admits that under bourgeois standards capitalist transactions are wholly moral and just, clearly he should be understood as insisting that these standards are false and ideological. For Marx the alleged freedom, equality, and utilitarian efficiency of the market are wholly illusory; worse, as ideological illusions these ideas are functional for maintaining a brutal, literally dehumanizing state of alienation.

The contrast between the "common interest" of a class and the ideological illusory "general good" suggested by the last two passages is noteworthy. Specifically, I think that it helps to explain an argument for the impossibility of constructing a (moral) theory of distributive justice found in the *Critique of the Gotha Programme*. According to Marx, because their political demands "contain nothing beyond the old democratic litany familiar to all: universal suffrage, direct legislation, popular rights, etc." the follower of Lassalle fail to see through the illusions of bourgeois morality (McLellan, p. 565). Marx is particularly harsh on the call for "a fair distribution of the proceeds of labor."

After criticizing the notion of 'proceeds of labor' as irredeemably vague, Marx attacks the notion of a just distribution of social goods. He begins by reminding his audience that from the point of view of bourgeois ideology, the capitalist distribution seems wholly just:

> What is a 'fair distribution'? Do not the bourgeoisie assert that the present-day distribution is 'fair'? And is it not, in fact, the only 'fair' distribution on the basis of the present-day mode of production? (McLellan, p. 566)

Marx is clearly not arguing that distribution under the capitalist mode of production is morally unimpeachable; however it is meant to be answered, the correct response to the first rhetorical question is not 'a fair distribution in capitalism is whatever the bourgeoisie say it is.' Rather, by the second question Marx emphasizes that the bourgeoisie rationalize away the immorality of the distribution by appealing to their ideological conception of justice. Finally, Marx's third question

highlights the fact that this immoral – 'fair' to the bourgeoisie – distribution is essential to the capitalist mode of production.

I said that Marx's argument in the *Critique of the Gotha Programme* is an argument against *any* proposed principles of just distribution, bourgeois or not. The text continues with Marx's rhetorical question, "Have not also the socialist sectarians have the most varied notions about 'fair' distribution?" (*ibidem*). Marx's description of the Lassalleans' proposed standards of distributive justice under communism prepares the way for his argument against all theories of distributive justice. He writes:

> To understand what is implied in this connection by the phrase 'fair distribution'... [we see that the Lassalleans hold that] 'the proceeds of labor belong undiminished with equal rights to all members of society'. 'To all members of society'? To those who do not work as well? What remains then of the 'undiminished proceeds of labor'? Only to those members who work? What remains then of the 'equal right' of all members of society? But 'all members of society' and 'equal right' are obviously mere phrases... (McLellan, pp. 566–567)

In Marx's eyes the Lassallean notions of 'undiminished proceeds' and 'equal rights' to those proceeds are, like their notion of 'proceeds of labor', hopelessly vague. But he also makes a more radical claim here, viz. that "'all members of society' and 'equal right' are obviously mere phrases." On the one hand, I think that Marx's radical claim is that it is in principle impossible for any standard of justice to embody a conception of 'all members of society' which captures all the morally relevant facts about those members. On the other (not unrelated) hand, Marx also claims that under no principles of distributive justice can there be a tenable conception of the way in which those principles guarantee people an "equal right" to the social goods being distributed.

I believe that Marx makes these claims largely for Aristotelian reasons. He continues by arguing that the Lassallean account of justice – which he identifies with the lesser stage of communism – doesn't really abandon the ideological bourgeois conception of "equal justice for all:"

> ... Only now do we come to the 'distribution' which the programme, under Lassallean influence, alone has in view in its narrow fashion, namely, to that part of the means of consumption which is divided among the individual producers of the cooperative society.
> ... What we have to deal with here is a communist society, not as it has developed on its own foundations, but, on the contrary, just as it emerges from capitalist society; which is thus in every respect, economically, morally, and intellectually still stamped with the birth marks of the old society from whose womb it emerges. Accordingly, the individual producer received back from society – after the [socially necessary] deductions have been made – exactly what he gives to it... He receives a certificate from society that he has furnished such and such

> an amount of labor (after deducting his labor for the common funds), and with this certificate he draws from the social stock of means of consumption as much as costs the same amount of labor. The same amount of labor which he has given to society in one form he receives back in another.
> Here obviously the same principle prevails as that which regulates the exchange of commodities [in capitalism]....the same principle prevails as in the exchange of commodity-equivalents: a given amount of labor in one form is exchanged for an equal amount of labor in another form. Hence, equal right here is still...bourgeois right.
> ... This equal right is still constantly stigmatized by a bourgeois limitation. The right of producers is proportional to the labor they supply; the equality consists in the fact that measurement is made with an equal standard, labor. (McLellan, pp. 567–568)

If Marx were only concerned with condemning the poverty of the proletariat under capitalism or with alleviating the most extreme forms of economic exploitation, one would expect that he would consider this communistic distribution morally acceptable. Yet he finds that in the society described, "equal right here is still bourgeois right." In fact, I believe that for Aristotelian reasons Marx finds the very notion of "equal rights" inherently immoral because mired in more subtle forms of alienation.

The heart of Marx's argument is the claim that no conception of 'equal right' to social goods is tenable because necessarily, under any such conception "the equality consists in the fact that measurement is made with an equal standard." To Marx this kind of equality is morally unacceptable because the use of such an equal standard contradicts his Aristotelian conception of human virtue. As Marx puts it:

> Right by its very nature can consist only in the application of an equal standard; but unequal individuals (and they would not be different individuals if they were not unequal) are measurable only by an equal standard in so far as they are brought under an equal point of view, are taken from one definite side only, for instance, in the present case, are regarded only as workers and nothing more is seen in them, everything else being ignored....To avoid all these defects, right instead of being equal would have to be unequal. (McLellan, p. 569)

The parenthetical comment is sometimes thought to show that Marx's argument relies on Leibniz's dubious principle of the identity of indiscernibles. This interpretation, I believe, is wrongheaded: Marx finds it wrong to bring unequal individuals "under an equal point of view" for wholly Aristotelian considerations about the moral uniqueness of individuals. In the passage Marx rejects all attempts to conceive of people as equal under any single description; so too would he reject "equal measure" by any *set* of morally-relevant criteria. For Marx's point is that no abstract principles of distributive justice can guarantee that people will not

suffer unnecessary alienation. On his conception, individuals are radically unequal because the conditions under which each can attain the Aristotelian virtues of a good life are unique. It follows, therefore, that individuals cannot be treated as "equal in principle" without violating the primary – and lexically prior – principle of his ethical theory.

Marx's condemnation of the 'general good' follows from this point. What he condemns as illusory is any supposed general good which is good for all people at all times. The moral good for Marx is de-alienated existence, which is both historically relative (viz., because the type of de-alienated existence to which a person can aspire at any time is a matter of historical possibility) and radically unique to each individual (because as de-alienated every individual consciously molds his own life). It follows that Marx cannot countenance conceptions of either trans-historical or historically-relative justice; for Marx all appeals to justice are ideological and immoral.

Thus we come to Marx's famous description of the "higher phase" of communism, where concerns of justice are simply irrelevant to de-alienated existence:

> In a higher phase of communist society, after the enslaving subordination of the individual to the division of labor, and therewith also the antithesis between mental and physical labor, has vanished; after labor has become not only a means of life but life's prime want; after the productive forces have also increased with the all-round development of the individual, and all the springs of co-operative wealth flow more abundantly – only then can the narrow horizon of bourgeois right be crossed in its entirety and society inscribe on its banners: from each according to his ability, to each according to his needs! (*Ibidem*)

Although crossing "the narrow horizons of bourgeois right" involves abandoning notions of distributive justice, Marx need not hold that this transition leaves behind all morality. I have argued that an Aristotelian reading of Marx's ethical theory is quite powerful, particularly because it shows that the famous charges of paradox are better off forgotten.

VII

If Marx's conception of justice – and his ethical condemnation of capitalism – does not focus on distribution of wealth, on what does it rest? I submit that, for Marx, as for Aristotle, justice has the most to do with the conditions that make it possible for an individual to realize potential and freedom within and through society; for both, I believe, the goal of justice is a society that fully develops its citizens, educates them, and raises them to reflective self-consciousness. If I am right, the key issues for understanding Marx's condemnation of capitalism center on the nature of social relationships within capitalist societies, the development of character within them, their formation of self-consciousness, and the possibilities of human potential they support. Thus, for example, these questions rise to the fore throughout Marx's Critique of the Gotha Program just as they do, in Books I and IX of Aristotle's *Politics* and in Book V of his *Nichomachean Ethics*.

Moreover, and as I have shown in section VI above, both thinkers explicitly connect their accounts of development and actualization of human potential and freedom within a society with a sharp critique of economic forms that block or retard that development and actualization. To reflect more deeply on this, consider George McCarthy's discussion of this theme in Aristotle. After noting of Aristotle that "anything that tends to undermine the harmonious social and political relations of the polis is criticized by him as destructive to the possibilities of human self-realization" (71), he glosses the first chapter of *Politics* Book IX as follows:

> With the introduction of commercial trade for profit – that is, commodity exchange – the art of household management turns into an unnatural process characterized by the search for wealth and profits without limit. In fact, the latter becomes the measure of the good life and the criterion by which happiness is calculated. This quest for money without limit is an example of living, but not living well; it is a form of pure existence without moral excellence, for it undermines the very foundations for the possibility of a virtuous and rational life in society....[For Aristotle], the economy has a crucial role to play in creation of a just society, for without it there can be no developed society at all. Exchange is a prerequisite for the economic survival of complex societies, the integration of their diverse needs, and the maintenance of a fair and just economic system. Opposed to *kapelike* we have economic reciprocity (*antipeponthos*), where need, equality, and mutual sharing maintain a just economic exchange for the common good. Here exchange is not viewed as a contest between individuals for private happiness and profit, but the economic basis for public happiness and social justice. Reciprocity and mutual sharing (*metadosis*) become the economic foundations of the polis and the education of the citizen. (73–74; also see Aristotle, *Politics*, 1258a 37–1258b 3).

To be sure, Aristotle's conception of justice is not Marx's, and their accounts of which members of society can benefit from a state of justice differ vastly (remember that Aristotle was the great defender of human slavery as necessary means for securing those benefits for some!), as do the particular details of exactly how just social relations constitute necessary conditions for promoting the full development of human capacities and freedom, and what social relations are necessary for justice. The crucial point is that Marx's critique of political economy can justly be seen as a form or return to Aristotle's conception of a broad account of justice that expands beyond issues of distributive justice in a way that integrates a positive philosophical anthropology that Miller's list of correspondences discussed above can be taken to articulate. To summarize via a slogan that McCarthy develops, "Like Aristotle, Marx believes that ethics is a form of political and economic knowledge" (111).

Recall Marx's clarion call from the German Ideology, discussed above, that "Only in the community has each individual the means of cultivating his gifts in all directions." For Marx, we have seen, the realization of human potential

depends on the organization and structure of the social relations of production. Despite the many differences of doctrine, and despite the many differences of perspective of an author writing so many centuries later, the heart of Marx's ethical condemnation of capitalism is Aristotelian: for Marx, as for Aristotle, rationality and morality represent full expressions of the humans' social nature, and for Marx capitalism is evil precisely because within capitalism individuals are alienated from the social forms on which rest the possibility of their freedom, their self-actualization, and their self-conscious existence. Although Marx's perspective incorporates philosophical, economic, and political developments of his time, at bottom his ethical critique of Enlightenment individualism rests on the teleological concept of a "social individual" developed within the Ethics and Politics of Aristotle.

Alan Gilbert makes a similar point by conceiving of Marx's communism as a modification of Aristotelian eudaimonism, which he defines broadly as "a theory that evaluates activities and states of characters by asking how they advance individual happiness *and affect the quality of human lives*" (Gilbert, 1971, p. 192; emphasis added). On Gilbert's view, Marx, like Aristotle,

> [O]ffered an objective interpretation of human happiness in which certain activities, states of character, and effects of fortune make possible a full and happy life. While Marx differed with Aristotle on many specific issues (the alleged incapacity for happiness among barbarians, or women, the primacy of contemplative activity), he also emphasized objective features of individual lives which would profoundly influence happiness (display of social connectedness in genuine forms of friendship and political community, freedom from exploitation, freedom from physical handicap, possession of the opportunity to realize one's capacities, perhaps realization of some of one's highest capacities)... in some fundamental respects, therefore, Marx could have regarded his argument as a correction and refinement of Aristotle's eudaimonism" (*ibidem*).

Allen Wood similarly sees a dialectic in Marx between individuals, their cultures, and their political and economic communities. While discussing the similarity of Marx's account of alienation to Aristotle's account of self-actualization, Wood reads Marx's central concept of alienation as one where people are unable to actualize themselves and to develop and exercise the powers that human beings can potentially exert. Thus he concludes:

> There are some reasons for thinking that the possibility of alienation is closely related to the essentially human trait of self-consciousness. To experience oneself or one's life as worthless or worthwhile, as meaningless or meaningful, seems to presuppose some conception of what is felt to be worthwhile or worthless. For this reason, only a being who has some sort of self-conception seems to be capable of either an alienated or a fulfilled life. Further, the possibility of alienation, at least

> for Marx, is closely bound up with the human trait of species consciousness. Marx often speaks of alienated life as one in which human beings fail to 'affirm' (*bejahen*), 'confirm' (*bestätigen*) or 'actualize' (*verwirklichen*) themselves. A human life which is self-affirming, self-confirming and self-actualizing is a meaningful life; a self which affirms, confirms and actualizes itself is a self which has worth, and recognizes the worth it has. But Marx also indicates that to affirm, confirm and actualize oneself is to affirm, confirm and actualize one's essence, that is, the human species-essence. The measure of this self-actualization, of an individual's satisfaction of a 'natural vocation' (*natürliche Bestimmung*), is 'the extent to which the human being as species being, as a human being, has become himself and grasped himself'. Alienation is thus conceived by Marx as a separation and estrangement of individuals from their human essence. Their 'being does not correspond to their essence', is not 'in harmony' with it; their lives are not lives in which 'the human essence feels itself satisfied'. (2004:21)

I have argued that Wood, Gilbert, McCarthy, and Miller are right to stress correspondences between Aristotle and Marx, although I have also provided an epistemological argument that shows that Miller and Wood are wrong to conclude that Marx's condemnation of capitalism is paradoxical.

I end this paper by noting that the failure of the two global criticisms of Marxism that I have addressed opens significant interpretive possibilities that readers of Marx in our own century should not ignore: if, on the one hand, the broad criticisms that seem to make concern for Marx's texts and arguments at best a pedantic irrelevance and, on the other hand, Marx's condemnation of capitalism promises to develop Aristotelian ideas while taking into account economic, political, and technological developments which Aristotle could not conceive, my optimistic conclusion there is much alive left in Marx's texts to benefit us as we confront the troubled and turbulent currents of our own time.

BIBLIOGRAPHY

Ashcraft, R., 'Marx and Political Theory', *Comparative Studies in Society and History* 26 [4] (1984), pp. 637–671.

Berki, R.N., 'On the Nature and Origins of Marx's Concept of Labor', *Political Theory* 7 [1] (1979), pp. 35–56.

Berlin, I., *Karl Marx*, 4edn, Oxford 1978.

Brenkert, G.G., 'Freedom and private property in Marx', *Philosophy and Public Affairs* 8 [2] (1979), pp. 122–147.

Buchanan, A.E., 'Marx, morality, and history: An assessment of recent analytical work on Marx', *Ethics* 98 [1] (1987), pp. 104–136.

Callinicos, A., *Marxist theory*, Oxford 1989.

Cohen, G.A., 'Karl Marx and the Withering Away of Social Science', *Philosophy and Public Affairs* 1 [2] (1972), pp. 182–203.

Cohen, G.A., *History, labor, and freedom: themes from Marx*, Oxford 1988.

Cohen, G.A., *Karl Marx's theory of history: a defence*, Oxford 1978.

Cohen. M., T. Nagel, and T. Scanlon, *Marx , justice, and history*, Princeton 1980.

Elster, J., *An introduction to Karl Marx*, Cambridge 1986.

Elster, J., *Making sense of Marx*, Cambridge 1985.

Finley, M.I., 'Aristotle and economic analysis', *Past and Present* 47 [47] (1970), pp. 3–25.

Fracchia, J., 'Marx's Aufhebung of Philosophy and the Foundations of a Materialist Science of History', *History and Theory* 30 [2] (1991), pp. 153–179.

Gilbert, A., 'An Ambiguity in Marx's and Engels's Account of Justice and Equality', *American Political Science Review* 76 [2] (1982), pp. 328–346.

Gilbert, A., 'Historical Theory and the Structure of Moral Argument in', *Political Theory* 9 [2] (1981), pp. 173–205.

Gilbert, A., 'Review of Marx's Theory of History by William H. Shaw', *Philosophical Review* 88 [3] (1979), pp. 476–480.

Kolakowski, L., *Main currents of Marxism, Volume 1: The Founders*, Oxford 1981.

Le Baron, B., 'Marx on Human Emancipation', *Canadian Journal of Political Science* 4 [4] (1971), pp. 559–570.

Lukes, S. *Marxism and Morality*, Oxford 1985.

Marx, K. and D. McLellan, *Karl Marx: selected writings*, New York 1977.

Marx, K. and F Engels, *Selected correspondence, 1846–1895*, Donna Torr, 3 edn, Moscow 1975.

Marx, K., T.B. Bottomore, and M. Rubel, *Selected writings in sociology & social philosophy*, New York 1964.

McCarthy, G.E., *Marx and the Ancients: Classical Ethics, Social Justice, and Nineteenth-Century Political Economy*, Savage 1990.

Meikle, S., 'Aristotle on Business', *The Classical Review. New Series* 46 [1] (1996), 138–151.

Meikle, S., 'Aristotle on equality and market exchange', *Journal of Hellenic Studies* 111 (1991), pp. 193–196.

Meikle, S., 'Aristotle on Money', *Phronesis* 39 [1] (1993), pp. 26–44.

Miller, D., and L. Siedentop, *The nature of political theory*, Oxford 1983.

Miller, R.W., 'Marx and Aristotle: A Kind of Consequentialism', *Canadian Journal of Philosophy*, suppl. vol. 7 (1981), pp. 323–352.

Nasser, A.G., 'Marx's Ethical Anthropology', *Philosophy and Phenomenological Research* 35 [4] (1975), pp. 484–500.

Pack, S.J., 'Aristotle's Difficult Relationship With Modern Economic Theory', *Foundations of Science* 13 [3-4] (2008), pp. 265–280.

Roemer, J.E., *Analytical Marxism*, Cambridge 1986.

Ruderman, R.S., 'Aristotle and the Recovery of Political Judgment', *American Political Science Review* 91 [2] (1997), pp. 409–420.

Schwartz, N.L., 'Distinction between Public and Private Life: Marx on the *Zoon Politikon*', *Political Theory* 7 [2] (1979), pp. 245-266.

Smith, T.W., 'Aristotle on the Conditions for and Limits of the Common Good', *American Political Science Review* 93 [3] (1999), pp. 625–636.

Springborg, P., 'Karl Marx on Democracy, Participation, Voting, and Equality', *Political Theory* 12 [4] (1984), pp. 537–556.

Springborg, P., 'Politics, Primordialism, and Orientalism: Marx, Aristotle, and the Myth of the Gemeinschaft', *American Political Science Review* 80 [1] (1986), pp. 185–211.

Wallach, J.R., 'Contemporary Aristotelianism', *Political Theory* 20 [4] (1992), pp. 613–641.

Wood, A.W., *Karl Marx*, New York 1981.

Wood, A.W., *Karl Marx*, 2 edn, New York 2004.

AESTHETICS

S t u d i a P h i l o s o p h i c a
W r a t i s l a v i e n s i a
Supplementary Volume, English Edition 2013

ADAM CHMIELEWSKI
University of Wrocław

Political Aesthetics and the Urban Change

Abstract

The purpose of this paper is to sketch out a conception of urban political aesthetics aimed to demonstrate a continuing relevance of philosophy to understanding the city. I shall attempt to show in particular that the political aesthetics outlined here may shed some new light on the problem of urban aesthetics of absence, the concept introduced to urban studies by Richard Shusterman. I argue that the tendency of individuals to withdraw from the present urban spaces, constituting a form of public agoraphobia, is encouraged by the processes of commodification which fuel rapid transformations of the city. I also stress that increasing levels of sophistication, required of individuals in order to participate in contemporary urban life, generates a phenomenon of interpassivity which adversely affects the civic agency of urban citizens.

Philosophy and the City

Plato's *Phaedrus* is his only dialogue set outside the city walls. Socrates is taken for a walk to the countryside by young Phaedrus who hopes to discuss the matters of love with him. On their way Phaedrus notices that Socrates, though knowledgeable about mythical events which are said to have happened in various locations around the city of Athens, is rather unfamiliar with its actual vicinity. He comments: "anyone would take you [...] for a stranger being shown the country [...] never leaving town to cross the frontier nor even, I believe, so much as setting foot outside the walls."[1] Begging for forgiveness, Socrates famously responds: "I am a lover of learning, and the trees and open country won't teach me anything, whereas men in the town do,"[2] adding that the reason he agreed to a countryside walk instead of strolling the city streets is the wisdom he hopes to learn from the speeches which Phaedrus promised to present to him. This passage signifies the humanistic turn in philosophy, which in fact had been effected earlier through the Sophistic opposition between *physis* and *nomos*, the laws of nature and the

[1] Plato, *Phaedrus*, 230c, transl. R. Hackforth, [in:] *The Collected Dialogues of Plato*, E. Hamilton, H. Cairns, Princeton 1985.
[2] *Ibidem*, 230d.

laws established by human communities. In this Platonic dialogue the opposition between natural and civilized forms of life is employed in order to endorse the values which may only be cultivated in a city.

As a matter of historical fact, philosophy rarely left the city, also in the sense that it rarely left the city outside the scope of its interest. Plato devoted himself to designing a perfect city-state, Aristotle saw in the city the highest form of political organization.[3] Protagoras, a leading figure of the pragmatically oriented Sophistic movement, drafted a constitution for the city of Thurii. The significance of philosophy for understanding the city may also be seen from the vast philosophical utopian literature aimed at formulating idealized visions of the perfect and just urban life.[4] Philosophers have also expressed negative opinions about the city; a critical approach towards the existing forms of urbanism, formulated by Rousseau, has inspired numerous dystopian visions of urban life and the establishment of an outspoken movement of the intellectual and political anti-urbanism.[5]

The general aim of this paper is to outline a conception of urban political aesthetics which demonstrates a continuing relevance of philosophy to understanding the city. More specifically, I shall attempt to show that the political aesthetics in question may shed some new light upon the problem of urban aesthetics of absence, i.e. a theme developed by the pragmatist philosopher Richard Shusterman[6] in his criticism of Georg Simmel's seminal paper on mental life in the metropolis,[7] and inspired by Richard Sennet's[8] urban studies, as well as Walter Benjamin's[9] aesthetic writings. In particular, I shall argue that mental and political withdrawal of individuals from the active participation in urban life, which I interpret as a form of the public agoraphobia, is engendered, among others, by the processes of commodification which encroach upon ever new spheres of human life, and by the phenomenon of interpassivity, which increasingly supplants the traditional, direct relationships among individuals and encourages vicarious modes of relationships between them. I also argue that these processes and phenomena are both results and causes of rapid transformations of contemporary city, while these transformations find expressions in, and are identified by, aesthetic characteristics specific to them.

The Urban Dynamics

As a form of organization of social life and satisfaction of human needs, the city consistently demonstrates its remarkable effectiveness and resilience. Despite

[3] Aristotle, *Politics*, 1252a 1–7.

[4] T. More, *Utopia*, 1516; T. Campanella, *The City of the Sun*, 1623; F. Bacon, *New Atlantis*, 1627.

[5] J.-J. Rousseau, *Emile, or on Education,* [1762] New York 1979.

[6] R. Shusterman, 'Urban Scenes and Unseens', *Filozofski Vestnik* 17 [2] (1996), pp. 171–179; also, in an expanded form: 'The Urban Aesthetics of Absence', *New Literary History* 28 (1997), pp. 739–755.

[7] G. Simmel, *The Metropolis and Mental Life*, [in:] *The Sociology of Georg Simmel*, transl. K. Wolff, New York 1950, pp. 409–424.

[8] R. Sennett, *The Conscience of the Eye: The Design* and *Social Life of Cities,* New York 1990.

[9] W. Benjamin, *On Some Motifs in Baudelaire*, [in:] W. Benjamin, *Illuminations*, New York 1988.

many negative aspects of urban life, cities remain compellingly attractive as a form of human cooperation. This is evidenced by the rapidity of the urbanization processes which have been particularly intense during the past two centuries. In 1800 only two cities in the world were inhabited by one million of people or more; in 1900 there were 17 such cities; in 1950 – 86; whereas in 2000 – 387. In 1800, the average size of the largest one hundred cities was 187 thousand people; in 1900 – 725 thousand; in 1950 – 2,2 million; whereas in 2000 it reached 6,3 million people.[10] In 1900, only 13 per cent of the world population lived in the cities; over the 20th century the number of cities-dwellers increased tenfold: in 2007 the urban population has exceeded the population living in rural areas. The urban population now exceeds 3 billion and is equal to the whole world population in 1960.

The global urbanization process is an effect of the concentration of the production of goods and services in urban centers which turns cities into economic powers on their own.[11] To illustrate this point one should observe that in 20th century the GNP produced in the cities exceeded the GNP of the rural territories; the volume of the urban GNP accounts now for 80 per cent of the global product, while in the developed countries it amounts to 85 per cent of GNP. The forecasts for the decades to come suggest that the urban dynamics will be even more intensive.[12]

In most general terms the urban success may be explained by cities' ability to generate knowledge and resources necessary for the satisfaction of human needs, as well as creation of new ones. The growth of knowledge in urban centers has helped them to develop in innovative ways which have been variously exemplified throughout millennia of urban history, most overwhelmingly in the industrialization of cities in 19th and 20th centuries, as well as in the current processes of their de-industrialization, associated with the dematerialization of production. Cities as economic, cultural and intellectual centers have now become the genuine agents of globalization, some of them becoming megacities or global cities.[13] Globalization has resulted in a decline of the political, economic and cultural role of the states and in an increase of the significance of cities. Technological development in transportation and related urban infrastructure have enabled the mobility of populations on a mass scale. Though motivated mainly by economic reasons, the mobility is also driven by aspiration for attractive leisure, now available to ever increasing masses of people. All forms of human mobility carry with themselves a great economic potential, but also grave and intricate problems in managing the fast growing urban human traffic, whereas attempts to tap the potential, and to solve those problems, deeply affect the ongoing transformation of contemporary cities.

[10] D. Satterthwaite, *The Scale of Urban Change Worldwide 1950–2000*, London 2006.
[11] J. Jacobs, *The Economy of Cities*, New York 1969.
[12] *Urban World*, MacKinsey Report 2011.
[13] S. Sassen, *Cities in a Word Economy*, London-Delhi 2000; S. Saskia, *Global Networks, Linked Cities*, London 2002; S. Sassen, *The Global City: New York, London and Tokyo*, Princeton 2000.

The City as a Work of Art

In the 19th century cities became centers of industrial capitalist production. The accompanying urbanist, demographic and cultural transformations have become a subject of sociology which has to a large extent superseded philosophy in its attempts to understand the city. The latest developments in the character of capitalist production, and in the nature of the goods produced, have tremendously affected the development of cities and transformed them from the industrial into post-industrial centers. This transformation has re-directed the attention of urban scholars toward the role of knowledge, culture and arts, i.e. symbolic goods, in the functioning and development of cities, as well as toward the rise of creative classes responsible for the creation of symbolic capital. It has also led to the establishment of urban studies as a discipline in its own right, which evolved from interdisciplinary research into the intense contemporary urbanization processes. However, even if the present-day urban scholars focus on cities as the centers of production, exchange and consumption of symbolic capital and cultural goods, they tend to perceive these goods in an instrumental way, i.e. as means of solving practical problems generated by the very form of urban life.[14] Despite their focus upon practical aspects of urban life, urban studies are unavoidably underpinned by philosophical and anthropological visions of individual and social life; and since human beings are inherently moral, aesthetic and artistic creatures,[15] ethical and aesthetic preconceptions constitute a natural and integral part of these visions. More often than not, however, within urban studies those preconceptions are taken for granted rather than critically appraised and debated. Since a critical discussion of these assumptions and preconception is a truly philosophical task, it seems to validate philosophy's bid for reclaiming its original role in understanding the contemporary urban life.

Another argument for the import of the philosophical approach in understanding the city is based upon a rather uncontroversial observation that the city itself and each aspect of urban life are cultural formations. The city is a complex embodiment of human excellence in various arts: planning, designing, engineering, construction, but also in the political arts of managing the social and moral life. In this sense the city is a work of art.[16] At the same time the city was the birthplace of arts and of their growth. Moreover arts, in an almost self-referential gesture, often turn their attention to the city itself in order to reflect upon it, depict it, and explain it in their own poetic manner. This dialectic relationship of the city

[14] *Cf*. e.g. D. Ilczuk, *Ekonomika kultury*, Warszawa 2012.

[15] E. Dissanayake, *Homo Aestheticus*, Washington 1995; D. Dutton, *Art Instinct*, New York 2009.

[16] Jane Jacobs claims that "there is a basic aesthetic limitation on what can be done with cities: a city cannot be a work of art" (*The Death and Life of Great American Cities: The Failure of Town Planning,* Harmondsworth 1965, p. 386, after: H. Smyth, *Marketing the City. The Role of Flagship Developments in Urban Regeneration*, London 1994, p. 225). Despite that I believe that cities may be considered as works of arts in the sense delineated above. This is because Jacobs's claim is based upon a unduly sharp distinction between "arts" and "life". Jacobs remark remains valid, however, in relation to attempts to endow cities with the status of the works of art in the narrow sense of this concept, especially by means of their spectacularisation, referred to below.

and arts as means of understanding, description and formation of the city, should be perceived as an essential characteristics of cities, for the city cannot be fully understood without reference to arts, but also arts cannot be understood without reference to the city, the milieu which made them possible.

In other words, cities are sites of production, distribution and consumption not only of material goods, but also of symbolic ones. Various symbolic values and goods play a number of instrumental roles in the city; but they are also endowed with autotelic status and as such cannot be reduced to their instrumental usefulness only. For they play not only the role of external values within urban practices, but also the role of internal ones, and as such they must be perceived as constitutive of these practices.[17] The city, as a specific cultural formation, is thus a place of creation, accumulation, transmission, distribution and consumption of cultural goods and symbolic values. I believe that these phenomena and processes may be understood most effectively from the perspective of urban political aesthetics, a philosophical discipline whose task is to explicate them and demonstrate their constitutive role for the practices of contemporary urban life. Such an approach seems especially justified in view of the intense growth of cities and their present role not only as economic and political, but also leading cultural agents.

The City as a Space

The predominance of the sense of sight in human cognition, memorably remarked upon by Aristotle,[18] indelibly and crucially affects the development of all human cultures. It thus seems that the proper point of departure for urban political aesthetics should be the fact that human cultures develop in a way which reflects the fundamental role played in human life by the perceptual, visual, and spectacular. In order to make this point one may take advantage of the Berkeleyan principle *esse est percipi* which served to express his epistemological subjectivism. A reinterpretation of this principle may be employed to explain contemporary cities as dominated by the culture of visibility. Many forms of urban developments, achievements and human behavior are performed in order to be noticed by others. Vying for attention is, on the one hand, among chief reasons for emergence of a specific urban human way of life, but also, on the other, turns cities into exposition places and sites of ceaseless rivalry for attention. One may thus say that in order to be in the urban spaces one must be perceived within them. Accordingly, urban spaces may be understood as areas of agonistic rivalry for recognition in the Hegelian sense, which is entered into by individual and collective subjects alike. From this point of view, cities as cultural formations are manifestations of human attempts to create and recreate spaces they inhabit in accordance with their desire for being seen which is satisfied in accordance with, or – ever more often today – in violation of, various and continuously changing customs and requirements of their cultures, determining the temporarily acceptable modes of public conduct.

The city is then a unique space which itself consists of many sub-spaces, created

[17] *Cf*. A. MacIntyre, 'After Virtue and Marxism: A Response to Wartofsky', *Inquiry* 27 [2–3] (1984), pp. 253–254.

[18] Aristotle, *Metaphysics*, 980a.

and recreated by people in order to live their life in them. Spaces of human life should be understood as products of historically changing social practices and as *loci* of continuous transformation of human life. The spatial nature of the city implies, in an yet another way, that if the city is to be understood, it needs to be approached from an aesthetic approach point of view. Thus the categories of space and its production, the conditions of its transformations, and the subjects taking part in, and affected by, these processes, constitute basic concepts of political aesthetics.[19]

The philosophical-anthropological understanding of space departs from common intuitions about it, which inform the view of the public space as an agora in which individuals gather in order to develop and exercise their agency. Through an analogy to contemporary physics, in which space exists in so far as there are particles which fill it with their kinetic energy, one should rather understand spaces of human life as existing in so far as there are social particles which fill them with their political, moral and aesthetic energy. The spaces are constituted by, and change according to, the energy generated by the individuals and communities moving in them.[20]

Spaces of human life, while they incessantly intersect and mutually blur each other's borderlines, are nevertheless distinct from each other: social aspects of human life are not identical with the public ones, and both are distinguishable from the private and intimate spheres. The intimate space is part of the private sphere, but cannot be fully identified with it for it is created by needs which cannot be satisfied by activities in any of the remaining spaces. Individuals function in these spaces as formed by nature and culture. They enter them as individuals, and as members of communities. The evolution of forms of social life has generated norms which regulate the conduct of individuals within each of these spaces, as well as at their intersections. Each of these spheres has its own history, and is governed by its own rules which require compliance with them. Actions performes in each of them have moral and political consequences, and are subject to an aesthetic assessment, for aesthetic categories constitute an integral part of human orientation in the spaces of human life.

Cities are dynamic social, political and cultural formations. Their dynamics is a result of deep-seated conflicts within them, which are contained, with varying success, by continuous attempts to impose some order upon them, while forms and shapes of those imposed orderings are in themselves an object of agonistic urban rivalry. This necessarily means that functioning in urban spaces involves a sustained effort; to be in them is to be prepared for a struggle for a place in them. This feature of urban spaces may be illustrated by saying that a place once assumed by an individual in such spaces cannot be claimed as her permanent property, for

[19] S. Langer, *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, 1942; H. Lefebvre, *The Production of Space*, Oxford 1991; *Writings on Cities*, Oxford 1996; *The Urban Revolution*, London 2003. See also ŁStanek, *Henri Lefebvre on Space: Architecture, Urban Research, and the Production of Theory*, Minneapolis 2011, and H. David, *Rebel Cities. From the Right to the City to the Urban Revolution*, London 2012.

[20] *Cf*. A. Chmielewski, *The Gaze and Touch in the Public Space. Toward the Political Aesthetics*, [in:] *Wrocław Non Stop: Urban Space*, Wrocław 2008, pp. 84–92.

while upon returning to it, she will often find it occupied, no less legitimately, by another person. For this reason urban spaces are continuously partitioned, while these partitions form the basic mode of the distribution of social goods. Urban spaces are constituted by objective conflicts and oppositions, which account for their incessant dynamics. It is this facet of the spheres of human life that turns them into places of ceaseless rivalries, while their rules constitute the rudimentary forms of the social distribution of goods.

The City as a Spectacle

The above-outlined perspective suggests that urban spaces are *loci* of continuous and wholesale transformations of human life which both reflect and rebounds on the on-going transformations of the modes of production and consumption. Material production is now increasingly superseded by the production of immaterial goods; as a result, the productive and symbolic aspects of social life undergo processes of an unprecedented spectacularisation.[21] Through privileging the culture of visibility, the ubiquitous media of mass communication additionally boost the spectacularisation and aesthetisation of social relations. Due to the increasing spectacularisation of human life, the visual aspect of the city has become much more pronounced and stimulates now efforts to turn the cities themselves into spectacles.[22] The contemporary city functions today both as an aesthetic object and as a venue for the display and marketing of cultural goods and symbolic values.

Contemporary cities, both as exposition places and as spectacles, attract new inhabitants thus increasingly becoming cosmopolitan centers. There are several forms of contemporary urban cosmopolitanism, each of them accompanied by a characteristic aesthetics of its own.

One of them, a quotidian cosmopolitanism, is a result of the mixing of various ethnic and cultural groups which migrate to the cities in search of livelihood. This kind of cosmopolitanism is usually characterized by an aesthetics of diversity and it engenders urban areas of great diversity which, though inclusive, may also become arenas of major internal conflicts. Due to their specificity, such areas are usually remarkably distinct from neighboring districts. Their distinctiveness leads sometimes to their ghettoisation; they are also subject to political and aesthetic actions aimed at transforming them into "theme parks", which is often perceived by their inhabitants as segregatory, divisive and unjust treatment.

The second kind of cosmopolitanism is also quotidian in character; it is a result of expansive forms of the production of space and of their commodification; these transformations have engendered the phenomenon of the horizontal urban sprawl. It was made possible by the growth of motorization and has been followed by a specific suburban culture and aesthetics of uniformity, often represented in arts; it is also accompanied by the culture and aesthetics of mobility.[23] These transforma-

[21] G. Debord, *Society of the Spectacle* (1967), Detroit 1970.

[22] M. Meskimmon, *Engendering the City: Women Artists and Urban Spaces*, London 1997; D. Massey, *Space, Place and Gender*, Cambridge 1994.

[23] F. Houben, *From Centre to Periphery. The Aesthetics of Mobility*, [in:] E. Charlesworth, *City Edge. Case Studies in Contemporary Urbanism*, Oxford 2005, pp. 100–117.

tions in many urban areas, characterized by the absence of any strong communal relationships, have been aptly captured by the slogan *Anywhere City*. It is worth remarking that cities in Central and Eastern Europe are an interesting variation of this development. After the World War 2 their spaces have been filled with stereotypical apartment blocks serving to produce and concentrate the working classes, necessary for the modernization of the backward states. Construction of such housing facilities allowed those classes to enjoy comforts of a more civilized life, and contributed to the growth of the urban population. At the same time, however, this model of housing architecture has deeply affected their social relations by breaking up the traditional family ties and leading to the privatization of their life. The relative weakness of the economies of post-socialist cities makes it impossible for them to replace these buildings with more socially functional ones. As a result, in the foreseeable future, they will remain part of the social and urban landscape in this region. The recent housing boom in these countries has been shaped, in case of individual home builders, by an aesthetics of ostentation typical of the *nouveau riche*; it has also encouraged the construction of new urban areas whose chaos defies any aesthetical categorization. The social harm resulting from the desire to possess a house or an apartment without respect for the quality of social relations, their harmoniousness with the existing urban and natural landscape, and their aesthetic features, has been great and will be irreversible in the near future.

The third cosmopolitanism, an elite one, is exclusive and is accountable for the emergence of a specific urban aesthetics which organizes the servicing of the mobile elite whose expectations of state-of-the-art urban comforts affect significant, usually crucial elements of the urban structure.[24] An aesthetics of uniformity, accompanying this kind of cosmopolitanism, induces the production of identical and conventional spaces irrespective of their geographical location. The elite urban cosmopolitanism promotes also an aesthetics of ostentation aimed to offset the boredom of the aesthetics of uniformity and in this way to attract global attention to the cities. It finds an expression especially in the spectacular "signature" architecture, as in the widely discussed reconstruction of Paris by Georges Hausmann, or in the more recent example of Bilbao, now emulated with various success by other cities. This kind of urban aesthetics of ostentation has received an additional boost from the latest technological innovations which have made possible novel and often striking forms of the urban vertical sprawl, as in the case of spectacular architectural constructions in Shanghai, Dubai and numerous other cities vying for global recognition.

The Urban Aesthetic Politics

The transformation of ever new urban spaces into exposition places and venues of ostentation is also an incentive for the consumption of the symbolic and prestigious goods, and is a driving force behind the emergence of the urban creative classes, consisting of the workers of the sectors of knowledge, education, culture,

[24] J. Binnie *et al.*, *Cosmopolitan Urbanism*, Oxford 2006.

arts and business management.[25] This also propels the innovative development of arts and their instrumental use in enhancing the attractiveness of the cities. Contemporary cities, themselves works of various arts, are increasingly shedding their character of industrial centers and are becoming chiefly venues of production of symbolic goods. As such, they are even more prone to the processes of spectacularisation mentioned above.

The de-industrialization of contemporary cities is accompanied by wide-ranging changes in the socio-economic structure of urban centers, especially the structure of the working classes. This change is leads to claims about the demise and dispersal of the working class, which, in the industrial period, was concentrated in the cities. In view of the unprecedented growth of the world population, and the correlated demand for all sorts of goods, both material and non-material ones, such claims seem *prima facie* contestable. They appear to be sustained by an obsolete yet persistent political aesthetics in the understanding of both a worker and the goods produced. It is remarkable that the worker continues to be perceived as a man „with a hammer in his hand and coal dust in his pores",[26] dressed in overalls with rolled-up sleeves, while the product continues to be envisaged as a locomotive, an internal combustion engine, or a brick.[27] It thus seems that this antiquated political aesthetics, which lags behind the realities of continuously changing capitalist modes of production, is a major obstacle in addressing properly the present dynamics of urban populations. Together with the de-materialization of production, a corresponding change should take place in the political aesthetics of the working classes, while obsolete stereotypes haunting the analyses of changes in the urban social stratification need to be questioned.

In virtue of their nature, symbolic goods are less amenable to measurement than other kinds of goods. It does not prevent them, however, from becoming marketable commodities, as it never has. The production of spectacles in itself has become a subject of an intense commodification. In the culture of visibility spectacularisation has become a precondition of the commodification of symbolic goods which has now reached an unprecedented scale. Commodification and spectacularisation are thus obviously correlated: the culture of visibility makes the spectacles a much sought commodity, while the mass media make them easily marketable. As a result, spectacles are now becoming the chief commodity of contemporary culture, and their production is most rapidly growing industry. Consequently, access to them is increasingly regulated by the market mechanisms. Because of the unrelenting commodification of public spaces, which in Western societies increasingly determines the forms of participation in them, gaining access to these spaces, and functioning in them, involves an ever-increasing effort. Commodification generates the sense alienation of individuals in cities, and affects the presence and perception of symbolic values in urban spaces.

The perception of the city both as a work of art and as a cradle of arts has

[25] R.L. Florida, *Cities and the Creative Class*, London 2005; *The Rise of the Creative Class: and How It Is Transforming Work, Leisure, Community and Everyday Life*, New York 2002.

[26] N. Klein, *The Shock Doctrine: The Rise of Disaster Capitalism*, London 2007, p. 173.

[27] *Cf*. A. Chmielewski, '*Postmodernizm i jego wrogowie*', *Odra*, 6 (2013).

already received more than ample attention in various conceptions of political aesthetics. Some of them stress the fact that cities are places of a variety of forms of inequality which affect also the extent to which their citizens participate in cultural goods. The level of participation of individuals and groups, or classes, in cultural productions, embodied in education, arts and artistic culture, may be taken as a measure of the social access to symbolic values and of the degree of involvement in the creation of symbolic capital.[28] Yet there is something to be added to this, for these issues have also a direct political dimension. The unequal access to cultural goods is as a rule additionally aggravated by the mechanisms of municipal subsidies for cultural productions. Their beneficiaries tend to be members of the local elites who, in virtue of their education and material status, do not need any additional economic incentives to enjoy arts, whereas the majority of the citizenry, especially the economically disadvantaged or otherwise deprived, are not among their beneficiaries. In effect, a lion's share of the taxes, which for the most part come from contributions of the lowest income groups, is being redirected in order to provide a sophisticated entertainment for the elites at the costs of the excluded groups. This widespread mechanism increases the severity of exclusions and significantly contributes to iniquities in many cities. These consequences of the urban aesthetic politics are yet another way to justify the claim that political aesthetic considerations should be an important element in understanding many aspects of urban life.

Absence as Agoraphobia

One of many non-economic factors stimulating the growth of cities is a relative anonymity of life in urban spaces which offer a greater scope of freedom from repressive means of social control still persisting in traditional, more tightly-knit non-urban communities. Cities are thus not only opportunities to establish relationships with others. They also enable individuals to select them at will, avoid them while continuously enjoying the tantalizing possibility, real of virtual, of re-entering them, or shun them altogether. That is why cities are also places of privacy and seclusion, isolation and solitude. The specific nature of social relations made possible by the urban settings turns them also into objects of aesthetic enjoyment and voyeuristic consumption epitomized in the Baudelairesque *flâneur* attitude.[29]

The intensity of interaction and the level of sophistication – prerequisites of successful functioning in contemporary society, resulting from technological advancements and increased level of rivalries in the spaces of human life – generate an attitude which may be described as public agoraphobia. Public agoraphobia is, on the one hand, enforced by economic and social exclusions, and, on the other, is self-induced as a form of self-exclusion from the public life. These phenomena, as effects of the commodification and the agonistic character of the spaces of human

[28] P. Bourdieu, *Distinction: A Social Critique of the Judgment of Taste*, Cambridge 1984.

[29] These phenomena may be interpreted as forms of a more general category of urban absence as discussed by Richard Shusterman in his critical reading of Simmel's view of on the mental life in metropolis, *cf*. *Urbans Scenes and Unseens*, p. 177.

life, tend to amplify each other. Public agoraphobia adversely affects the political and social agency of individuals, engendering attitudes of the civic passivity in them.

Public agoraphobia is also fuelled, yet at the same time partly alleviated, by a phenomenon which may be described as interpassivity. In simplest terms: in order to be inter*active*, one is supposed to be personally involved in a social relationship; in order to be inter*passive*, one is supposed to substitute one's own activity with an object, or another person, within such a relationship. Cities are a particularly fertile ground for the development of interpassive attitudes. This is due to the growth of the division of labor, as well as creation, professionalization and specialization of services which are, obviously, available to greater extent within urban forms of collective life than within rural ones. As a result, in the urban environment individuals feel positively discouraged from performing many duties, tasks and chores which they would otherwise perform themselves, were it not for the easy availability of professionals ready to perform them instead of them. Through these processes mutual meaningful relations among individuals are increasingly mediated by expedient third parties. One has to emphasize that an essential part in this expediency lies in the fact that those intermediaries are themselves easily replaceable and conveniently disposable. Contemporary urban societies increasingly approximate the ideal type of "abstract societies" in which "men practically never meet face to face..., in which all business is conducted by individuals in isolation who communicate by typed letters or by telegrams, and who go about in closed motor-cars."[30] The interpassive attitude, together with public agoraphobia and commodification, mutually enhance each other and affect adversely both relationships among the individuals as well as their civic agency.

Politics and Ostentation

In conclusion it has to be observed that the aesthetics of ostentation, some forms of which were mentioned above, grows also on a fertile soil of local urban politics which, for the most part, has not been immune to the overwhelming processes of spectacularisation, much to the contrary. It is an outcome of decisions of the city managers who quite often treat their jobs as ego-rides and occasions to realize their ambitions by involving the cities they manage into spectacular but often miscalculated, ill-designed, excessively costly, and sometimes ruinous undertakings. Such ventures in ostentation adversely affect the stability of a growing number of cities across the globe. As a matter of fact the greatest current danger to the stability of the Chinese economy, and thus to the world economy as a whole, comes now from a huge debt incurred by innumerable Chinese municipalities, large and small which, restrained only by their ambitions and imagination, have become involved in misguided developments schemes. Rather curiously, in some cases it took the form of constructing large scale or even identical replicas of some European cities. As a result, Chinese municipalities are now responsible for 25 per cent

[30] K. Popper, *The Open Society and its Enemies*, London [1945] 1962, vol. I, p.174.

of their country's overall debt.[31] In view of this, a recent claim that mayors may be able to solve problems of the world economy better than nation states[32] seems grossly flawed for it is evident that mayors, instead solving the world's economic problems, are now actually busy creating new ones, even greater and less solvable than those actually at hand.

This kind of urban governance, which is by no means confined to the Chinese territories, is accompanied by various pathologies. One of them is rampant urban cronyism and clientelism of groups that stand behind the privatization of the urban public sphere and make up the core of the new urban patriciate. It is worth noticing that the class of new patricians is easily identified by an aesthetics of ostentation of its own.

More serious among these pathologies, however, is insufficient control of the municipal management by the urban democratic systems. For urban ostentation goes hand in hand with urban agoraphobia, the latter being a precondition of the former. Civic passivity of urban citizens, who forfeit their right to participation in the management of their cities, is no less damaging than adventurousness of their mayors. For through their withdrawal, they are becoming invisible and, thus, politically insignificant; through their public absence, they condone the transformation of urban forms of self-government into local despotisms; through their silence, they encourage the privatization of the traditional urban agora, contributing in this way to the atrophy of the public sphere in contemporary cities. Thus, if urban governance today resembles a set of crowd-management techniques rather than a joint democratic effort for the common good of the city inhabitants, it is due not only to the despotic leanings of the city managers, but also, to a greater extent, to the public desertion of the urban citizens.

These developments, which are both a cause and a consequence of prevalent public agoraphobia, are a reason why Socratic perambulations through the city streets in search of people to learn from, being now confined to shrinking urban enclaves, have been all but replaced by a detached *flaneurish* consumption of their appearances.

[31] The debt of Chinese municipalities is estimated at 13 trillion yuan (ca US$ 2,22 trillion).

[32] *Cf*. B. Barber, *If Mayors Ruled the World: Dysfunctional Nations, Rising Cities*, New Haven–London 2013. Barber writes: "Cities are increasingly networked into webs of culture, commerce and communication that encircle the globe. These networks and the cooperative complexes they embody can be helped to do formally what they now do informally: govern through voluntary cooperation and shared consensus. If mayors ruled the world, the more than 3,5 billion people (over half of the world's population) who are urban dwellers and many more in the exurban neighbourhoods beyond could participate locally and globally at the same time – a miracle of civil 'glocality' promising pragmatism instead of politics, innovation rather than ideology, solutions in place of sovereignty" (p. 5).